

TU WIEN

TECHNISCHE PHYSIK

---

EINFÜHRUNG IN DIE ALLGEMEINE RELATIVITÄTSTHEORIE

---

*Autor*

Alexander MIETHLINGER

29. Februar 2024

# Inhaltsverzeichnis

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Einleitung</b>                                       | <b>2</b>  |
| <b>2</b> | <b>Begriffe der SRT</b>                                 | <b>3</b>  |
| 2.1      | Die Raumzeit . . . . .                                  | 3         |
| 2.2      | Gleichzeitigkeit und Längenkontraktion . . . . .        | 4         |
| 2.3      | Die Lorentztransformation . . . . .                     | 7         |
| <b>3</b> | <b>Multilineare Algebra</b>                             | <b>11</b> |
| 3.1      | Vektorräume . . . . .                                   | 11        |
| 3.2      | Der Dualraum . . . . .                                  | 14        |
| 3.3      | Multilineare Funktionen . . . . .                       | 17        |
| 3.4      | Der Doppel-Dual-Raum . . . . .                          | 19        |
| 3.5      | Tensoralgebra . . . . .                                 | 20        |
| <b>4</b> | <b>Topologien und topologische Räume</b>                | <b>25</b> |
| 4.1      | Begriffe der Topologie . . . . .                        | 25        |
| 4.2      | Abbildungen auf topologischen Räumen . . . . .          | 26        |
| 4.3      | Trennungsaxiome . . . . .                               | 28        |
| 4.4      | Mannigfaltigkeiten . . . . .                            | 31        |
| 4.4.1    | Vektorfelder auf Mannigfaltigkeiten . . . . .           | 37        |
| 4.4.2    | Parallelverschiebung und kovariante Ableitung . . . . . | 44        |
| <b>5</b> | <b>Einstein'sche-Feldgleichungen</b>                    | <b>52</b> |
| 5.1      | Autoparallele Kurven . . . . .                          | 52        |
| 5.2      | Lie-Ableitung . . . . .                                 | 55        |
| 5.3      | Der Einstein-Tensor . . . . .                           | 59        |
| <b>6</b> | <b>Schluss</b>  | <b>63</b> |

# Kapitel 1

## Einleitung

Dieses Skriptum ist eine Zusammenfassung des Stoffes der gleichnamigen Lehrveranstaltung an der Technischen Universität Wien. Ziel ist es, dem Leser einen Einblick in die Allgemeine Relativitätstheorie zu geben, mit Betonung auf die mathematischen Hintergründe.

Im ersten Teil des Skriptums wird die Spezielle Relativitätstheorie nochmals wiederholt<sup>1</sup>, um dem Leser einige Begrifflichkeiten, sowie Überlegungen darzubringen, welche für die spätere Auseinandersetzung mit der Allgemeinen Relativitätstheorie wichtig sind.

Danach wird das mathematische Fundament gelegt, um die Gleichungen der ART zu verstehen und anwenden zu können.

Es soll hier noch gewarnt werden: Dieses Skriptum wurde von einem Studenten erstellt, der sich mit dem zentralen Thema der Lehrveranstaltung noch nie auseinandergesetzt hat. Deshalb fühle ich mich verpflichtet dem Leser hier einige seriöse Quellen zu empfehlen:

1. *General Relativity* (Robert Wald): Dieses Buch wird von Prof. Balasin persönlich auf der TISS-Seite der Lehrveranstaltung empfohlen.
2. *General Relativity: The Theoretical Minimum* (Leonard Susskind): Dieses Buch gibt einen sanften Einstieg ins Thema, kann aber natürlich nicht mit einem Lehrbuch für Bachelor-Studenten mithalten. Ebenso kann ich die online Vorlesungen von Prof. Susskind auf Youtube sehr empfehlen.
3. Im letzten Kapitel wird etwas Elektrodynamik vorkommen, was nicht sonderlich relevant für das Verständnis der Feldgleichungen ist, aber es hilft dennoch sich vertiefend hineinzulesen. Dafür wird auf das Skriptum der LVA „Elektrodynamik I“ und auf den dritten Teil von *The Theoretical Minimum: Special Relativity and classical field theory* verwiesen.

Falls Fehler oder Unklares aufkommt, bitte an [a.miethlinger@gmail.com](mailto:a.miethlinger@gmail.com) melden! Ich wünsche der Leserschaft hiermit viel Spaß mit der Allgemeinen Relativitätstheorie.

---

<sup>1</sup>Der Stoff der LVA beginnt aber erst ab Kapitel 2.3!

# Kapitel 2

## Begriffe der SRT

### 2.1 Die Raumzeit

Die Darstellung der 4-dimensionalen Raumzeit ist an sich schwierig, doch man kann sie limitiert visualisieren. Man stelle sich den  $\mathbb{R}^3$ , wobei wir die x-y Ebene als den „wirklichen“  $\mathbb{R}^3$  identifizieren werden und die z-Achse als unsere Zeitachse t. Nun kann man sich klarmachen, dass jede Ebene bei  $t = \text{const}$  einen Moment in der Raumzeit darstellt. Ähnlich symbolisieren Geraden, die im Winkel  $90^\circ$  zu solch einer Zeitpunktsebene stehen, einen Punkt im Raum, zu jedem Zeitpunkt. Wir nennen einen Punkt, welcher die 4-dimensionalen Raumzeit-Koordinaten  $\mathbf{x}^\mu = (ct, x, y, z)^T$  hat, ein *Ereignis*.

In Abbildung 2.1 sieht man eine solche Visualisierung. Die Ebenen zeigen Ereignisse mit gleicher  $x^0$ -Koordinate an ( $t = \text{const}$ ). Die Punkte A, B und C sind solche Ereignisse die in einer Zeitpunktsebene liegen. Durch B geht eine Gerade, das bedeutet, dass ein Körper sich in der Raumzeit bewegt und zum Zeitpunkt  $t = 0$  am Punkt B gewesen ist. Der Winkel  $\alpha = 68.45$  zwischen der Gerade und der Ebene  $t = 0$  spiegelt die Geschwindigkeit des Körpers relativ zu einem ruhenden System an, je größer der Winkel, desto schneller bewegt sich der Körper. Es gilt nämlich für den Winkel und der Geschwindigkeit eines Körpers:  $\tan(\alpha) = \frac{x}{ct} = \frac{v}{c}$ . Man kann erkennen, dass ein Winkel von  $0^\circ$  einer Geschwindigkeit von 0 m/s entspricht, deshalb bewegt sich der Körper von C nach G *räumlich* nicht, aber in der Raumzeit schon, nämlich auf einer Geraden, die wir ab sofort *Weltlinien* nennen werden, mit einem Winkel von  $90^\circ$  zu einer Zeitpunktsebene.

Beim Punkt A wurde ein Kegel mit Öffnungswinkel  $90^\circ$  eingezeichnet, also einem Winkel  $\alpha = 45$ . Deshalb können wir die Geschwindigkeit mit der sich dieses *Ding* bewegt mit  $v = c$  identifizieren. Wir werden später sehen, dass das einzige was sich mit Lichtgeschwindigkeit fortbewegen kann, nur Licht selbst ist. Daraus folgt, dass dieser *Lichtkegel* alle Richtungen in der Raumzeit anzeigt, in welche sich ein Lichtblitz, der bei A entstanden ist, ausbreiten kann. D.h. alle Weltlinien innerhalb eines Lichtkegels bewegen sich mit  $v < c$ , alle auf dem Lichtkegel mit  $v = c$ .

Ein großer Unterschied zwischen der klassischen Mechanik und der Relativitätstheorie sind die

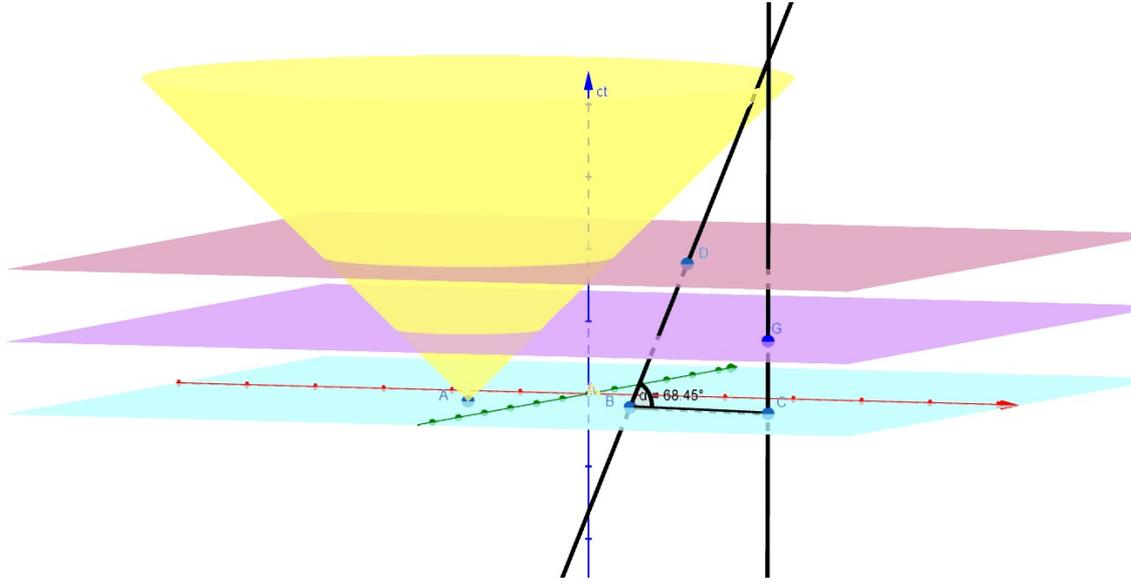


Abbildung 2.1: Visualisierung der Raumzeit mit Zeitpunktebenen und Ereignissen.

maximalen Geschwindigkeiten, mit welcher sich ein Körper fortbewegen kann. Bei Newton gibt es tatsächlich eine maximale Geschwindigkeit, diese beträgt  $\infty$ , was einem Öffnungswinkel einer Weltlinie von  $90^\circ$  entspricht. Bei Einstein ist sie die Lichtgeschwindigkeit  $c$ .

Zwei Begriffe müssen hier noch ergänzt werden: Die invariante Geschwindigkeit und die Grenzgeschwindigkeit. Die invariante Geschwindigkeit ist in allen Bezugssystemen gleich, wobei die Grenzgeschwindigkeit von keinem Körper überschritten werden kann. In der Speziellen Relativitätstheorie sind diese Geschwindigkeiten *gleich*.

## 2.2 Gleichzeitigkeit und Längenkontraktion

Die Spezielle Relativitätstheorie von Einstein wirft den Begriff der *Gleichzeitigkeit* aus dem Fenster, sie wird nur relativ zum Bezugssystem wirksam. Man betrachte ein Bezugssystem  $S$ , welches für uns ruhend ist, und ein bewegtes System  $S'$ , welches sich mit konstanter Geschwindigkeit  $v$  bewegt. Nun kann man ein *Minkowski-Diagramm* erstellen, in dem System  $S$  und  $S'$  eingezeichnet werden, siehe Abbildung 2.2.

Als Beispiel stelle man sich Albert Einstein auf einem Zug und Isaac Newton am Bahnhofsgleis vor. Hierbei ist nun Einstein unser bewegtes System und Newton unser ruhendes. Am Bahnhofsgleis steht eine Uhr und beide wollen wissen, ob ihre Uhr denn überhaupt richtig läuft, deshalb wollen sie mit ihrem universellen Zeitgefühl die Zeit stoppen, um zu sehen, ob sie falsch läuft<sup>1</sup>. Der Zug fährt

<sup>1</sup>Obwohl beide geniale Physiker waren, fällt ihnen anscheinend nicht auf, dass man so keine Uhr auf ihre Richtigkeit überprüfen kann.

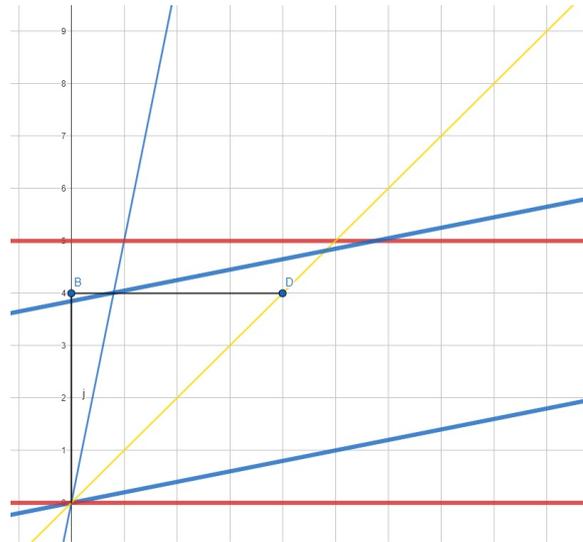


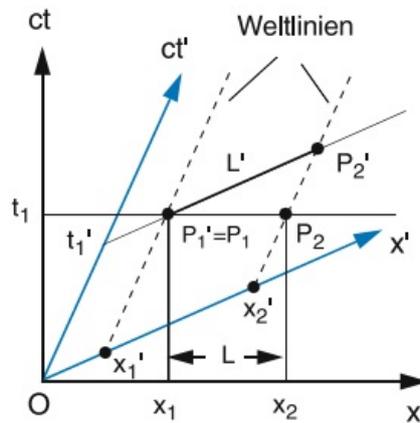
Abbildung 2.2: Gleichzeitigkeit hängt vom Bezugssystem ab

los und beide schauen auf den Sekundenzeiger ihrer Uhr. Sie haben sich ausgemacht, dass nach 4 Sekunden einer „Stopp!“ schreit.

Es mag überraschen, aber Newton sagt als Erster stopp, noch vor Einstein. Dieser aber sagt, Newton habe zu früh gestoppt, für Einstein sind noch keine 4 Sekunden vergangen. Tja, wer von den beiden hat nun Recht? Die Antwort ist, dass beide recht haben. Newton, unser ruhendes System, misst die Zeit zum Zeitpunkt B, für ihn nach  $\Delta t = 4s$  Sekunden. Einstein wiederum ist in einem bewegten System, welches in Abbildung 2.2 durch die blauen Linien gekennzeichnet wurde. Für Einstein vergeht, **relativ zu Newton weniger** Zeit. Equivalent wäre es zu sagen, dass für Einstein die Uhr *langsamer* ist, relativ zu Newton. Für Einstein verging eine Zeit  $\Delta t' < 4s$ , da seine Zeit, **relativ zu Newton** gestreckt wurde, eine Sekunde dauert bei Einstein länger als bei Newton. Der Leser soll über diese Überlegung kurz nochmal reflektieren, denn hier steckt viel Physik dahinter<sup>2</sup>. Als Abschluss soll dieser Absatz nochmals zusammengefasst werden: *Bewegte Uhren laufen langsamer.*

Da die beiden nach ihrer Zankerei über die Uhren feststellten, dass etwas nicht stimmt, wollen sie nun etwas anderes ausprobieren. Sie wollen überprüfen ob ihr Meterstab wirklich 1 m beträgt. Wieder, Einstein steigt in den Zug und fährt los. In Abbildung 2.3 ist das Minkowski-Diagramm dieser Anordnung aus der Sicht Einsteins dargestellt. Bei den Ereignissen  $\mathbf{x}_1 = (t_1, x_1)$  und  $\mathbf{x}_2 = (t_1, x_2)$  wird der Meterstab von Newton gemessen, für ihn *gleichzeitig*. Einstein fährt vorbei und beginnt auch zum Zeitpunkt  $t_1$  zu messen, aber Newton sieht, dass Einstein seinen Rollmeter zu früh wieder ablegte. Für Newton dauerte Einsteins Messung  $\Delta t$  Sekunden. Später besprechen sich die beiden: Einstein sagt, dass ihr Meterstab keineswegs 1 m betragen könnte, er sei zu kurz dafür. Newton sagt, es sei sehrwohl 1 m lang und Einstein irre sich. Nun, wer hat dieses Mal Recht?

<sup>2</sup>Der Autor musste auch darüber erneut nachdenken.



■ **Abb. 3.26** Graphische Darstellung der Lorentz-Kontraktion im Minkowski-Diagramm für einen im bewegten System  $S'$  ruhenden Längenmaßstab  $L'$

Abbildung 2.3: Räumliche Länge hängt vom Bezugssystem ab; Demtröder 1

Wieder haben beide, relativ zu ihrem Bezugssystem Recht. Für Einstein bewegt sich der Meterstab, wodurch seine Längenskala, **relativ zu Newton** verkürzt wird. Wenn Einstein nun einen Rollmeter zur Messung nimmt (welcher sich relativ zu ihm nicht bewegt), erscheint der Meterstab *kürzer*. Für Newton ist Einsteins Rollmeter kürzer, wodurch er ein „falsches“ Ergebnis erhält. Auch hier soll der Leser wieder darüber nachdenken, was eigentlich passiert ist. Als Abschluss soll dieser Absatz nochmals zusammengefasst werden: *Für Bewegte Körper werden Längen kürzer.*

Nach diesen kleinen Gedankenexperimenten soll nun die Transformation von einem Bezugssystem ins andere hergeleitet werden, um die einzelnen Aussagen zu verifizieren.

## 2.3 Die Lorentztransformation

Ausgehend des Relativitätsprinzips „Jedes Inertialsystem ist gleichwertig“ können wir schon einige Eigenschaften der gesuchten Transformation angeben:

1. Homogenität
2. Isotropie (keine ausgezeichneten Richtungen)
3. Universalität der Transformation
4. Die Transformation darf nur von der *Relativgeschwindigkeit* abhängen

Der allgemeinste Ansatz für solch eine Transformation für die einzelnen Raumzeit-Koordinaten des Systems  $\mathbf{x}^\mu$  in ein System  $\mathbf{x}'^\mu$  muss folgendermaßen aussehen:

$$t' = a(v)t + b(v)v^j x^j \quad (2.1)$$

$$x'^i = d(v)v^i t + e(v)x^j + f(v)v^i(v^j x^j) \quad (2.2)$$

wobei die Einstein'sche Summenkonvention verwendet wird.  $v$  ist der Betrag der Relativgeschwindigkeit  $v = \sqrt{v^j v^j}$ . In Matrixschreibweise würden die Ausdrücke oben so ausschauen:

$$\begin{pmatrix} t' \\ x'^i \end{pmatrix} = \begin{pmatrix} a(v) & b(v)v^j \\ d(v)v^i & e(v)\delta^{ij} + f(v)v^i v^j \end{pmatrix} \begin{pmatrix} t \\ x^i \end{pmatrix} \quad (2.3)$$

Im Folgenden werden wir uns aus reinen physikalischen Eigenschaften, die diese Transformation haben *muss*, ihre Form herleiten. Dafür können wir drei Eigenschaften identifizieren:

1. Die Inverse angewandt auf die Transformation muss die Einheitsmatrix  $\mathbb{I}$  ergeben.
2. Die Inverse muss der Transformation mit negativer Relativgeschwindigkeit entsprechen;  $A^{-1}(v) = A(-v)$ .
3. Die Anwendung der Transformation von einem Bezugssystem mit Geschwindigkeit  $v$  in ein Bezugssystem  $v$  soll die räumlichen Koordinaten nicht verändern.
4. Transformiert man von ein System  $I$  in ein System  $I'$  mit  $v$  und vom System  $I'$  in ein System  $I''$  mit Relativgeschwindigkeit  $w$ , dann soll es eine Transformation von  $I$  nach  $I''$  mit einer (unbekannten) Geschwindigkeit  $u$  geben.

Die ersten beiden Bedingungen an die Transformation kann man in einer Formel darstellen:

$$A(-v)A(v) = \mathbb{I} \quad (2.4)$$

$$\begin{pmatrix} a(v) & -b(v)v^j \\ -d(v)v^i & e(v)\delta^{ij} + f(v)v^i v^j \end{pmatrix} \cdot \begin{pmatrix} a(v) & b(v)v^k \\ d(v)v^j & e(v)\delta^{jk} + f(v)v^j v^k \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \delta^{ik} \end{pmatrix} \quad (2.5)$$

Daraus können die fünf Gleichungen aufgestellt werden:

$$\begin{aligned} \text{I: } & a^2 - bdv^2 = 1 \\ \text{II: } & ab - be - bfv^2 = 0 \\ \text{III: } & -ad + ed + fdv^2 = 0 \\ \text{IV: } & -bdv^i v^k + e^2 \delta^{ik} + f^2 v^2 v^i v^k + 2efv^i v^k = \delta^{ik} \\ \text{V: } & -bd + f^2 v^2 + 2ef = 0 \end{aligned}$$

Aus der Gleichung **IV** wurde ein Koeffizientenvergleich ( $\delta^{ik}$ ) gemacht um Gleichung **V** zu erhalten. Nun können wir  $e^2 = 1$  identifizieren, also  $e(v) = \pm 1$ . Da bei  $A(0) = \mathbb{I}$  gelten soll, ist  $e(0) = 1$  und somit  $e(v) = 1$ . Damit haben wir unsere erste Unbekannte erhalten! Die Gleichungen vereinfachen sich nun zu:

$$\begin{aligned} \text{I: } & a^2 - bdv^2 = 1 \\ \text{II: } & b(a - 1 - fv^2) = 0 \\ \text{III: } & d(-a + 1 + fv^2) = 0 \\ \text{V: } & -bd + f^2 v^2 + 2ef = 0 \end{aligned}$$

Aus Gleichungen **II** und **III** kann die Beziehung  $f = \frac{a-1}{v^2}$  gewonnen werden. Mit der Bedingung drei kann folgende Gleichung aufgestellt werden:

$$\begin{pmatrix} a(v) & -b(v)v^j \\ -d(v)v^i & \delta^{ij} + f(v)v^i v^j \end{pmatrix} \cdot \begin{pmatrix} t \\ tv^i \end{pmatrix} = \begin{pmatrix} * \\ 0^i \end{pmatrix} \quad (2.6)$$

wobei  $x^i$  mit  $tv^i$  ersetzt wurde. Mit der neuen Beziehung zwischen f und a können wir aus diesem Gleichungssystem eine weitere Beziehung gewinnen:

$$dv^i t + tv^i f v^2 v^i t = 0 \quad (2.7)$$

$$d + 1 + a - 1 = 0 \quad (2.8)$$

$$d = -a \quad (2.9)$$

Wobei wir einen Koeffizientenvergleich für  $v^i$  und dann die Beziehung  $f = \frac{a-1}{v^2}$  eingesetzt haben. **I** kann nun auf b umgeformt werden und man erhält  $b = \frac{a^2-1}{-av^2}$ .

Mit der Bedingung vier kann folgende Gleichung aufgestellt werden:

$$A(w) \cdot A(v) = A(u) \quad (2.10)$$

$$\begin{pmatrix} a(w) & \frac{a^2(w)-1}{-a(w)w} e^i \\ -a(w)we^k & \delta^{ki} + (a(w)-1)e^k e^i \end{pmatrix} \cdot \begin{pmatrix} a(v) & \frac{a^2(v)-1}{-a(v)v} e^j \\ -a(v)ve^i & \delta^{ij} + (a(v)-1)e^i e^j \end{pmatrix} = A(u) \quad (2.11)$$

$$\text{I: } a(w)a(v) + \frac{a^2(w)-1}{a(w)w} a(v)v = a(u)$$

$$\text{II: } a(w) \frac{a^2(v)-1}{-a(v)v} e^j + \frac{a^2(w)-1}{-a(w)w} (e^j + (a(v)-1)e^j) = \frac{a^2(u)-1}{-a(u)u} e^j$$

$$\text{III: } -a(v)a(w)we^k - a(v)v(e^k + (a(w)-1)e^k) = -a(u)ue^k$$

IV:

$$\frac{a^2(v)-1}{a(v)v} a(w)we^k e^j + \delta^{kj} + e^k e^j (a(w)-1)(a(v)-1) + (a(v)-1)(a(w)-1)e^k e^j = \delta^{ik} + (a(u)-1)e^i e^k$$

Für die weitere Rechnung brauchen wir nur Gleichungen I und IV denn mit einem Koeffizientenvergleich bei IV ( $e^i e^k$ ) kann man erhalten:

$$\frac{a^2(w)-1}{a(w)w^2} = \frac{a^2(v)-1}{a(v)v^2} \quad (2.12)$$

Und diese Gleichung ist nur erfüllt wenn beide Seiten einer Konstanten  $k$  entsprechen. Dadurch kann unsere letzte Unbekannte  $a$  durch eine (noch) unbekannte  $k$  ausgedrückt werden:

$$a(v) = \sqrt{\frac{1}{1-kv^2}}. \quad (2.13)$$

Somit können wir nun die Lorentztransformation anschreiben:

$$A(v) = \begin{pmatrix} a & -kav^j \\ -av^i & \delta^{ij} + \frac{a^2-1}{v^2} v^i v^j \end{pmatrix}. \quad (2.14)$$

Zur Herleitung der Minkowski-Metrik müssen wir unsere Zeitkomponente mit unseren räumlichen Komponenten der Raumzeit gleichsetzen,  $t^2 = x^i x^i$  <sup>3</sup>. Diese Gleichung beschreibt jene Punkte, welche sich *auf* dem Lichtkegel befinden. Da das aber laut dem Relativitätsprinzips in jedem Koordinatensystem gelten sollte, können wir auch  $t'^2 = x'^i x'^i$  schreiben und unsere Transformation benutzen <sup>4</sup>:

$$t' = at - ka(vx) \quad (2.15)$$

$$x'^i = -av^i t + x^i \frac{a-1}{v^2} v^i(vx), \quad (2.16)$$

wobei  $v = \sqrt{v^k v^k}$  und  $x = \sqrt{x^k x^k}$  gilt. Wenn wir nun die Gleichungen quadrieren erhalten wir:

$$t'^2 = a^2 t^2 + k^2 a^2 (vx)^2 - 2a^2 k (vx)t \quad (2.17)$$

<sup>3</sup>Beachte, dass  $t = ct$  gilt!

<sup>4</sup>Was wir hier tun ist folgendes: Wir sagen, dass jeder Punkt auf dem Lichtkegel in einem Bezugssystem ein Punkt auf dem Lichtkegel in einem beliebigen anderen Bezugssystem ist. Kurzgesagt: Unter Lorentztransformation bleiben lichtartige Punkte lichtartige Punkte.

$$\begin{aligned}
x'^i x'^i &= av^2 t^2 + x^i x^i + \frac{a-1}{v^2} (vx)^2 - 2a(vx)t - 2a(a-1)(vx)t \\
&\quad + 2\frac{a-1}{v^2} (vx)^2
\end{aligned} \tag{2.18}$$

Durch Koeffizientenvergleich bei  $t'^2$  erhalten wir:

$$-a^2 + a^2 v^2 + 1 = 0, \tag{2.19}$$

$$a^2(1 - v^2) = 1. \tag{2.20}$$

Wenn wir nun zurück zur Herleitung der Lorentztransformation schauen, erkennen wir, dass für lichtartige Punkte die Konstante  $k = 1$  gilt. Doch mithilfe der Lorentztransformation kann man noch allgemeinere Aussagen erzielen. Setzen wir mal im Ausdruck  $-t'^2 + x'^i x'^i$  die Lorentztransformation ein, bzw. wir schauen wie dieser Ausdruck im Bezugssystem  $\mathbf{x}^\mu$  aussieht:

$$-t'^2 + x'^2 = -(at - a(vx))^2 + (-av^i t + x^i + \frac{a-1}{v^2} v^i (vx))^2 \tag{2.21}$$

$$\begin{aligned}
-t'^2 + x'^2 &= -at^2 + 2a^2(vx)t - a^2(vx)^2 + a^2 v^2 t^2 + x^2 + \frac{a-1}{v^2} (vx)^2 \\
&\quad - 2a(vx)t - 2a(a-1)(vx)t + 2\frac{a-1}{v^2} (vx)^2
\end{aligned} \tag{2.22}$$

$$-t'^2 + x'^2 = -t^2(a^2 - a^2 v^2) + x^2 + 2(vx)t(a^2 - a - a(a-1)) \tag{2.23}$$

$$-t'^2 + x'^2 = -t^2 + x^2 \tag{2.24}$$

Man sieht nun, dass dieser Ausdruck anscheinend *invariant* bzgl. dem Bezugssystem ist! Später werden wir diesen Term als „Raumzeit Intervall“ bzw. als „Eigenzeit“ bezeichnen. Mithilfe einer Matrix können wir den Term etwas knapper anschreiben:

$$-t^2 + x^2 = \eta_{\mu\nu} x^\mu x^\nu \tag{2.25}$$

Dieses  $\eta_{\mu\nu}$  ist die sogenannte *Minkowski-Metrik*. In der Speziellen Relativitätstheorie hat sie nur den Nutzen einer einfacheren Notation, aber hier in der Allgemeinen Relativitätstheorie werden wir sehen, dass ihr ein tiefer geometrischer Zusammenhang zur Gravitation und Krümmung der Raumzeit zugrunde liegt.

# Kapitel 3

## Multilineare Algebra

### 3.1 Vektorräume

Ein Vektorraum ist eine Menge bestehend aus Vektoren und einem Körper. Der  $\mathbb{R}^3$  wäre ein Vektorraum aller 3er Tupel von reellen Zahlen. Die 3er Tupel sind die Vektoren und die reellen Zahlen bilden den Körper. Formal werden Vektorräume so angeschrieben:  $(V, \Gamma)$ , wobei  $V$  die Menge der Vektoren und  $\Gamma$  den Körper darstellt. Wir können nun für die einzelnen Mengen der Vektoren und des Körpers Operatoren für Vektoren  $v, w \in V$  und  $\lambda, \mu \in \Gamma$  einführen:

1.  $+$ :  $V \times V \rightarrow V$

(a)  $(v, w) = v + w = w + v$

(b)  $(0, v) = 0 + v = v$

(c)  $(-v, v) = -v + v = 0$

2.  $\cdot$ :  $\Gamma \times V \rightarrow V$

(a)  $(\lambda, v) = \lambda \cdot v$

(b)  $(0, v) = 0 \cdot v = 0$

(c)  $(1, v) = 1 \cdot v = v$

3.  $+$ :  $\Gamma \rightarrow \Gamma$

(a)  $(\lambda, \mu) = \lambda + \mu$ , mit 0 als neutrales Element

4.  $\cdot$ :  $\Gamma \rightarrow \Gamma$

(a)  $(\lambda, \mu) = \lambda \cdot \mu$ , mit 1 als neutrales Element

Wir haben die Operatoren paarweise nicht umsonst mit dem gleichen Zeichen versehen, denn wir „fordern?“ nun, dass diese Operationen Gleichheit haben:

1.  $\lambda \cdot (v+w) = \lambda v + \lambda w$
2.  $(\lambda + \mu)v = \lambda v + \mu v$
3.  $(\lambda \cdot \mu)v = \lambda \cdot (\mu \cdot v) = \mu \cdot (\lambda \cdot v)$

Mit diesen Werkzeugen können wir nun wichtige Erkenntnisse über unseren (beliebigen) Vektorraum erhalten. Sei die Menge  $(v_\alpha)_{\alpha \in I} \subseteq V$ , also eine Menge bestehend aus den Vektoren  $v_\alpha$ , wobei der Index  $\alpha$  eine Nummerierung darstellt. Wir werden die Summe  $\lambda^\alpha v_\alpha \in V$  als *Linearkombination* bezeichnen. Hierbei sind die  $\lambda^\alpha$  beliebige Elemente aus  $\Gamma$ , also sind sie auch eine Teilmenge von  $\Gamma$ :  $(\lambda^\alpha)_{\alpha \in I} \subseteq \Gamma$ . Wir fordern übrigens, dass die *Koeffizienten*,  $\lambda^\alpha$ , für endlich viel Elemente ungleich voneinander sind. Warum hier bei den Zahlen erhöhte Indizes und bei den Vektoren erniedrigte Indizes gewählt wurden, wird später erläutert.

Nun kommt ein Begriff, der uns aus der Linearen Algebra sehr bekannt ist: *lineare Unabhängigkeit*. Wir fordern nun, dass unsere Teilmenge des Vektorraums  $V$   $(v_\alpha)_{\alpha \in I}$  linear unabhängig ist. D.h. es gilt  $\lambda^\alpha v_\alpha = 0 \Rightarrow \lambda^\alpha = 0 \forall \alpha \in I$ . Wir können nun beliebige Vektoren mithilfe der Linearkombination darstellen, denn diese ist *eindeutig*: Sei  $v \in V$  und  $v = \lambda^\alpha v_\alpha$ , aber auch  $v = \mu^\alpha v_\alpha$ . Aus  $v - v = 0$ , kann man durch einsetzen folgern, dass  $\lambda^\alpha v_\alpha - \mu^\alpha v_\alpha = (\lambda^\alpha - \mu^\alpha)v_\alpha = \nu^\alpha v_\alpha = 0$ . Da die Benennung der Koeffizienten irrelevant ist, muss  $\nu^\alpha = 0 \forall \alpha \in I$ . Das muss aber heißen, dass  $\lambda^\alpha = \mu^\alpha \forall \alpha \in I$ , deshalb ist die Darstellung eines Vektors in einer linear unabhängigen Linearkombination eindeutig.

Nun kann die Frage gestellt werden: Gibt es denn eine maximale linear unabhängige Teilmenge  $M = (v_\alpha)_{\alpha \in I} \subseteq V$ . Sei  $v \in V$ , dann gilt  $v \in M \vee v \notin M$ . Gilt  $v \in M$ , dann gilt offensichtlich  $v = \lambda^\alpha v_\alpha$ , da  $v$  selbst eines der  $v_\alpha$  ist. Dann kann man schreiben  $v = \lambda^{\alpha_0} v_{\alpha_0}$ ,  $\lambda^{\alpha_0} = 1$ . Dann gilt  $\lambda^\alpha = 0 \forall \alpha \neq \alpha_0$ . Der interessantere Teil ist wenn  $v \notin M$  gilt. Da wir angenommen haben,  $M$  sei linear unabhängig, können wir annehmen, dass  $v \cup M$  nicht linear unabhängig ist, da wir sonst  $v$  zu  $M$  hinzufügen können. Da sie nicht linear unabhängig sind, gilt für die Summe  $\lambda v + \lambda^\alpha v_\alpha = 0$  folgendes: Wenn  $\lambda = 0$  gilt, haben wir wieder die Definition einer linear unabhängigen Linearkombination. Deshalb fordern wir, dass  $\lambda \neq 0$  gilt, dann können wir die Summe auf  $v$  umformen und erhalten:  $v = -\frac{\lambda^\alpha}{\lambda} v_\alpha = \mu^\alpha v_\alpha$ , was wiederum eine eindeutige Linearkombination ist. Wenn wir jeden Vektor im Vektorraum  $V$  durch eine Linearkombination darstellen können, dann nennen wir die Menge der linear unabhängigen Vektoren eine *Basis*. Also kann man auch fragen, gibt es solch eine Basis? Die Antwort ist *ja* und für ihre Begründung müssen wir in die Mengenlehre gehen.

Sei  $(A, \leq)$  eine *teilgeordnete* Menge. D.h. sie hat folgende Eigenschaften;

1.  $a \leq a \forall a \in A$
2.  $a \leq b \wedge b \leq a \rightarrow a = b \forall a, b \in A$
3.  $a \leq b \wedge b \leq c \rightarrow a \leq c \forall a, b, c \in A$

Wir sehen, dass die Elemente in der Menge  $A$  mit diesen Eigenschaften geordnet werden können. Mit dieser Definition können wir weiterarbeiten.

Sei  $M$  eine Menge. Dann ist  $P(M)$  die *Potenzmenge* von  $M$ , d.h. sie beinhaltet alle Teilmengen von  $M$ . Hier ein Beispiel: Sei  $M = \{1, 2, 3\}$ , also eine Menge mit den drei Elementen 1, 2 und 3. Die Potenzmenge würde nun so ausschauen:  $P(M) = \{\{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}, \{\emptyset\}\}$ , wobei wir fordern, dass die leere Menge  $\emptyset$  auch ein Element der Potenzmenge ist. Mathematisch kann man die Potenzmenge so anschreiben:  $P(M) = \{U \mid U \subseteq M\}$ .

Nun brauchen wir nur noch einen Begriff, und zwar das *Lemma von Zorn*. Das Lemma von Zorn besagt, dass wenn in einer teilgeordneten Menge  $(A, \leq)$  jede Kette eine obere Schranke besitzt, dann existiert ein maximales Element  $K$  in  $A$ . Was ist nun eine Kette? Eine Kette  $T$  ist eine *totalgeordnete* Teilmenge von  $A$ , d.h. sie erfüllt alle Eigenschaften der teilgeordneten Menge  $A$ , nur mit einer zusätzlichen Eigenschaft:  $a \leq b \vee b \leq a, \forall a, b \in T$ , d.h. für alle  $a$  und  $b$  gilt stets nur, dass eines größer gleich als das andere ist. Nun fordern wir für das Lemma von Zorn, dass diese Ketten eine obere Schranke besitzen, d.h. Für jede Kette  $T$  in  $A$  existiert ein  $s \in A$ , sodass  $t \leq s \forall t \in T$  gilt. Also bedeutet das, dass es ein Element  $s$  in unserer teilgeordneten Menge  $A$  gibt, welches größer ist als alle Elemente der Kette  $T$ . Nun folgt das Lemma von Zorn: In  $A$  gibt es nun ein maximales Element  $m$ , wofür es kein größeres Element in  $A$  existiert. Wenden wir nun dieses Lemma auf unseren Vektorraum an, um zu verifizieren, dass es eine Basis dafür gibt.

Sei  $(\mathbb{L}, \leq)$  eine Teilgeordnete Menge für die gilt:  $\mathbb{L} = \{L \subseteq V \mid L \text{ ist lin. unab.}\}$ .  $\mathbb{L}$  ist also eine Menge bestehend aus Mengen von Vektoren unseres Vektorraums  $V$ , die linear unabhängig sind. Elemente von  $\mathbb{L}$  wären z.B.  $L_1$  und  $L_2$ . Für diese Elemente gilt  $L_1 \leq L_2 \Leftrightarrow L_1 \subseteq L_2$ . Sei nun  $\mathbb{K}$  eine Kette auf  $\mathbb{L}$ , d.h.  $\forall L_1, L_2 : L_1 \geq L_2 \vee L_2 \geq L_1$ . Nun nennen wir in weiser Voraussicht  $L^*$  die Vereinigung aller Elemente  $L$  aus  $\mathbb{K} : L^* = \cup_{L \in \mathbb{K}} L$ . Dadurch gilt  $L_i \leq L^*, \forall L_i \in \mathbb{K}$ . D.h. die Menge  $L^*$  enthält alle linear unabhängigen Mengen aus  $\mathbb{K}$ . Nun müssen wir zeigen, dass  $L^*$  ein Element von  $\mathbb{L}$  ist. Das würde bedeuten, dass  $L^*$  linear unabhängig ist und mit dem Lemma von Zorn dass  $L^*$  das maximale Element von  $\mathbb{L}$  ist.

Also wir müssen zeigen, dass  $L^* = (v_\alpha)_{\alpha \in I}$ , wobei  $\lambda^\beta v_\beta = 0, \lambda^\beta = 0 \forall \lambda^\beta$  gilt. Für Linearkombinationen gilt aber, dass sie endlich sein müssen, deshalb betrachten wir nur eine endliche Anzahl  $n$  der Koeffizienten:  $\lambda^{\beta_1}, \dots, \lambda^{\beta_n}$  und  $v_{\beta_1}, \dots, v_{\beta_n}$ . Da  $L^* = (v_\alpha)_{\alpha \in I}$  gelten soll, aber  $L^*$  eine Vereinigung von Elementen  $L$  aus  $\mathbb{K}$  ist, können wir sagen, dass jedes  $v_{\beta_i}$  in mindestens einem  $L_{\beta_i}$  enthalten ist. Das können wir für alle  $v_{\beta_i}$  behaupten:  $v_{\beta_1} \in L_{\beta_1} \in \mathbb{K}, \dots, v_{\beta_n} \in L_{\beta_n} \in \mathbb{K}$ . Die Teilmenge  $(L_{\beta_1}, \dots, L_{\beta_n})$  von  $\mathbb{K}$  ist wie  $\mathbb{K}$  eine Kette, d.h. wir können wieder die Elemente mit  $\leq$  vergleichen. D.h. wenn ohne Beschränkung der Allgemeinheit wir nun sagen, dass  $L_{\beta_1} \leq L_{\beta_2}$  gilt, können wir behaupten, dass  $v_{\beta_1}$  sowohl in  $L_{\beta_1}$  als auch nun in  $L_{\beta_2}$  ist. Usw. kann man dieses Spiel weiter spielen, bis man zum Schluss kommt, dass es eine obere Schranke  $L_{\beta^*}$  gibt, die alle  $v_{\beta_i}$  beinhaltet und Teil der Kette ist. D.h. in  $L_{\beta^*}$  sind unsere Vektoren enthalten und sie ist linear unabhängig, weil  $L_{\beta^*}$  Teil der Kette ist, wobei die Kette ein Element von  $\mathbb{L}$  ist und da sind ja alle unseren linear unabhängigen Mengen drinnen. Aus diesen Überlegungen können wir nun sagen, dass  $L^*$  linear unabhängig ist. Somit

haben wir bewiesen, dass es eine **maximale** Anzahl an linear unabhängigen Vektoren in unserem Vektorraum gibt.

## 3.2 Der Dualraum

Wie Prof. Balasin gerne sagt, wenn man einen Vektorraum  $V$  hat, bekommt man immer etwas gratis dazu, nämlich den *dualen* Vektorraum  $\tilde{V}$ . Der duale Raum sei folgendermaßen definiert:  $\tilde{V} = \{\phi : V \rightarrow \Gamma \mid \phi \text{ ist linear}\}$ , d.h. die Elemente des dualen Raums sind Funktionen, die Vektoren als Input haben, und als Output Skalare hat. Da diese Eigenschaft nicht reicht um einen Vektorraum aufzubauen, müssen wir noch fordern, dass  $\phi$  *linear* ist. D.h. es gilt:

1.  $\phi(v_1 + v_2) = \phi(v_1) + \phi(v_2), \forall v_1, v_2 \in V$
2.  $\phi(\lambda v) = \lambda \cdot \phi(v) \forall v_1 \in V \wedge \lambda \in \Gamma$

Daraus folgen auch diese Eigenschaften:

1.  $(\phi + \psi)(v_1 + v_2) = \phi(v_1) + \psi(v_1) + \phi(v_2) + \psi(v_2), \forall \phi, \psi \in \tilde{V}$
2.  $(\phi + \psi)(\lambda v) = \lambda(\phi(v) + \psi(v))$

Alles schön und gut, die dualen Vektoren sind also Funktionen, die durch die Wirkung auf Vektoren definiert sind. Durch die Linearität folgen noch viele andere Eigenschaften, die die Wirkung von dualen Vektoren auf Basisvektoren beschreibt: Sei  $(E_\alpha)$  eine Basis von  $(V, \Gamma)$  und  $\phi \in \tilde{V}$ , dann nennen wir die Wirkung von  $\phi$  auf  $E_\alpha$ , also auf einen beliebigen Basisvektor,  $\phi(E_\alpha) = \phi_\alpha \in \Gamma$ . Sei nun  $v$  ein beliebiger Vektor in  $V$ , welcher folgende Linearkombination hat:  $v = v^\alpha E_\alpha$ , wobei  $v^\alpha$  nur Zahlen sind, d.h.  $\in \Gamma$ . Was passiert nun wenn wir  $\phi$  auf unseren Vektor  $v$  wirken lassen? Wir würden folgendes erhalten:

$$\phi(v) = \phi(v^\alpha E_\alpha) = v^\alpha \phi(E_\alpha) = v^\alpha \phi_\alpha \in \Gamma. \quad (3.1)$$

Wir bekommen also eine Summe von Termen, die die Produkte aus den Koeffizienten zur jeweiligen Basis und der Basisvektoren, ausgewertet an  $\phi$ . Nun sind wir sehr schlau und führen eine *duale Basis* ein, die wir nun einfach über ihre Wirkung auf eine *beliebige* Basis von  $V$  definieren:

$$e^\alpha(E_\beta) = \delta_\beta^\alpha. \quad (3.2)$$

Hier soll noch einmal auf die Indizes geachtet werden: Ein oberer Index steht für ein Element des dualen Raums und stellt eine Funktion dar, während ein unterer Index ein normaler Vektor unseres Vektorraum  $V$  ist. Doch wie soll man sich jetzt eine *Basis* einer *Funktion* vorstellen? Dafür werden wir hier ein kleines Beispiel durchrechnen.

### Mathematical Interlude: Die Duale Basis

Für eine duale Basis brauchen wir natürlich eines: eine Basis des Vektorraums! Doch welche nehmen wir? Die Frage ist hinfällig, da für jede Basis eine andere duale Basis existiert. Nehmen wir an, wir haben folgende Basis des  $\mathbb{R}^2$  gegeben:

$$\vec{E}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \vec{E}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad (3.3)$$

Nun wissen wir, dass unsere duale Basis mit dieser Basis folgende Beziehung hat:

$$e^\alpha(E_\beta) = \delta_\beta^\alpha. \quad (3.4)$$

Explizit ausgeschrieben heißt das:

$$e^1(E_1) = 1 \quad (3.5)$$

$$e^1(E_2) = 0 \quad (3.6)$$

$$e^2(E_1) = 0 \quad (3.7)$$

$$e^2(E_2) = 1 \quad (3.8)$$

Was machen wir nun? Wir können annehmen, dass unsere Funktionen  $e^1$  und  $e^2$  die Komponenten unserer Basis mit einer Konstante multiplizieren und diese Terme addieren<sup>1</sup>:

$$e^{\alpha 1} E_{\beta 1} + e^{\alpha 2} E_{\beta 2} = \delta_\beta^\alpha \quad (3.9)$$

Wobei die griechischen Indizes für die jeweiligen Vektoren und Funktionen sind. In Matrixschreibweise angeschrieben heißt das

$$\begin{pmatrix} e^{11} & e^{12} \\ e^{21} & e^{22} \end{pmatrix} \begin{pmatrix} E_{11} & E_{21} \\ E_{12} & E_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (3.10)$$

Für die Basisvektoren eingesetzt erhalten wir:

$$\begin{pmatrix} e^{11} & e^{12} \\ e^{21} & e^{22} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (3.11)$$

Ausmultipliziert gibt das:

$$\begin{pmatrix} e^{11} + e^{12} & e^{11} - e^{12} \\ e^{21} + e^{22} & e^{21} - e^{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (3.12)$$

---

<sup>1</sup>Warum können wir das? Na weil die Dualen Vektoren *lineare* Funktionen sind, also *können* die Komponenten nur mit Potenz 1 vorkommen.

Dieses Gleichungssystem ist leicht zu lösen und erhalten dadurch:

$$\begin{pmatrix} e^{11} & e^{12} \\ e^{21} & e^{22} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \quad (3.13)$$

Also haben unsere linearen Funktionen folgende Form:

$$e^1(E_1) = \frac{1}{2}(1 + 1) = 1 \quad (3.14)$$

$$e^2(E_2) = \frac{1}{2}(1 - (-1)) = 1 \quad (3.15)$$

$$e^1(E_2) = \frac{1}{2}(1 + (-1)) = 0 \quad (3.16)$$

$$e^2(E_1) = \frac{1}{2}(1 - 1) = 0, \quad (3.17)$$

So sehen die genauen Darstellungen der Wirkungen der dualen Vektoren auf die Basis. Schön und gut, aber nicht jeder Vektor ist eine Basis, wir wollen wissen was das  $e^1(v)$  ist, für beliebige Vektoren  $v$ . Tja hier kommt die wichtige Beziehung (3.1) ins Spiel. Sei  $v$  ein Vektor mit folgender Linearkombination mit unseren Basisvektoren:  $v = xE_1 + yE_2$ . In der Basis  $(E_\alpha)$  würde der Vektor also trivialerweise so ausschauen:  $v = \begin{pmatrix} x \\ y \end{pmatrix}$ . Wenn wir nun wissen wollen was z.B.  $e^1(v)$  ist, müssen wir nur für  $v$  unsere Linearkombination einsetzen und die Linearität ausnutzen:

$$e^1(v) = e^1(xE_1 + yE_2) = xe^1(E_1) + ye^1(E_2) = x1 + y0 = x \quad (3.18)$$

Also erhalten wir die *Komponente* des Vektors bezüglich der Basis, wenn wir den jeweiligen dualen Vektor draufjagen! Und so erhalten wir den Wert eines dualen Vektors, angewendet auf einen beliebigen Vektor, wenn wir ihn in der Basis gegeben haben. Diese Beziehung wird uns noch oft begegnen! Deshalb betone ich diese Beziehung nocheinmal: Mit einem dualen Vektor können wir die Komponenten eines normalen Vektors *extrahieren*.

Doch wie können  $e^1$  und  $e^2$  eine Basis eines Vektorraums sein? Betrachten wir eine lineare Funktion  $\phi \in \tilde{V}$ , ausgewertet an einer beliebigen Basis  $(E_\alpha)$ :

$$\phi(E_\alpha) = \phi_\alpha = \phi_\beta \delta_\alpha^\beta = \phi_\beta e^\beta(E_\alpha) \quad (3.19)$$

Aha! Wir können die Werte der Funktion  $\phi$ , ausgewertet an der Basis  $(E_\alpha)$  anschreiben als eine *Linearkombination* von unserer Dualen Basis  $e^\beta$ ! Doch wenn links die Funktionswerte von  $\phi$  an der Basis ausgewertet *gleich* der Funktionswerte einer anderen Funktion für alle möglichen Inputs ist, dann sind diese Funktionen gleich! Anders gesagt: Sind die Wertetabellen zweier Funktionen für alle Argumente gleich, sind die Funktionen gleich. D.h. für einen endlichdimensionalen dualen Vektorraum stellt die Menge  $(e^\alpha)$  eine Basis für *alle* anderen linearen Funktionen dar!

### 3.3 Multilineare Funktionen

Wir haben uns nun intensiv mit linearen Funktionen auseinandergesetzt, jetzt wird es spannend! Nun stellen wir uns die beliebigen *bilinearen* Abbildungen vor:  $B(V, \Gamma) = \{f: V \times V \rightarrow \Gamma \mid f \text{ ist linear in jedem Argument} \leftrightarrow \text{bilinear}\}$ . In Worten ist das die Menge der Funktionen, die zwei Vektoren unseres Vektorraums  $V$  als Input nehmen und diese dann in unseren Körper  $\Gamma$  abbilden, darüber hinaus müssen alle Funktionen  $f$  bilinear sein, also linear in beiden seiner Argumente. D.h.

$$\forall f, g \in B \vee v, w \in V : (f + g)(v, w) = f(v, w) + g(v, w) \quad (3.20)$$

$$(\lambda \cdot f)(v, w) := \lambda \cdot f(v, w) \quad (3.21)$$

Als kleine Sidenote: Stellt man sich nun vor der Vektor  $w$  wäre festgehalten, würde eine bilineare Funktion  $f$  nur noch einen Input haben und dadurch nicht mehr bilinear, sondern nur linear sein.

Man betrachte nun wieder eine Basis  $(E_\alpha)$ . Setzt man jeweils den Basisvektor  $E_\alpha$  und  $E_\beta$  in die beiden Argumente einer bilinearen Funktion  $f$  ein, so folgt:

$$f(E_\alpha, E_\beta) = f_{\alpha\beta} \in \Gamma. \quad (3.22)$$

Für beliebige Vektoren  $v$  und  $w$  gilt nun, analog zu den dualen Vektoren:

$$f(v, w) = f(v^\alpha E_\alpha, w^\beta E_\beta) = v^\alpha w^\beta f_{\alpha\beta}. \quad (3.23)$$

Das ist das bilineare Analogon zu Gl. (3.1). Statt einer Summe, erhalten wir hier zwei Summen über zwei Indizes  $\alpha$  und  $\beta$ .

Nun können wir uns die Frage stellen: Kann man eine bilineare Funktion aus  $B$  als ein Produkt von zwei linearen Funktionen darstellen? Für die Antwort betrachten wir nun zwei lineare Funktionen  $\phi$  und  $\psi \in \tilde{V}$ . Nun behaupten (?) wir, dass das *Tensorprodukt*  $\otimes$  zwischen  $\phi$  und  $\psi$  ein Element aus  $B(V, \Gamma)$  ist. Was ist ein Tensorprodukt? Naja, wir definieren es so:

$$(\phi \otimes \psi)(v, w) := \phi(v) \cdot \psi(w) \quad (3.24)$$

Einfach das Produkt der Funktionen, ausgewertet an dem jeweiligen Vektor. D.h. zu zwei Elementen aus dem dualen Raum, gibt es ein Element im bilinearen Raum.

Betrachten<sup>2</sup> wir nun einen Vektorraum, der *nur* aus solchen Tensorprodukten aufgebaut ist:  $\tilde{V} \otimes \tilde{V} = \otimes^2 \tilde{V} = \{\sum_{(\phi, \psi) \in \tilde{V} \otimes \tilde{V}} \lambda_{(\phi, \psi)} \phi \otimes \psi\} \subseteq B(V, \Gamma)$ . Was heißt das jetzt? Nun es ist ein Vektorraum, dessen Elemente Tensorprodukte von dualen Vektoren sind, die in einer Linearkombination aller

---

<sup>2</sup>Ja wir betrachten sehr viele Dinge hier.

möglichen  $\phi$ 's und  $\psi$ 's dargestellt werden. Seien nun  $f_1$  und  $f_2 \in \otimes^2 \tilde{V}$ , dann gilt:

$$f_{1,2} = \Sigma_{(\phi,\psi)} \lambda_{(\phi,\psi)}^{1,2} \phi \otimes \psi, \in \otimes^2 \tilde{V}, \quad (3.25)$$

$$f_1 + f_2 = \Sigma(\lambda^1 + \lambda^2) \phi \otimes \psi, \in \otimes^2 \tilde{V}, \quad (3.26)$$

$$\mu \cdot f_1 = \Sigma \mu \lambda^1 \phi \otimes \psi, \in \otimes^2 \tilde{V}. \quad (3.27)$$

Nun haben wir schon einen Verdacht, der lautet: Ist unser gerade gebauter Vektorraum  $\otimes^2 \tilde{V}$  gleich  $B(V, \Gamma)$ ? Diese Antwort können wir beantworten, wenn wir zeigen können, dass ein beliebiges Element von  $B$  auch als Element aus  $\otimes^2 \tilde{V}$  geschrieben werden kann. Sei wieder  $f \in B(V, \Gamma)$  und  $(E_\alpha)$  eine Basis von  $V$ :

$$\begin{aligned} f(E_\alpha, E_\beta) &= f_{\alpha\beta} = f_{\gamma\delta} \delta_\alpha^\gamma \delta_\beta^\delta = f_{\gamma\delta} e^\gamma(E_\alpha) e^\delta(E_\beta) = \\ &= f_{\gamma\delta} (e^\gamma \otimes e^\delta)(E_\alpha, E_\beta) = (f_{\gamma\delta} e^\gamma \otimes e^\delta)(E_\alpha, E_\beta) \end{aligned} \quad (3.28)$$

Wir haben es gerade geschafft die Funktion  $f$  mit einer anderen Funktion, die eine Linearkombination von Tensorprodukten ist, gleichzusetzen. Aber beide Funktionen werden an den selben beliebigen Argumenten ausgewertet, d.h. dass diese Funktionen dieselben sind! Damit haben wir bewiesen, dass jedes beliebige Element von  $B$  als ein Element aus  $\otimes^2 \tilde{V}$  dargestellt werden kann. Da das also für jedes Element, also für jede Funktion gilt, gilt offenbar  $B(V, \Gamma) = \otimes^2 \tilde{V}^3!$

Nun hat das Studium uns völlig verrückt gemacht und wir wollen nun *mehr*. Nun betrachten wir Funktionen, die in beliebig vielen Argumenten, nennen wir die Anzahl  $p$ , bilinear sind. Diese sind Elemente des Vektorraums  $\otimes^p \tilde{V} = \text{Mult}_p(V, \Gamma) = \{f : V \times V \times \dots \times V \rightarrow \Gamma \mid f \text{ ist multilinear}\}$ . D.h. es handelt sich um Funktionen die  $p$  Argumente haben, und in jedem Argument ist die Funktion linear. Ohne Beweis gelten folgende Sachverhalte:

1. Sei  $f \in \text{Mult}_p$ :  $f(E_{\alpha_1}, \dots, E_{\alpha_p}) = f_{\alpha_1 \dots \alpha_p} e^{\alpha_1} \otimes \dots \otimes e^{\alpha_p}(E_{\alpha_1}, \dots, E_{\alpha_p})$ .
2.  $e^{\alpha_1} \otimes \dots \otimes e^{\alpha_p}$  Stellt eine Basis in  $\text{Mult}_p$  dar.

Also wir wissen, wie alle Funktionen auf eine Basis wirkt und wie wir die Funktionen in einer Basis darstellen können.

Bevor wir weiter gehen wollen, möchte ich nochmals rekapitulieren was wir hier eigentlich gemacht haben: Wir sind von einem Vektorraum ausgegangen und haben uns die Frage gestellt, was man mit linearen Funktionen alles machen kann. Wir haben gesehen, dass man mithilfe der dualen Vektoren die Komponenten von Vektoren extrahieren kann und dass Funktionen eine Basis haben können, die fest mit einer ganz bestimmten Basis des Ausgangsvektorraums verbunden ist<sup>4</sup>. Weil ein Argument langweilig ist, haben wir auch bilineare Abbildungen analysiert und haben festgestellt, dass diese auch eine Basis besitzen und sehr ähnliche Eigenschaften haben, wie die normalen dualen Vektoren. Nun als letztes haben wir noch kurz multilineare Funktionen betrachtet und für die gelten dieselben

<sup>3</sup>Der letzte Schritt in der Umformung gilt nur, wenn der Vektorraum endlich dimensional ist

<sup>4</sup> $e^\alpha(E_\beta) = \delta_\beta^\alpha$

Dinge wie zuvor. Als nächstes wollen wir wissen, was eigentlich der duale Raum des dualen Raums eigentlich ist.

### 3.4 Der Doppel-Dual-Raum

Also, wir haben uns nun gründlichst mit dem dualen Raum beschäftigt, aber was wäre der duale Raum des Dualen Raums? Sagen wir mal, es gibt ihn, dann würden wir ihn als  $\tilde{\tilde{V}}$  bezeichnen. Was sind seine Elemente? Keine Ahnung, sagen wir mal, sie heißen  $\Phi$ . Die Funktionen des Dualen Raums haben Vektoren des Vektorraums genommen und in den Körper  $\Gamma$  abgebildet. Also kann man die Verbindung machen und sagen, dass die Elemente des Doppel-Dual-Raums Funktionen nehmen und diese in den Körper  $\Gamma$  abbilden. Was das eigentlich nun heißen soll, wird in Kürze erklärt. In Zeichen heißt das:  $\Phi \in \tilde{\tilde{V}}, \Phi : \tilde{V} \rightarrow \Gamma | \Phi$  ist linear. Aber eine Funktion von  $\tilde{V}$  macht erst Sinn, wenn man sie an einen Vektor  $v$  aus  $V$  auswertet. Deshalb sagen wir nun:

$$\Phi_v(\phi) := \phi(v). \quad (3.29)$$

In sehr ähnlicher Form funktioniert die bekannte Dirac-Delta Distribution:  $\delta[\phi(x)] = \phi(0)$ . Da wir fordern das  $\Phi$  linear ist, gilt:

1.  $\Phi_v(\phi_1 + \phi_2) = (\phi_1 + \phi_2)(v) = \phi_1(v) + \phi_2(v)$
2.  $\Phi_v(\lambda\phi) = \lambda\phi(v)$

Da  $v$  ein beliebiger Vektor ist, können wir sagen, dass  $v = v_1 + v_2$ . Dadurch folgt

$$\begin{aligned} \Phi_{v_1+v_2}(\phi) &= \Phi_v(\phi) = \phi(v) = \phi(v_1 + v_2) = \\ &\phi(v_1) + \phi(v_2) = \Phi_{v_1}(\phi) + \Phi_{v_2}(\phi) = (\Phi_{v_1} + \Phi_{v_2})(\phi) \end{aligned} \quad (3.30)$$

und

$$\Phi_{\lambda v} = \lambda\Phi_v(\phi) = \lambda\phi(v). \quad (3.31)$$

Irgendwie hat man das Gefühl, dass es egal ist, ob ich nun ein Element für einen Vektor  $v$  aus dem Doppel-Dual-Raums nimm und ihn an einer Funktion auswerte oder ob ich gleich die Funktion an  $v$  auswerte. Wir werden deshalb die folgende Notation verwenden:  $v(\phi) = \Phi_v(\phi) = \phi(v)$ . Deshalb können wir sagen, dass  $V \subseteq \tilde{\tilde{V}}$  ist. Nun stellt sich die Frage, ob die Gleichheit zwischen diesen Größen gilt. Dafür betrachten wir ein  $\Phi \in \tilde{\tilde{V}}$  und werten es an einem dualen Basisvektor  $e^\alpha$  aus:

$$\Phi(e^\alpha) = \Phi^\alpha = \Phi^\beta \delta_\beta^\alpha = \Phi^\beta e^\alpha(E_\beta) \quad (3.32)$$

Da wir in unserer neuen Notation  $v(\phi) = \Phi_v(\phi) = \phi(v)$  schreiben können, können wir auch  $e^\alpha(E_\beta) = E_\beta(e^\alpha)$  schreiben:

$$\Phi(e^\alpha) = \Phi^\beta e^\alpha(E_\beta) = \Phi^\beta E_\beta(e^\alpha). \quad (3.33)$$

Wir haben somit  $\Phi$  als eine Linearkombination der Basisvektoren  $E_\alpha$  geschrieben. Aber halt, eine Linearkombination von Vektoren ist doch... ein Vektor? Das würde heißen, jedes Element aus dem Doppel-Dual-Raum ist ein normaler Vektor unseres Vektorraums  $V$ ! Und dem ist auch so, der Doppel-Dual-Raum ist unser altbekannter Vektorraum  $V$ .

Warum haben wir diese abstruse Identität hergeleitet? Dass Vektoren Vektoren sind, haben wir schon vorher gewusst. Aber nun haben wir eine neue Eigenschaft unserer Vektoren gefunden, wir können sie als Elemente des Doppel-Dual-Raums auffassen und sie so als Funktionen verstehen. Nun können wir von doppel-dualen Vektoren zu doppel-multilinearen Funktionen gehen, so wie wir es schon beim normalen Dualen Raum gemacht haben. Unser  $\text{Mult}_q \tilde{V} : \{f : \tilde{V} \times \tilde{V} \dots \tilde{V} \mid f \text{ ist multilineare}\}$  ist nun ein Vektorraum der multilinearen doppel-dualen Vektoren, oder equivalent:  $\otimes^q V = \otimes^q \tilde{V}$ . D.h. seine Elemente sind  $(v_1 \otimes \dots \otimes v_q)(\phi_1, \dots, \phi_q) = v_1(\phi_1) \cdot \dots \cdot v_q(\phi_q)$ , also Tensorprodukte von Vektoren, ausgewertet an Funktionen. Doch können wir ja den Spieß umdrehen und sagen, dass  $v_1(\phi_1) \cdot \dots \cdot v_q(\phi_q) = \phi_1(v_1) \cdot \dots \cdot \phi_q(v_q)$  eigentlich ist.

Nun gehen wir einen sehr großen Schritt nach vorne. Jetzt verbinden wir das schon gelernte um eine Abbildung zu erschaffen, die  $p$  Duale-Vektoren und  $q$  Vektoren als *Argumente* hat. Eine solche Funktion ist ein Element des Vektorraums  $\text{Mult}_{p,q} : \otimes_q^p V \{t : \tilde{V} \times \dots \times \tilde{V} \times V \times V \dots \times V \rightarrow \Gamma \mid t \text{ ist in jedem Argument linear}\}$ <sup>5</sup>. D.h. die Elemente des Vektorraums  $\otimes_q^p V = (\otimes^p V) \otimes (\otimes_q \tilde{V})$  sind Funktionen die  $p$  Argumente haben, die normale Vektoren sind, und  $q$  Argumente haben, die Duale-Vektoren sind. Wir können die Elemente als Linearkombination wieder anschreiben:  $t = \sum \lambda_{\phi_1, \dots, \phi_p, v_1, \dots, v_q} (\phi_1 \otimes \dots \otimes \phi_p \otimes v_1 \otimes \dots \otimes v_q)$ . Hier muss man vorsichtig nun sein:  $t$  hat  $p$  Argumente, die Vektoren sind, deshalb braucht  $t$  in der Linearkombination auch  $p$  Duale-Vektoren im Tensorprodukt, um die Argumente auszuwerten. Dasselbe gilt für die  $q$  Argumente, die Duale-Vektoren sind.

Mit diesen Werkzeugen in den Händen können wir uns nun auf die Tensoralgebra stürzen.

### 3.5 Tensoralgebra

Die Objekte des letzten Abschnittes werden auch *Tensoren* genannt, wobei  $p$  und  $q$  jeweils eine beliebige natürliche Zahl sein kann. Was wäre dann ein  $\otimes_0^0 V$ ? Das ist ein Tensor der Stufe  $(0,0)$ ; Skalar, einfach eine Zahl. Ein  $\otimes_0^1 V$ ? Um das zu überprüfen, gehen wir zurück zur Definition unseres  $\text{Mult}_{p,q} : \otimes_q^p V \{t : \tilde{V}_1 \times \dots \times \tilde{V}_p \times V_1 \times V_2 \dots \times V_q \rightarrow \Gamma \mid t \text{ ist in jedem Argument linear}\}$ <sup>6</sup>. Hier haben wir

<sup>5</sup>Da die Notation hier *sehr* wichtig ist, konnte sich der Autor seiner eigenen Mitschrift nicht mehr sicher sein. Deshalb wird von nun an der Merktzettel zur ART von <https://goldsilberglitzer.at/> verwendet.

<sup>6</sup> $V_i$  kennzeichnet hier nur eine Nummerierung, um die Anzahl der Argumente hervorzuheben. An sich sind es immer dieselben Vektorräume.

gesagt, dass t p *Argumente* Duale-Vektoren und q Argumente Vektoren hat. Als Regel muss man sich das nun merken:

$\otimes_q^p V$  bedeutet ein Tensor der Stufe (p,q). Der **obere** (p) Index zeigt an aus wie vielen **Vektoren** der Tensor besteht<sup>7</sup> und der **untere** (q) zeigt an wie viele **Duale-Vektoren**<sup>8</sup>.

D.h. ein Tensor der Stufe (1,0) ist ein normaler Vektor. Ein Tensor der Stufe (0,1) ist ein Dualer-Vektor. Der Begriff „woraus der besteht“ wird klarer, wenn wir gleich die Basis eines (p,q) Tensors einführen.

Sei t ein Tensor der Stufe (p,q), also er hat p Duale-Vektoren und q normale Vektoren als Input. Wenden wir den Tensor t auf eine Basis ( $E_\alpha$ ) und eine duale Basis ( $e^\beta$ ) an, können wir schreiben:

$$\begin{aligned} t(e^{\beta_1}, \dots, e^{\beta_p}, E_{\alpha_1}, \dots, E_{\alpha_q}) &= t_{\alpha_1 \dots \alpha_q}^{\beta_1 \dots \beta_p} = t_{\gamma_1 \dots \gamma_q}^{\delta_1 \dots \delta_p} \delta_{\delta_1}^{\beta_1} \dots \delta_{\delta_p}^{\beta_p} \delta_{\alpha_1}^{\gamma_1} \dots \delta_{\alpha_q}^{\gamma_q} = \\ & t_{\gamma_1 \dots \gamma_q}^{\delta_1 \dots \delta_p} E_{\delta_1}(e^{\beta_1}) \dots E_{\delta_p}(e^{\beta_p}) e^{\gamma_1}(E_{\alpha_1}) \dots e^{\gamma_q}(E_{\alpha_q}) = \\ & (t_{\gamma_1 \dots \gamma_q}^{\delta_1 \dots \delta_p} E_{\delta_1} \otimes \dots \otimes E_{\delta_p} \otimes e^{\gamma_1} \otimes \dots \otimes e^{\gamma_q})(e^{\beta_1}, \dots, e^{\beta_p}, E_{\alpha_1}, \dots, E_{\alpha_q}). \end{aligned} \quad (3.34)$$

Und schon wieder haben wir zwei Funktionen, ausgewerten an den selben Argumenten, gleichgesetzt. D.h. ( $E_{\delta_1} \otimes \dots \otimes E_{\delta_p} \otimes e^{\gamma_1} \otimes \dots \otimes e^{\gamma_q}$ ) stellt eine Basis des Vektorraums  $\otimes_q^p V$  dar! D.h. wir können alle beliebigen Tensoren als Linearkombination von dieser Basis und normalen Skalaren darstellen.

### Mathematical Interlude: Abstrakte Indexschreibweise

Da wir alle wissen, dass Physiker im Herzen alle faul sind, müssen wir uns nun eine Schreibweise angewöhnen, die diese Faulheit widerspiegelt. Ihr Name lautet *Abstrakte Indexschreibweise*. Mit dieser Notation können wir ohne explizite Nennung der Basis angeben, um welches Objekt es sich handelt. Z.B. was könnte v sein? In unserer bisherigen Notation kann das Alles mögliche sein, ein Vektor, aber auch ein Tensor der Stufe (6,9)! Hier findet man die Stärke der abstrakten Indexschreibweise.

Allgemein kann man nun sagen, dass wir Tensoren t mit beliebiger Stufe (p,q)<sup>9</sup> anschreiben werden als  $t \rightarrow t_{b_1 \dots b_q}^{a_1 \dots a_p}$ . Wir werden also die *Art*, um welchen Tensor es sich hier handelt mit lateinischen Buchstaben anschreiben, während wir seine *Komponenten* bzgl. einer Basis mit griechischen Indizes bezeichnen werden. In Tabelle 3.1 ist die neue Notation für wichtige Beispiele illustriert. Als Faustregel kann man hier nun sagen:

Ein **oberer** abstrakter Index zeigt ein Element des Vektorraums V an, ein **unterer** abstrakter Index zeigt ein Element des Dualen Raums an.

Dabei ist aber der lateinische Index kein Summenindex, auch keine Nummerierung. Er zeigt einfach die *Stufe* des Tensors direkt an, ohne eine Basis explizit hinzuschreiben. Am Anfang macht es noch nicht viel Sinn, aber mit der Zeit erkennt man den Wert dieser Notation. Z.B. können wir den Vektor

<sup>7</sup>Bzw. wie viele *Argumente* Duale-Vektoren sind.

<sup>8</sup>Oder wie viele *Argumente* normale Vektoren sind.

<sup>9</sup>Die Stufe ist ja aus dem Buchstaben t nicht ersichtlich.

| Objekt        | Alte Schreibweise | Neue Notation                       |
|---------------|-------------------|-------------------------------------|
| Vektor        | $v$               | $v^a$                               |
| Dualer-Vektor | $\omega$          | $\omega_a$                          |
| Skalar        | $f$               | $f$                                 |
| Tensor (p,q)  | $t$               | $t_{b_1 \dots b_q}^{a_1 \dots a_p}$ |
| Tensor (0,2)  | $g$               | $g_{ab}$                            |

Tabelle 3.1: Abstrakte Indexschreibweise

$v$  jetzt anschreiben als

$$v^a = v^\alpha E_\alpha^a. \quad (3.35)$$

Wobei über  $\alpha$  summiert wird, aber  $a$  nur die Art des Tensor anzeigt. Weil  $v^a$  und auch natürlich die Basisvektoren ( $E_\alpha^a$ ) Vektoren sind, erhalten diese beiden Objekte abstrakte Indizes  $a$ , wobei darauf geachtet werden muss, dass dieselben freien abstrakten Indizes auf beiden Seiten der Gleichung auftauchen. Ein Element des dualen Raums wäre dann:

$$\omega_a = \omega_\alpha e_a^\alpha \quad (3.36)$$

Die lineare Funktion  $\omega_a$  an  $v^a$  ausgewertet, wäre dann:

$$\omega(v) = \omega_a v^a \quad (3.37)$$

also zeigt ein kontrahierter<sup>10</sup> abstrakter Index eine Anwendung einer Funktion an! Hier können wir noch weiter machen, ein Tensorprodukt wäre:

$$(\omega \otimes \phi)(v, w) = \omega_a \phi_b v^a w^b = \omega_a v^a \phi_b w^b \quad (3.38)$$

D.h. ein beliebiger Tensor der Stufe (1,3) könnte dann als Tensorprodukt von den einzelnen Tensoren angeschrieben werden

$$t_c^{ab} = r^{ab} \Theta_c = v^a w^b \Theta_c \quad (3.39)$$

Für beliebige Vektoren und Duale-Vektoren. Doch was ist, wenn wir einen Tensor in einer Basis darstellen wollen? Das würde so gehen:

$$t_{b_1 \dots b_q}^{a_1 \dots a_p} = t_{\gamma_1 \dots \gamma_q}^{\delta_1 \dots \delta_p} E_{\delta_1}^{a_1} \dots E_{\delta_p}^{a_p} e_{b_1}^{\gamma_1} \dots e_{b_q}^{\gamma_q} \quad (3.40)$$

Wie versprochen, wird über alle griechischen Indizes summiert, während die lateinischen Indizes die

<sup>10</sup>↔ in einem Term zweimal vorkommender

Art des Tensors anzeigen. Und nun ergibt unsere vorherige Notation des  $\otimes_q^p V$  mehr Sinn, denn wir können nun den oberen Index nun direkt mit einem Vektor  $v^a$  und den unteren Index mit einem Dualen-Vektor  $\omega_a$  identifizieren.

Nun gibt es besondere Tensoren, mit besondere Eigenschaften: Die Metrik. Wir kennen schon eine Metrik: die Minkowski-Metrik der SRT. Mit einer Metrik kann man die abstrakten Indizes hinauf und hinunterschieben:

$$v^a \eta_{ab} = v_b \in \tilde{V} \quad (3.41)$$

Wir können also mit der Metrik aus Vektoren Duale-Vektoren machen! Aber wir wissen noch mehr! Für eine orthonormale Basis ( $E_\alpha^a$ ) des Vektorraums gilt:

$$\eta_{ab} E_\alpha^a E_\beta^b = \eta_{\alpha\beta} \leftrightarrow \text{diag}(-1, 1, 1, 1) \quad (3.42)$$

Also hat die Metrik die Form einer Diagonalmatrix mit einem negativen Eintrag. Hier muss aber unterschieden werden:  $\eta_{ab}$  ist ein Tensor der Stufe (0,2), d.h. in ihrer Basis hat die Metrik zwei Duale-Vektoren:

$$\eta_{ab} = \eta_{\alpha\beta} e_a^\alpha e_b^\beta, \quad (3.43)$$

Also hat die Metrik bzgl. zu einer orthonormalen dualen Basis die Form einer Diagonalmatrix. Hier ist der große Unterscheid zwischen ihrer Form bzgl. einer Basis und dem eigentlichen Tensor. So wie man die Komponenten des Vektors  $v = \begin{pmatrix} x \\ y \end{pmatrix}$  bzgl. einer Basis nicht als den Vektor selbst ansieht. Hat man nun eine andere ONB ( $\tilde{E}_\alpha^a$ ) gegeben, gibt es eine lineare Transformation zwischen den beiden Basen:

$$\tilde{E}_\alpha^a = \Lambda_\alpha^\beta E_\beta^a \quad (3.44)$$

Dann kann man die Einträge  $\eta_{\alpha\beta}$  der Metrik bzgl. der einen Basis mit der anderen gleichsetzen:

$$\eta_{\alpha\beta} = \eta_{ab} E_\alpha^a E_\beta^b = \eta_{ab} \tilde{E}_\alpha^a \tilde{E}_\beta^b \quad (3.45)$$

Setzen wir nun das Transformationsgesetz ein, erhalten wir:

$$\eta_{\alpha\beta} = \eta_{ab} \Lambda_\alpha^\delta E_\delta^a \Lambda_\beta^\gamma E_\gamma^b = \Lambda_\alpha^\delta \Lambda_\beta^\gamma \eta_{\delta\gamma}, \quad (3.46)$$

wobei wir wieder die Beziehung zwischen der Metrik und der Basis ( $\tilde{E}_\alpha^a$ ) benutzt haben. Wie sehen nun Tensoren aus, wenn sie mit der Metrik verändert wurden? Dazu betrachten wir einen Vektor  $v^a$

und sehen uns seine kovarianten Komponenten mithilfe der Metrik an:

$$v^a = v^\alpha E_\alpha^a \quad (3.47)$$

$$v^a \eta_{ab} = v^\alpha E_\alpha^a \eta_{ab} \quad (3.48)$$

$$v_b = v^\alpha E_\alpha^a \eta_{\delta\gamma} e_a^\gamma e_b^\delta \quad (3.49)$$

$$v_b = v^\gamma \eta_{\delta\gamma} e_b^\delta \quad (3.50)$$

$$v_b = v_\delta e_b^\delta \quad (3.51)$$

Explizit gilt für die Komponenten eines 4-Vektor in der SRT:

$$v_\delta = v^\gamma \eta_{\delta\gamma} \quad (3.52)$$

$$v_0 = v^\gamma \eta_{0\gamma} = -v^0 \quad (3.53)$$

$$v_1 = v^\gamma \eta_{1\gamma} = v^1 \quad (3.54)$$

$$v_2 = v^\gamma \eta_{2\gamma} = v^2 \quad (3.55)$$

$$v_3 = v^\gamma \eta_{3\gamma} = v^3, \quad (3.56)$$

also sind die räumlichen Komponenten eines 4-Vektors gleich, nur die zeitliche Komponente hat ein anderes Vorzeichen. Die Inverse der Metrik existiert auch und lautet  $\eta^{ab} \in \otimes^2 V$ . Für sie gilt:

$$\eta^{ab} = \eta^{\alpha\beta} E_\alpha^a E_\beta^b \quad (3.57)$$

$$\eta^{ab} \eta_{bc} = \eta^{\alpha\beta} E_\alpha^a E_\beta^b \eta_{\gamma\delta} e_b^\gamma e_c^\delta = \eta^{\alpha\beta} \eta_{\gamma\delta} \delta_\beta^\gamma E_\alpha^a e_c^\delta = \delta_\delta^a \delta_c^\delta = \delta_c^a \quad (3.58)$$

Nun haben wir vieles zu Vektorräumen und multilinearer Algebra gelernt. Wir haben herausgefunden, dass jeder Vektorraum eine Basis hat, und die Vektoren eine eindeutige Basisdarstellung haben. Weiters fanden wir einen zusätzlichen Vektorraum, den Dualen-Raum. Hier haben wir ebenfalls eine Duale-Basis gefunden, welche fest mit unserer Basis des Vektorraums verbunden ist. Dann haben wir uns an multilineare Funktionen herangetraut, für beliebige Argumente. Dort fanden wir die allgemeinste Form eines Tensors, nämlich als multilineare Funktion mit Argumenten aus  $V$  und  $\tilde{V}$ . Weiters haben wir eine neue Notation eingeführt, um Basisunabhängig unsere Objekte anzuschreiben. Als letztes haben wir uns noch die Metrik genauer untersucht. Nun können wir zum nächsten Kapitel vordringen: *Die topologischen Räume und Mannigfaltigkeiten*.

# Kapitel 4

## Topologien und topologische Räume

### 4.1 Begriffe der Topologie

Die entscheidenden Räume, in denen sich die Allgemeine Relativitätstheorie aufhält, sind die topologischen Räume. Man kann sie als verallgemeinerte Vektorräume ansehen. Zwei Objekte sind für topologische Räume wichtig: Eine Menge  $X$  und eine Topologie  $\mathfrak{X}$ . Was eine Menge ist, wissen wir: Ein Verbund von unterscheidbaren Elementen. Z.B. wäre  $\{1,2,3,4,5,6\}$ , die Augenzahlen eines Würfels eine Menge. Was soll nun eine Topologie sein? Eine Topologie (auf  $X$ ) ist die Menge von Teilmengen von  $X$ , die mindestens 2 Elemente hat: die Leere Menge  $\emptyset$  und  $X$  selbst. D.h. die *Elemente* einer Topologie  $\mathfrak{X}$  sind *Mengen* selbst. Ohne weitere Bedingungen wäre eine Topologie für unseren Würfel folgende Menge:  $\mathfrak{X} = \{O_1, O_2, X, \emptyset\} = \{\{1\}, \{3\}, \{1, 3\}, \{1, 2, 3, 4, 5, 6\}, \emptyset\}$ , man beachte, dass  $\mathfrak{X}$   $X$  und  $\emptyset$  beinhaltet und, dass  $O_i$  „richtige“ Teilmengen von  $X$  sind. Aber da unser Leben sonst viel zu öde ist, gibt es noch zusätzliche Forderungen an unseren topologischen Raum, drei an der Zahl:

1.  $\emptyset, X \in \mathfrak{X}$ <sup>1</sup>
2.  $O_1, O_2 \in \mathfrak{X} \rightarrow O_1 \cap O_2 \in \mathfrak{X}$
3.  $(O_\alpha) \subseteq \mathfrak{X} \rightarrow \cup_\alpha O_\alpha \in \mathfrak{X}$

Was bedeutet das? Die erste Bedingung ist trivial und schon besprochen. Die zweite ist interessanter: Wir nehmen zwei Teilmengen der Menge  $X$ <sup>2</sup>, die Elemente der Topologie  $\mathfrak{X}$  sind, und bilden den Durchschnitt. Und dieser Durchschnitt soll auch wieder ein Element von  $\mathfrak{X}$  sein. Die dritte Bedingung

---

<sup>1</sup>Wie bereits erwähnt

<sup>2</sup>auch offene Mengen  $O_\alpha$  genannt!

ist ähnlich: Dieses Mal nehmen wir eine beliebige Anzahl von Teilmengen von  $X$ , die wie immer Elemente von  $\mathfrak{X}$  sind, bilden die Vereinigung davon und diese Menge soll auch wieder in  $\mathfrak{X}$  liegen. Wegen diesen Bedingungen ist auch unsere „Topologie“ des Würfels keine richtige Topologie, denn wenn wir den Durchschnitt von  $\{1, 3\}$  mit  $\{1, 2, 3, 4, 5, 6\}$  bilden, ergibt das die Menge  $\{2, 4, 5, 6\}$ , welche nicht in  $\mathfrak{X}$  enthalten ist, ergo keine Topologie.

Wenn nun  $X$  eine Topologie  $\mathfrak{X}$  hat, nennt man das Paar  $(X, \mathfrak{X})$  einen *topologischen Raum*. Es gibt nun zwei besondere Arten von Topologien  $\mathfrak{X}$ : die Indiskrete Topologie  $\mathfrak{I}$  und die diskrete Topologie  $\mathfrak{D}$ . Die indiskrete ist das Minimum, was eine Topologie können muss, denn sie besteht nur aus der Leeren Menge  $\emptyset$  und  $X$  selbst:  $\mathfrak{I} = \{\emptyset, X\}$ . Die diskrete Topologie ist sozusagen das Gegenteil, sie beinhalten *alle* möglichen Teilmengen von  $X$ . D.h. sie beinhaltet die sogenannte *Potenzmenge*  $\mathcal{P}(X)$ :  $\mathfrak{D} = \{\mathcal{P}(X)\}$ . Die Potenzmenge ist so definiert:  $\mathcal{P}(X) = \{U | U \subseteq X\}$ , also alle Teilmengen von  $X$ .

Nun können wir den Begriff der *Konvergenz* auf einem topologischen Raum definieren:

Sei  $(x_n) \subseteq X, x_n \rightarrow x \in X : \Leftrightarrow$  zu jedem  $O \in \mathfrak{X}: x \in O \exists N_O \in \mathbb{N}$  sodass  $x_n \in O$  für  $n \geq N_O$ .

Umgangssprachlich heißt das: Die Menge der Folgenglieder  $(x_n)$  sei eine Teilmenge unserer Menge  $X$ . Wenn die Folge gegen einen Grenzwert  $x$  konvergiert, bedeutet das, dass es eine offene Umgebung  $O$  um den Grenzwert  $x$  gibt. Darüber hinaus gibt es einen Folgenindex  $N_O$  für den gilt, dass *alle* Folgenglieder mit einem größeren Folgenindex in dieser Umgebung um  $x$  sind. Dieses Konvergenzkriterium ist sehr ähnlich zur  $\epsilon$ -Umgebung aus Analysis 1.

Für unsere beiden Topologien  $\mathfrak{I}$  und  $\mathfrak{D}$  bedeutet das nun folgendes:

1. In topologischen Räumen mit der indiskreten Topologie  $\mathfrak{I}$  konvergiert *jede* Folge, da es nur eine Umgebung  $O$  um den Grenzwert geben kann, und zwar die ganze Menge selbst,  $O = X$ . Da aber jedes Folgenglied in der Menge enthalten ist, konvergiert jede Folge gegen *jeden* Wert von  $X$ . Kurzgesagt: *Alle Folgen konvergieren gegen jeden Wert bei der indiskreten Topologie.*
2. In topologischen Räumen mit der diskreten Topologie  $\mathfrak{D}$  konvergieren nicht alle Folgen. Die Folgen, die jedoch nach unserem Konvergenzkriterium konvergieren, sind nach dem minimalen Index  $N_O$  dann konstant. Man kann sich das so vorstellen: Die Folge konvergiert gegen ihren Grenzwert, d.h. es gibt eine Umgebung davon. Aber da wir die diskrete Topologie haben, können wir annehmen, dass die Umgebung nur den Grenzwert selbst  $O = \{x\}$  enthält. Für jeden Index  $n \geq N_O$  ist dann jedes Folgenglied in der Umgebung, aber da die Umgebung nur  $x$  enthält, sind alle Folgenglieder danach konstant. Kurzgesagt: *Konvergente Folgen sind in der diskreten Topologie nach dem Grenzwert konstant.*

## 4.2 Abbildungen auf topologischen Räumen

Nun können wir uns über Abbildungen zwischen topologischen Räumen  $(X, \mathfrak{X})$  und  $(Y, \mathfrak{Y})$  unterhalten. Sei  $f$  eine Abbildung von  $X$  nach  $Y$ :  $f: X \rightarrow Y$ . D.h. alle Elemente von  $X$  werden auf  $Y$

abbgebildet, wobei nicht jedem Element von  $Y$  einem Wert auf  $X$  entsprechen muss. Das Urbild  $f^{-1}(B) \in X$  sei die Menge aller  $x \in X$  für die gilt:  $f(x) \in B$ . Wir nennen  $f$  stetig wenn gilt:

$$f \text{ ist stetig} :\Leftrightarrow f^{-1}(O') \in \mathfrak{X}, \forall O' \in \mathfrak{Y}$$

D.h. wenn die Umkehrabbildung  $f^{-1}$  angewandt auf irgendeine beliebige Teilmenge von  $Y$ , also ein Element aus  $\mathfrak{Y}$ , ein Element von  $X$  ist, dann ist  $f$  stetig. Wenn zwei Abbildungen  $f: X \rightarrow Y$  und  $g: Y \rightarrow Z$ , mit  $(Z, \mathfrak{Z})$ , beide stetig sind, ist ihre Komposition  $f \circ g$  auch stetig, wobei  $(g \circ f)(x) := g(f(x))$  und  $C \subseteq Z$  gilt:

$$x \in (g \circ f)^{-1}(C) \Leftrightarrow (g \circ f)(x) \in C \Leftrightarrow g(f(x)) \in C \Leftrightarrow f(x) \in g^{-1}(C) \Leftrightarrow x \in f^{-1}(g^{-1}(C))$$

$$(g \circ f)^{-1}(C) = g^{-1}(f^{-1}(C)) \quad (4.1)$$

Umgangssprachlich ausgedrückt heißt das folgendes: Sei  $x$  ein Element aus der Inversen von  $(g \circ f)(C)$  bzw. gilt natürlich  $x \in X$ . Diese Aussage ist equivalent zu der, dass  $(g \circ f)$  angewandt auf  $x$  in  $C$  liegt, einfach die Inverse ausgenutzt. Dann haben wir unsere Definition einer Komposition angewendet  $(g \circ f)(x) := g(f(x))$ , was auch wieder in  $C$  liegen muss. Dann muss aber  $f(x)$  in der Inversen von  $g(C)$  liegen, und dass ist equivalent, dass  $x$  in  $f^{-1}(g^{-1}(C))$  ist.

Stellen wir uns nun vor, wir sind in der misslichen Lage eine Abbildung  $f$  zwischen zwei Mengen  $X$  und  $Y$  zu haben, wobei nur  $Y$  eine Topologie besitzt und  $X$  nicht! Schrecklich... können wir uns da irgendwie behelfen? Die Antwort ist ja! Wir können eine Topologie  $\mathfrak{X}_f$  mithilfe der Urbilder von  $\mathfrak{Y}$  unter  $f$  konstruieren<sup>3</sup>. Es gilt:  $\mathfrak{X}_f = \{O : \exists O' \in \mathfrak{Y} : O = f^{-1}(O')\}$ , also  $\mathfrak{X}_f$  besteht aus den Bildern der Inversen von  $f$  angewandt auf beliebige Teilmengen von  $Y$ . Hier die Beweise, dass es sich tatsächlich um eine Topologie handelt:

1.  $\emptyset = f^{-1}(\emptyset); X = f^{-1}(Y)$ .
2. Sei  $O'_1, O'_2, \dots \in \mathfrak{Y}$  und  $O_1 = f^{-1}(O'_1), O_2 = f^{-1}(O'_2), \dots \in \mathfrak{X}_f$ . Da  $(Y, \mathfrak{Y})$  gilt, gilt  $O'_1 \cap O'_2 \in \mathfrak{Y}$ . Dann gilt aber:  $f^{-1}(O'_1 \cap O'_2) \in \mathfrak{X}_f \Leftrightarrow f^{-1}(O'_1) \cap f^{-1}(O'_2) \in \mathfrak{X}_f \Leftrightarrow O_1 \cap O_2 \in \mathfrak{X}_f$ .
3. Sei  $(O'_\alpha) \subseteq \mathfrak{Y} \rightarrow \cup_\alpha O'_\alpha \in \mathfrak{Y} \rightarrow f^{-1}(\cup_\alpha O'_\alpha) \in \mathfrak{X}_f \Leftrightarrow \cup_\alpha f^{-1}(O'_\alpha) \in \mathfrak{X}_f \Leftrightarrow \cup_\alpha O_\alpha \in \mathfrak{X}_f$ .

In einer sehr ähnlichen Form können wir nun aus einer Menge  $X$  und einer Teilmenge  $U$  eine Topologie konstruieren, ohne explizite Abbildung  $f$ . Die *Spurtopologie* auf  $U$  wird vom topologischen Raum  $(X, \mathfrak{X})$  „geerbt“. Sei  $i_U$  die Einbettungsabbildung:  $i_U : U \rightarrow X$ , d.h. ein Element aus  $U$  wird in sich selbst abgebildet, welches wiederum ein Element aus  $X$  ist:  $i_U(x) = x; x \in i_U^{-1}(O) = i_U(x) \Leftrightarrow x \in U \cap O$ . Nun können wir die Spurtopologie definieren:  $\mathfrak{X}_{i_U} = \{O' : O' = O \cap U \wedge O \in \mathfrak{X}\}$ . D.h. jede Untermenge kann eine Topologie erben.

<sup>3</sup>Das nennt man topologisieren.

<sup>4</sup>Beziehung hier ohne Beweis angeführt.

<sup>5</sup>Hier ebenso ohne Beweis angeführt

Nun brauchen wir noch einen Begriff: Den Homöomorphismus. Sei  $f$  eine stetige Abbildung  $f: X \rightarrow Y$  und  $g$  ebenso stetig  $g: Y \rightarrow X$ , mit  $X$  und  $Y$  beide topologische Räume. Weiters soll gelten, dass  $(g \circ f) = \text{id}_X$  und  $(f \circ g) = \text{id}_Y$ , also die Kompositionen sind die Identitätsabbildungen für die jeweiligen topologischen Räume  $X$  und  $Y$ .  $f^{-1}$  ist im Allgemeinen nicht stetig, auch nicht  $g^{-1}$ . Da beide Abbildungen  $f$  und  $g$  selber aber stetig sind, gilt  $f^{-1}(\mathfrak{Y}) \subseteq \mathfrak{X} \wedge g^{-1}(\mathfrak{X}) \subseteq \mathfrak{Y}$ . Daraus folgt:

$$\begin{aligned}
 f^{-1}(\mathfrak{Y}) &\subseteq \mathfrak{X} \\
 g^{-1}(f^{-1}(\mathfrak{Y})) &\subseteq g^{-1}(\mathfrak{X}) \subseteq \mathfrak{Y} \\
 \text{id}_Y(\mathfrak{Y}) = \mathfrak{Y} &\subseteq g^{-1}(\mathfrak{X}) \subseteq \mathfrak{Y} \\
 g^{-1}(\mathfrak{X}) &= \mathfrak{Y},
 \end{aligned} \tag{4.2}$$

wobei  $f^{-1}(\mathfrak{Y}) = \mathfrak{X}$  gleichermaßen gilt. Wenn also  $f$  und  $g$  stetig sind und  $(g \circ f) = \text{id}_X$  und  $(f \circ g) = \text{id}_Y$  gilt, sind diese Funktionen Homöomorphismen. Kurz, aber schlampig formuliert: *Ein Homöomorphismus ist eine bijektive Abbildung zwischen zwei topologischen Räumen.*

### 4.3 Trennungsaxiome

Die Trennungsaxiome werden wichtig sein, um unsere Mannigfaltigkeiten zu definieren. Sie bringen eine gewisse Struktur zu unseren topologischen Räumen und geben uns mehr Informationen über jene. Insgesamt gibt es fünf Trennungsaxiome, wobei wie in der SRT mit 0 zu zählen begonnen wird. Bevor wir die Trennungsaxiome besprechen, muss noch gesagt werden, was offene und geschlossene Menge auf einer Topologie sind. Man nennt  $A \subseteq X$  eine geschlossene Menge, wenn gilt:  $\exists O \in \mathfrak{X} : A = X - O$ , d.h. es soll eine offene Menge  $O$  in  $\mathfrak{X}$  geben, wodurch man  $A$  als „ $X$  ohne  $O$ “ anschreiben kann, bzw. es gilt  $X - A = O \in \mathfrak{X}$ . Jetzt wo diese Formalität aus dem Weg ist, können wir die Trennungsaxiome auflisten. Das sind die Trennungsaxiome, wobei  $(X, \mathfrak{X})$  wieder ein topologischer Raum ist:

1.  $T_0 : \forall x \neq y \in X \leftrightarrow (x \in O_X \wedge y \notin O_X) \vee (x \notin O_Y \wedge y \in O_Y)$
2.  $T_1 : \forall x \neq y \in X \leftrightarrow (x \in O_X \wedge y \notin O_X) \wedge (x \notin O_Y \wedge y \in O_Y)$
3.  $T_2 : \forall x \neq y \in X \leftrightarrow \exists x \in O_X \wedge y \in O_Y : O_X \cap O_Y = \emptyset$
4.  $T_3 : A \subseteq X, x \notin A \rightarrow \exists x \in O_X, O \supseteq A : O_X \cap O = \emptyset$
5.  $T_4 : A_1; A_2$  abgeschlossen:  $A_1 \cap A_2 = \emptyset \rightarrow \exists O_1, O_2 \in \mathfrak{X} : O_1 \supseteq A_1, O_2 \supseteq A_2 : O_1 \cap O_2 = \emptyset$

Um Verwirrungen aufzuklären, was diese Axiome heißen werden sie nun nacheinander abgehandelt und erklärt. Das Trennungsaxiom  $T_0$  ist das gröbste und sagt folgendes aus: Für zwei beliebige ungleiche Elemente  $x$  und  $y$  aus  $X$  gibt es eine offene Umgebung, die  $x$  umgibt und  $y$  nicht, *oder* umgekehrt. Dieses *oder* ist als logisches oder zu verstehen, also  $1 \vee 1 = 1$ , also es können beide

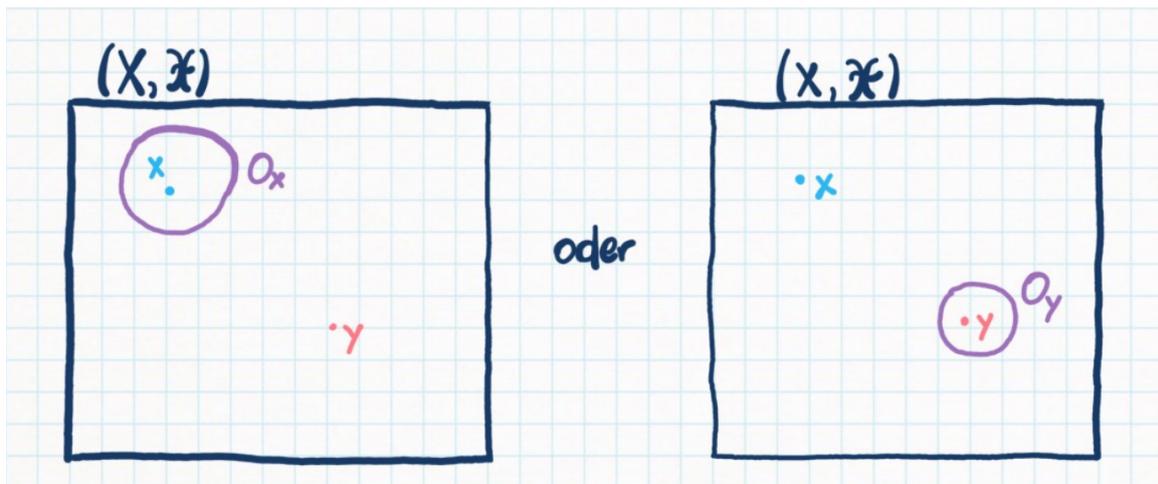


Abbildung 4.1: Zur Illustration des Trennungsaxiom  $T_0$ .

Punkte offene Umgebungen haben, die den jeweils anderen Punkt nicht enthalten, müssen es aber nicht, siehe Abb. 4.1.

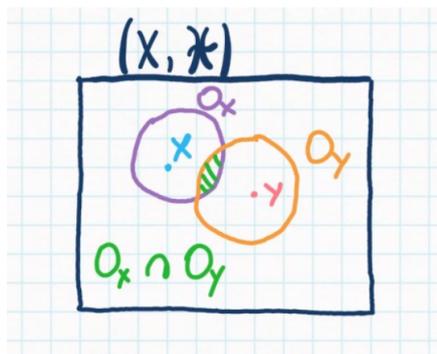


Abbildung 4.2: Zur Illustration des Trennungsaxiom  $T_1$ .

$T_1$  ist sehr ähnlich zu  $T_0$  nur, dass wir hier fordern, dass es gleichzeitig zwei offene Umgebungen um  $x$  und  $y$  gibt, die den jeweils anderen Wert nicht enthält. Über den Durchschnitt dieser beiden Umgebungen (= Teilmengen von  $X$ ) treffen wir noch keine Aussagen, siehe Abb. 4.2. Durch  $T_1$  folgt, dass die Menge, die nur  $x$  enthält  $\{x\}$  eine abgeschlossene Menge ist. Es gilt:  $\forall x \in X : X - \{x\} = \mathcal{X}$  ist offen;  $y \in \mathcal{X} = X - \{x\} \rightarrow y \neq x \rightarrow \exists y \in O_y \wedge x \notin O_y \rightarrow \{y\} \subseteq O_y \subseteq X - \{x\} \forall y \in X - \{x\} \rightarrow X - \{x\} = \cup_{y \in X - \{x\}} y \subseteq \cup_{y \in X - \{x\}} O_y \subseteq X - \{x\}$ ;  $\cup_{y \in X - \{x\}} O_y = X - \{x\} \rightarrow x$  ist eine abgeschlossene Menge. Wir haben es also geschafft  $X - \{x\}$  als eine Zusammenfügung von Elementen von  $\mathcal{X}$  zu schreiben, was bedeutet, dass  $\{x\}$  eine abgeschlossene Menge ist.

$T_2$  ist wieder sehr ähnlich zu seinem Vorgänger, doch hier stellen wir die Bedingung, dass der Durchschnitt der Umgebungen die Leere Menge sein soll, sie sollen also disjunkt sein, siehe Abb. 4.3.

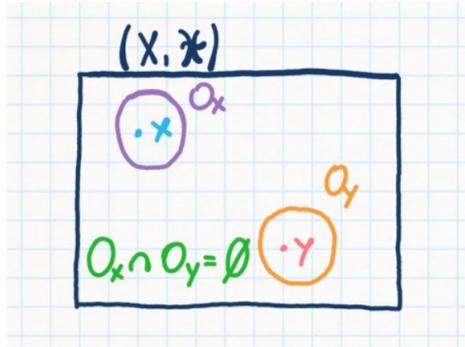


Abbildung 4.3: Zur Illustration des Trennungsaxiom  $T_2$ .

Wenn ein topologischer Raum  $T_2$  erfüllt, nennt man ihn auch Hausdorff-Raum.  $T_2$  hat die Eigenschaft, dass jede konvergente Folge genau nur einen Grenzwert hat bzw. der Grenzwert ist eindeutig. Zum Beweis nehmen wir an eine Folge  $(x_n)$  konvergiert gegen zwei unterschiedliche Grenzwerte  $x$  und  $y$ . Dann existieren zwei Umgebungen  $O_X$  und  $O_Y$  um  $x$  und  $y$  für die es zwei unterschiedliche Indizes  $N_{O_X}$  und  $N_{O_Y}$  gibt, für die sich alle Folgenglieder danach in der jeweiligen Umgebung befinden. Aber für den Konvergenzbegriff brauchen wir nur einen Index  $N$ , also nehmen wir den größeren Wert der Beiden und nennen ihn  $N^*$ . D.h.  $\forall n \geq N^* : x_n \in O_X \wedge x_n \in O_Y \rightarrow x, y \in O_X \cap O_Y \rightarrow O_X \cap O_Y \neq \emptyset$ , was zur Folge hat, dass dieser Raum nicht  $T_2$  ist.

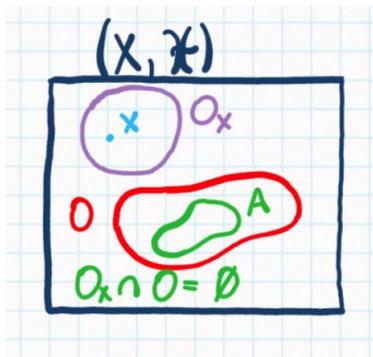


Abbildung 4.4: Zur Illustration des Trennungsaxiom  $T_3$ .

Ein Raum, der  $T_3$  erfüllt, wird auch regulär genannt.  $T_3$  sagt aus, dass jede abgeschlossene Menge  $A$  ein gewisses  $x$  nicht beinhaltet. Für beide Objekte kann man nun Umgebungen, offene Mengen, finden, für denen der Durchschnitt die Leere Menge ist.

Einen  $T_4$ -Raum nennt man auch normal. Um zwei beliebige abgeschlossenen Mengen  $A_1$  und  $A_2$  deren Durchschnitt die Leere Menge ist, gibt es zwei offene Umgebungen  $O_1$  und  $O_2$  um diese Mengen, deren Durchschnitt auch wieder die Leere Menge ist.

Nun sind wir endlich am Ende dieses Kapitel angekommen. Zur Rekapitulation: Wir haben

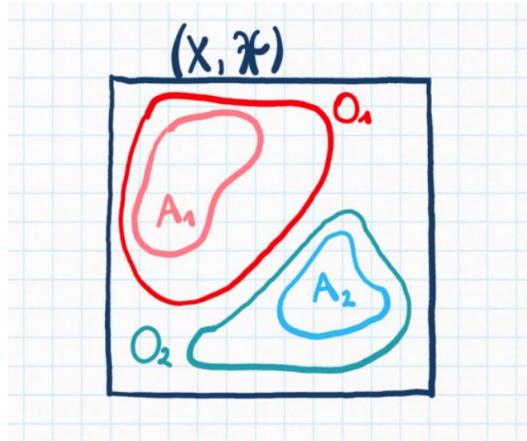


Abbildung 4.5: Zur Illustration des Trennungsaxiom  $T_4$ .

Topologien und topologische Räume betrachtet. Diese können als verallgemeinerte Vektorräume angesehen werden. Ein topologischer Raum hat eine Topologie  $\mathfrak{X}$ , was eine Menge der Teilmengen einer betrachteten Menge  $X$  ist. Diese Topologie muss drei Bedingungen genügen, um eine Topologie zu sein. Dann haben wir uns Abbildungen angeschaut und haben erkannt, dass man mit Abbildungen Topologien selber konstruieren kann. Darüber hinaus fanden wir den Begriff des Homöomorphismus und der Spurtopologie, einer Topologie für eine Teilmenge  $U$  von  $X$ , die von  $(X, \mathfrak{X})$  geerbt wurde. Als letztes haben wir die Trennungsaxiome analysiert und darin interessante Eigenschaften für die topologischen Räume gefunden. Nun können wir auf eine besonderen Art von topologischen Räumen hinbewegen, nämlich den Mannigfaltigkeiten.

## 4.4 Mannigfaltigkeiten

Halt halt, nicht so schnell, bevor wir zu den Mannigfaltigkeiten gehen können, müssen wir noch die Standardtopologie auf dem  $\mathbb{R}^n$  besprechen. Die Standardtopologie des  $\mathbb{R}^n$ ,  $\mathfrak{R}$  besteht aus dem Verbund von offenen Kugeln:  $K(x, \epsilon) = \{y \in \mathbb{R}^n \mid |x - y| < \epsilon\}$ , also die Menge aller Punkte  $y$ , für die der Abstand vom Mittelpunkt  $x$  kleiner als der Radius  $\epsilon$  ist. Nun behaupten wir, dass die Vereinigung solcher offenen Kugeln die offenen Mengen unserer Topologie  $\mathfrak{R}$  sind, also  $\mathfrak{R} \ni O \subseteq \mathbb{R}^n \leftrightarrow O = \cup_{\alpha} K_{\alpha}, K_{\alpha}(x_{\alpha}, \epsilon)$ . Jetzt müssen diese Vereinigungen unsere drei Bedingungen erfüllen, der Bequemlichkeit liste ich sie hier nochmals auf:

1.  $\emptyset, X \in \mathfrak{X}$
2.  $O_1, O_2 \in \mathfrak{X} \rightarrow O_1 \cap O_2 \in \mathfrak{X}$
3.  $(O_{\alpha}) \subseteq \mathfrak{X} \rightarrow \cup_{\alpha} O_{\alpha} \in \mathfrak{X}$

Die erste Bedingung können wir schnell beweisen: Eine Kugel mit Radius  $\epsilon = 0$  hat keine Punkte, also stellt sie die Leere Menge dar:  $\cup K(x, 0) = \cup \emptyset = \emptyset^6$ . Sei  $K(0, n)$  eine Kugel mit Mittelpunkt 0 und Radius gleich einer natürlichen Zahl  $n$ . Dann ist die Vereinigung solcher Kugeln mit immer größer werdenden Radius gleich dem  $\mathbb{R}^n$ :  $x \in \mathbb{R}^n : |x| = d, \exists n \in \mathbb{N} : n > d \rightarrow x \in K(0, n) \rightarrow \cup_n K(0, n) = \mathbb{R}^n$ . Wir sagen also, dass es für jede beliebige Zahl  $x$  eine natürliche Zahl  $n$  gibt, für die der Betrag von  $x$  kleiner ist als  $n$ . Dann wissen wir, dass  $x$  innerhalb der Kugel mit Radius  $n$  ist. D.h. wir können so den gesamten  $\mathbb{R}^n$  so aufspannen.

Die zweite Bedingung ist nicht so trivial wie die erste. Denn allgemein ist der Durchschnitt von zwei Kugeln keine Kugel. Zuerst müssen wir beweisen, dass jeder Punkt innerhalb des Durchschnitts zweier Kugeln in einer Kugel mit Radius  $\epsilon^*$  ist:

Sei  $x_\alpha$  und  $x_\beta$  die Mittelpunkte zweier Kugeln  $K_\alpha(x_\alpha, \epsilon_\alpha)$  und  $K_\beta(x_\beta, \epsilon_\beta)$ , und deren Durchschnitt ist nicht 0:  $K_\alpha \cap K_\beta \neq \emptyset$ . Dann definieren wir  $\eta_\alpha = |x - x_\alpha| < \epsilon_\alpha$  und  $\eta_\beta = |x - x_\beta| < \epsilon_\beta$  für beliebige  $x \in K_\alpha \cap K_\beta$ . Dann wählen wir zwischen den zwei Termen  $\epsilon_\alpha - \eta_\alpha$  und  $\epsilon_\beta - \eta_\beta$  den kleineren Wert, also  $\min\{\epsilon_\alpha - \eta_\alpha, \epsilon_\beta - \eta_\beta\} = \epsilon^*$ . Sei nun  $K(x, \epsilon^*)$  eine Kugel im Durchschnitt mit unserem Radius  $\epsilon^*$  und als Mittelpunkt ein beliebiger Punkt im Durchschnitt. Dann gilt

$K(x, \epsilon^*) \subseteq K_\alpha \cap K_\beta, \forall x \in K_\alpha \cap K_\beta$ . Nun betrachten wir ein beliebiges  $z \in K(x, \epsilon^*)$ . Für dieses  $z$  gilt:  $|z - x_\alpha| = |z - x + x - x_\alpha| \leq |z - x| + |x - x_\alpha| = |z - x| + \eta_\alpha < \epsilon^* + \eta_\alpha = \epsilon_\alpha - \eta_\alpha + \eta_\alpha = \epsilon_\alpha$ . Hier wurde angenommen, dass  $\epsilon^* = \epsilon_\alpha - \eta_\alpha$  ist, wenn dem nicht der Fall ist, kann die Ungleichung

für  $x_\beta$  analog gemacht werden. Nun haben wir gezeigt, dass  $|z - x_\alpha| < \epsilon_\alpha$  gilt, also, dass  $z$  innerhalb von  $K_\alpha$  ist. Doch wenn das für  $K_\alpha$  gilt, gilt das auch für  $K_\beta$ ! D.h.

$$z \in K_\alpha \cap K_\beta \rightarrow K(x, \epsilon^*) \subseteq K_\alpha \cap K_\beta.$$

Nun müssen wir beweisen, dass die Vereinigung von diesen Kugeln  $K(x, \epsilon^*)$  den Durchschnitt aufspannen:

$$\begin{aligned} \text{Sei } x \in K_\alpha \cap K_\beta \rightarrow \{x\} \subseteq K(x, \epsilon^*) \subseteq K_\alpha \cap K_\beta, \forall z \in K(x, \epsilon^*) \rightarrow K_\alpha \cap K_\beta = \cup_{x \in K_\alpha \cap K_\beta} \{x\} \subseteq \\ \cup_{x \in K_\alpha \cap K_\beta} K(x, \epsilon^*) \leftrightarrow K_\alpha \cap K_\beta = \cup_{x \in K_\alpha \cap K_\beta} K(x, \epsilon^*) \rightarrow O_1 \cap O_2 = (\cup_\alpha K_\alpha) \cap (\cup_\beta K_\beta) = \\ \cup_\alpha K_\alpha \cap K_\beta = \cup K(x, \epsilon^*) \in \mathfrak{A} \end{aligned}$$

Die dritte Bedingung ist wieder leichter: Die Vereinigung von beliebig vereinten offenen Kugeln ist natürlich auch eine Vereinigung von Kugeln, d.h. sie sind Elemente von  $\mathfrak{A} : (O_\alpha) \subseteq \mathfrak{A} \rightarrow \cup_\alpha O_\alpha = \cup_\alpha \cup_\beta (K_\beta)_\alpha \in \mathfrak{A}$ .

Nun sind wir endlich in der Lage zu beschreiben, was eine Mannigfaltigkeit ist: Eine Mannigfaltigkeit  $(M, \mathfrak{M})$  ist ein topologischer Raum, der auch wieder drei Bedingungen erfüllen muss:

1. Er muss  $T_2$  sein.
2. Er muss parakompakt sein.

---

<sup>6</sup>Wieso ist nicht der Punkt 0 innerhalb solch einer Kugel? Weil wir fordern, dass  $|x - y| < \epsilon$  gilt, und  $|0 - 0| \not< 0$ .

3. Er muss lokal homöomorph zum topologischen Raum  $(\mathbb{R}^n, \mathfrak{A})$  sein.

Eine wichtige Folgerung dieser Eigenschaften ist, dass die Mannigfaltigkeit um einen Punkt *topologisch ununterscheidbar* zum  $\mathbb{R}^n$  ist. Was bedeutet nun, dass er lokal homöomorph zum  $\mathbb{R}^n$  ist? D.h. es gibt um einen Punkt  $p$ <sup>7</sup> eine *Kartenabbildung*  $\phi$ , die die Umgebung  $O$  des Punktes in den  $\mathbb{R}^n$  bijektiv abbilden kann: zu jedem  $p \in M \exists O \in \mathfrak{M} : p \in O \wedge \phi : O \rightarrow \mathbb{R}^n$ , sodass  $(O, \mathfrak{M} \cap O)$  homöomorph zu  $(\mathbb{R}^n, \mathfrak{A})$  ist. Nun können wir uns vorstellen, wie in Abb. 4.6, dass es zwei Umgebungen um den Punkt  $p$  gibt,  $O$  und  $\tilde{O}$ . Nun gibt es dann zwei unterschiedliche Abbildungen  $\phi : O \rightarrow \mathbb{R}^n$  und  $\tilde{\phi} : \tilde{O} \rightarrow \mathbb{R}^n$ . Da diese zwei Umgebungen einen nicht-verschwindenden Durchschnitt haben, können wir sicherlich eine Abbildung  $\psi$  definieren, die von einem  $\mathbb{R}^n$  in den anderen  $\mathbb{R}^n$  abbilden kann, wobei diese nicht die Identitätsabbildung ist, denn allgemein sind die Bilder von  $\phi$  und  $\tilde{\phi}$  unterschiedlich.

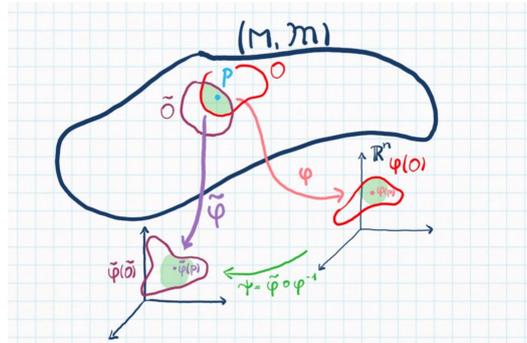


Abbildung 4.6: Zur Illustration der Kartenwechselabbildung.

In Abb. 4.7 ist dieser Vorgang illustriert. Auf unserer Mannigfaltigkeit  $M$  gibt es zwei Abbildungen  $\phi$  und  $\tilde{\phi}$  und die *Kartenwechselabbildung*  $\psi$  ist definiert als:  $\psi = \tilde{\phi} \circ \phi^{-1}$ . Man kann sich das auch bildhaft wirklich vorstellen: Als erstes ist man im  $\mathbb{R}^n$  und kann durch die stetige Inverse von  $\phi$ ,  $\phi^{-1}$ , in die Mannigfaltigkeit abgebildet werden. Dann kann man im Durchschnitt wieder mit  $\tilde{\phi}$  zurück in den  $\mathbb{R}^n$  abbilden. Was ist wenn wir drei Umgebungen haben, die alle einen nicht-verschwindenden Durchschnitt haben? Das nennt man die Dreifach-Durchschnittsbedingung und ist in Abb. 4.7 dargestellt. Nun haben wir drei Umgebungen um  $p$ ,  $O_\alpha, O_\beta$  und  $O_\gamma$ , und drei Kartenabbildungen,  $\phi_\alpha, \phi_\beta$  und  $\phi_\gamma$ . Nun fordern wir, dass die Vereinigung der Kartenwechselabbildungen  $\psi_{\beta\alpha}$  und  $\psi_{\gamma\beta}$  gleich ist, wie die Kartenwechselabbildung  $\psi_{\gamma\alpha}$ :  $\psi_{\gamma\alpha} = \psi_{\gamma\beta} \circ \psi_{\beta\alpha}$ . Umgangssprachlich bedeutet das, dass es keinen Unterschied macht, ob wir von  $\phi_\alpha(O_\alpha \cap O_\beta \cap O_\gamma)$  nach  $\phi_\beta(O_\alpha \cap O_\beta \cap O_\gamma)$  und dann nach  $\phi_\gamma(O_\alpha \cap O_\beta \cap O_\gamma)$  gehen, oder gleich von  $\phi_\alpha(O_\alpha \cap O_\beta \cap O_\gamma)$  nach  $\phi_\gamma(O_\alpha \cap O_\beta \cap O_\gamma)$  wechseln. Diese Bedingung ist leicht zu beweisen, wir gehen von der Definitionen der Kartenabbildungen aus:  $\psi_{\gamma\beta} \circ \psi_{\beta\alpha} = \phi_\gamma \circ \phi_\beta^{-1} \circ \phi_\beta \circ \phi_\alpha^{-1} = \phi_\gamma \circ id \circ \phi_\alpha^{-1} = \psi_{\gamma\alpha}$ . Diese Dreifach-Bedingung kann beliebig oft iteriert werden, sodass sie für 4 und 5-fache Durchschnitte auch gilt, aber Dreifacher Durch-

<sup>7</sup>Also nicht global!

schnitt reicht uns.  $\mathcal{A}$  heißt Atlas von  $M$  und ist eine Ansammlung von Kartenabbildungen und deren zugehörige Umgebung  $O_\alpha$ , wobei  $\cup_\alpha O_\alpha = M$  gilt.

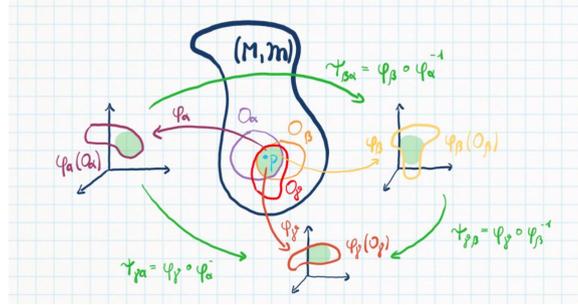


Abbildung 4.7: Zur Illustration der Dreifach-Durchschnittsbedingung.

Nun wollen wir skalare Funktionen betrachten. Hierfür werden wir die Bilder unserer Kartenabbildungen, also die Koordinaten in einem  $\mathbb{R}^n$ , als  $x_\alpha^i$  bzw.  $x_\beta^i$  bezeichnen, wobei  $i$  der Index der  $i$ -ten Komponente unseres  $n$ -dimensionalen  $\mathbb{R}^n$  ist und  $\alpha$  bzw.  $\beta$  kennzeichnet unsere Kartenabbildung in den  $\mathbb{R}^n$ . Nun betrachten wir wirklich eine skalare Funktion  $F: M \rightarrow \mathbb{R}$ , also eine Funktion, die von der Mannigfaltigkeit in die reellen Zahlen abbildet. Wollten wir nun eine skalare Funktion, die einen Vektor des  $\mathbb{R}^n$  als Input hat, konstruieren, sodass wir die Eigenschaften der Mannigfaltigkeit auch erben, könnten wir diese skalare Funktion  $f_\alpha$  folgendermaßen definieren:  $f_\alpha: \mathbb{R}^n \rightarrow \mathbb{R}; f_\alpha(x_\alpha^i) = (F \circ \phi_\alpha^{-1})(x_\alpha^i)$ . Somit haben wir die Möglichkeit, vom Punkt im  $\mathbb{R}^n$  in die Mannigfaltigkeit  $M$  zu wechseln und dann damit einen Skalar zu produzieren. Natürlich hängt  $f_\alpha$  damit mit der Kartenabbildung ab, die wir wählen. Im Durchschnitt ist nicht gegeben, welche wir nehmen sollen.

Nun betrachten wir wieder die Koordinaten  $x_\alpha^i$  bzw.  $x_\beta^i$ . Diese zwei Koordinatensysteme müssen doch irgendwie zusammenhängen, wenn wir den Durchschnitt der Urbilder betrachten. Seien  $f_\alpha$  und  $f_\beta$  zwei Skalare Funktionen, so wie wir sie zuvor definiert haben, wobei  $x_\alpha^i \in \phi_\alpha(O_\alpha \cap O_\beta)$  bzw.  $x_\beta^i \in \phi_\beta(O_\alpha \cap O_\beta)$  gilt. D.h. es muss eine Kartenwechselabbildung  $\psi_{\beta\alpha}$  geben, mit der man zwischen den Koordinatensystemen wechseln kann. Betrachten wir nun die Funktion  $f_\beta(x_\beta^i)$ . Dann gilt:

$$f_\beta(x_\beta^i) = f_\beta(\psi_{\beta\alpha}(x_\alpha^k)) = (f_\beta \circ \psi_{\beta\alpha})(x_\alpha^k). \quad (4.3)$$

Nun können wir die Definitionen von  $f_\beta$  und  $\psi_{\beta\alpha}$  verwenden:

$$(f_\beta \circ \psi_{\beta\alpha})(x_\alpha^k) = (F \circ \phi_\beta^{-1} \circ \psi_{\beta\alpha} \circ \phi_\alpha^{-1})(x_\alpha^k) = (F \circ \phi_\alpha^{-1})(x_\alpha^k) = f_\alpha(x_\alpha^k). \quad (4.4)$$

Insgesamt können wir also schreiben:

$$f_\beta(x_\beta^i) = f_\alpha(x_\alpha^k), \forall x_\beta^i \in \phi_\beta(O_\alpha \cap O_\beta) \wedge x_\alpha^k \in \phi_\alpha(O_\alpha \cap O_\beta). \quad (4.5)$$

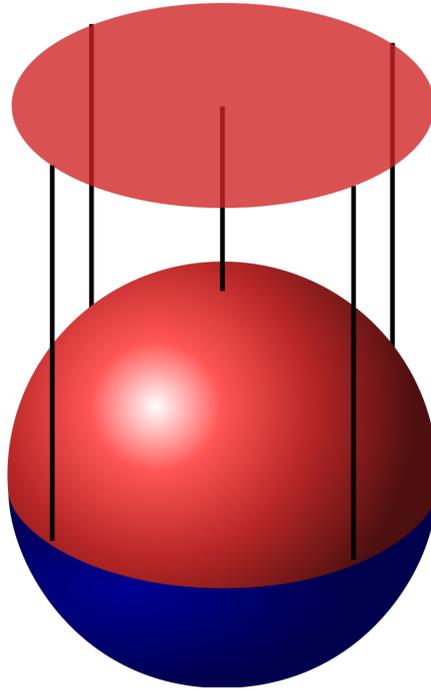


Abbildung 4.8: Zur Illustration der Kugel als Mannigfaltigkeit.

D.h. zwei skalare Funktionen angewandt auf Koordinaten im Durchschnitt der Mannigfaltigkeit sind gleich. Mit einer neuen Notation  $x_\beta^i = \psi_{\beta\alpha}(x_\alpha^k) = x_\beta^i(x_\alpha^k)$  können wir die Gleichheit noch etwas umschreiben:

$$f_\beta(x_\beta^i(x_\alpha^k)) = f_\alpha(x_\alpha^k). \quad (4.6)$$

### Mathematical Interlude: Die Kugel als Mannigfaltigkeit

In Abb. 4.8 sieht man eine Kugel, deren obere Halbkugel auf den  $\mathbb{R}^2$  projiziert wird. Die Abbildungsfunktion für eine Kugel mit  $f(x,y,z) = x^2 + y^2 + z^2 - 1 = 0$  nimmt nur die  $x$ - und  $y$ -Werte:  $\phi(x, y, z) = (x, y)^T$ . Aber das ist nicht die einzige Art, wie man die Kartenabbildung konstruieren kann. Man könnte auch wie mit einer gewöhnlichen Karte vorgehen: In Abb. 4.9 ist dieses Vorgehen erläutert. Man bildet die ganze Kugel auf einen  $\mathbb{R}^2$  ab, wobei der Nord/Südpol jeweils nicht inkludiert werden kann. Bildlich gesprochen nimmt man die Erde und schneidet den Nord/Südpol hinaus, dann drückt man diese offene Fläche platt und versucht den Rand, der um dem Nord/Südpol ist, auszuweiten. Dann erhält man Abb. 4.9 und deshalb ist diese Abbildung der Erde so verzerrt, da je weiter man vom Süd/Nordpol weg ist, desto gestreckter werden die Linien. Zur weiteren Behandlung der Kugel als Mannigfaltigkeit werden wir die zweite Methode der Kartenabbildung verwenden, denn wir wollen nun die Kartenwechselabbildung herleiten.

Man betrachte Abb. 4.10 in der die Projektion der Punkte auf der Kugel in den  $\mathbb{R}^2$  dargestellt ist, wobei die Koordinaten, welche den Südpol nicht inkludieren mit  $\tilde{x}^i$  und welche die den Nordpol

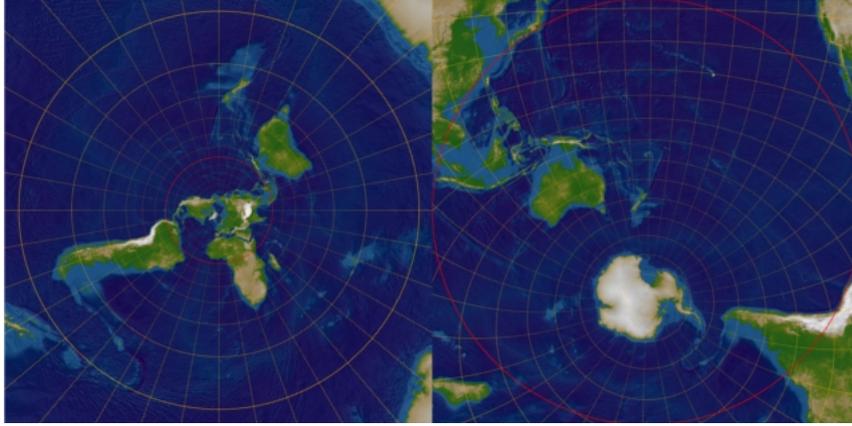


Abbildung 4.9: Zur Illustration der möglichen Kartenabbildungen.

nicht inkludieren mit  $x^i$  bezeichnen. Zur Herleitung der Kartenwechselabbildung soll nun Abb. 4.11 betrachtet werden, diese stellt nur eine 2D Ansicht von Abb. 4.10 dar. Zur Vereinfachung wird der Radius der Kugel mit 1 normiert.  $\rho$  und  $\tilde{\rho}$  bezeichnen nun den Abstand vom Süd/Nordpol in den jeweiligen  $\mathbb{R}^2$ . D.h. es gilt  $\rho = \sqrt{x^i x^i}$  und  $\tilde{\rho} = \sqrt{\tilde{x}^i \tilde{x}^i}$ . Da jeder Punkt bijektiv auf den jeweiligen  $\mathbb{R}^2$  abgebildet wird, können wir  $\rho$  und  $\tilde{\rho}$  miteinander geometrisch in Beziehung setzen. Aus Abb. 4.11 kann man herauslesen, dass  $\tan \alpha = \frac{\rho}{1}$  und  $\tan \beta = \frac{\tilde{\rho}}{1}$  gilt. Da  $\beta = \frac{\pi}{2} - \alpha$  gilt, da diese ja ein Dreieck bilden, können wir schreiben:

$$\tilde{\rho} = \tan \beta = \frac{\sin(\frac{\pi}{2} - \alpha)}{\cos(\frac{\pi}{2} - \alpha)} = \frac{\cos \alpha}{\sin \alpha} = \frac{1}{\rho}. \quad (4.7)$$

Somit haben wir eine Beziehung zwischen  $\rho$  und  $\tilde{\rho}$  gefunden! Nun können wir damit die Kartenwechselabbildung herleiten. Mit Polarkoordinaten können wir mit  $\rho$  und  $\tilde{\rho}$  in die Koordinaten  $\tilde{x}^i$  und  $x^i$  umrechnen, denn diese sind ja die „normalen“ kartesischen Koordinaten für den jeweiligen  $\mathbb{R}^2$ . Es gilt:

$$\tilde{x}^i = \tilde{\rho} e^i(\tilde{\phi}) = \frac{1}{\rho} e^i(\tilde{\phi}) = \frac{\rho}{\rho^2} e^i(\phi) = \frac{x^i}{\rho^2} \quad (4.8)$$

Das ist unsere Kartenwechselabbildung von  $x^i$  nach  $\tilde{x}^i$ , die uns erlaubt von einer (buchstäblichen) Karte in die andere zu wechseln. Es ist hier nur Zufall, dass die Inverse, also von  $\tilde{x}^i$  nach  $x^i$  dieselbe Form hat:

$$x^i = \frac{\tilde{x}^i}{\tilde{x}^k \tilde{x}^k} \quad (4.9)$$

Allgemein sind die Kartenwechselabbildungen beliebig komplizierter Natur. Diese Kartenwechselabbildungen haben den Wertebereich des  $\mathbb{R}^2$ , ohne den jeweiligen Ursprung:  $\tilde{x}^i(x^k) : \mathbb{R}^2 - \{O\} \rightarrow$

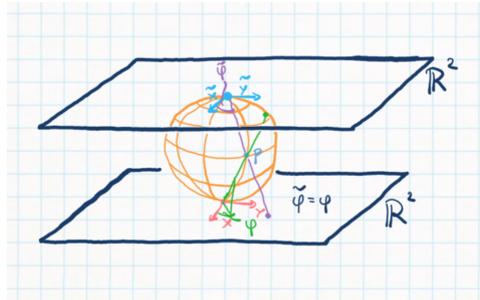


Abbildung 4.10: Zur Illustration der Herleitung der Kartenwechselabbildung.

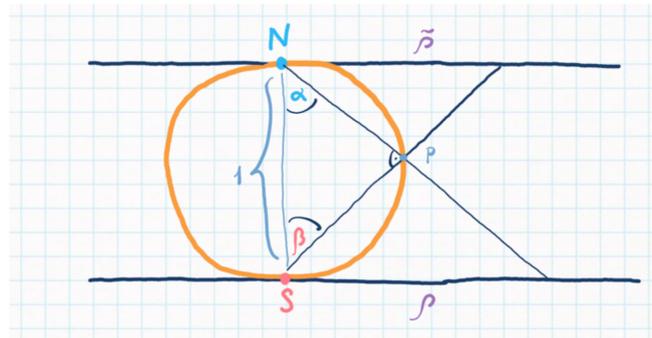


Abbildung 4.11: Zur Illustration der Herleitung der Kartenwechselabbildung.

$\mathbb{R}^2 - \{O\}$  und umgekehrt. Nun wissen wir auch, wie eine Funktion  $f(x^i)$  in den anderen Koordinaten aussehen würde. Sei z.B.  $f(x^i) = \frac{1}{x^2 + y^2 + 1}$ , dann könnte man diese Funktion zu  $\tilde{f}(\tilde{x}^i)$  umschreiben:

$$f(x^i) = \frac{1}{x^2 + y^2 + 1} = \frac{1}{\rho + 1} = \frac{1}{\frac{1}{\tilde{\rho}^2} + 1} = \frac{\tilde{\rho}^2}{1 + \tilde{\rho}^2} = \tilde{f}(\tilde{x}^i) \quad (4.10)$$

Wie man sehen kann, ist die Kartenwechselabbildung nicht linear und unendlich oft differenzierbar.

#### 4.4.1 Vektorfelder auf Mannigfaltigkeiten

Um Vektorfelder auf Mannigfaltigkeiten zu betrachten, müssen wir uns noch mit Kurven und Geraden auf Mannigfaltigkeiten beschäftigen. Sei  $x^i(t)$  eine Kurve im normalen  $\mathbb{R}^n$  mit dem Kurvenparameter  $t$ . Dann benennen wir den Punkt im  $\mathbb{R}^n$  mit Kurvenparameter  $t = 0$  als  $x_0^i = x^i(t = 0)$ . Nun definieren wir den Tangentenvektor an dieser Kurve als  $v^i = \frac{dx^i}{dt} \Big|_{t=0}$ . Wir wissen aber vom vorherigen Abschnitt, dass wir diese Kurve in einem anderen Koordinatensystem  $\tilde{x}^i$  beschreiben können, solange diese Kurve im Durchschnitt zweier Umgebungen um einen Punkt auf der Mannigfaltigkeit existiert<sup>8</sup>. Ist dies der Fall können wir  $\tilde{x}^i(t) = \tilde{x}^i(x^k(t))$  schreiben, wobei die Nota-

<sup>8</sup>Denn hier existiert die Kartenwechselabbildung!

tion meint, dass wir hier die Kartenwechselabbildung anwenden, um jeden Kurvenpunkt im  $x^i$ -Koordinatensystem auf einen Punkt im  $\tilde{x}^i$ -Koordinatensystem abzubilden. Wir schreiben für den Punkt an  $t=0$   $\tilde{x}^i(x^k(0)) = \tilde{x}^i(0) = \tilde{x}_0^i$ .

Nun können wir uns die Frage stellen, wie sieht nun der Tangentenvektor im  $\tilde{x}^i$ -Koordinatensystem aus? Dazu definieren wir wieder, dass  $\tilde{v}^i = \frac{d\tilde{x}^i}{dt}$  gelten soll. Für  $\tilde{x}^i$  können wir nun unsere Beziehung einsetzen und die Kettenregel anwenden, also innere Ableitung mal äußere Ableitung:

$$\tilde{v}^i = \frac{d\tilde{x}^i}{dt} = \frac{\tilde{x}^i(x^k(t))}{dt} = \frac{d\tilde{x}^i}{dx^k} \cdot \frac{dx^k}{dt} = \frac{d\tilde{x}^i}{dx^k} v^k \quad (4.11)$$

Weil diese Beziehung so wichtig ist, bzw. sie wird oft in der Literatur als bedeutsame Transformation angepriesen, schreibe ich sie hier noch einmal an:

$$\tilde{v}^i = \frac{d\tilde{x}^i}{dx^k} v^k \quad (4.12)$$

Das ist das Transformationsgesetz für die *Komponenten* eines Vektoren. Objekte, die so transformieren heißen *kontravariant*. Also man bildet die Ableitung der Kartenwechselfunktion nach der jeweiligen Koordinate und bildet dann die Summe mit dem Vektor. Nun haben wir eine *lineare* Transformation von einem Koordinatensystem ins andere gefunden<sup>9</sup>! Und wenn etwas linear ist, können wir sehr gut weiter damit arbeiten. Denn nun definieren wir die Addition von zwei Vektoren unter der Transformation.

Seien  $v^i$  und  $w^i$  zwei Tangentenvektoren bei  $x_0^i$ . Da die Kurve völlig willkürlich sein kann<sup>10</sup>, bauen wir uns nun die einfachste Kurve, die wir finden können, deren Tangentenvektor  $v^i + w^i$  lautet. Diese ist trivialerweise  $x^i(t) = x_0^i + t(v^i + w^i)$ . Nun können wir die Transformation ausführen:

$$(\tilde{v} + \tilde{w})^i = \frac{d\tilde{x}^i}{dt} = \frac{d\tilde{x}^i}{dx^l} (v^l + w^l) = \tilde{v}^i + \tilde{w}^i. \quad (4.13)$$

Als kurze Randnotiz: Wir nennen Kurven, die für sich verschieden verlaufen, aber den Anfangswert  $x_0^i$  gleich haben als Äquivalenzklassen. Also für die gilt:

$$x_1^i(t=0) = x_0^i = x_2^i(t=0) \rightarrow \frac{d}{dt} x_1^i(t=0) = v^i = \frac{d}{dt} x_2^i(t=0), \quad (4.14)$$

d.h. ihr Tangentenvektor ist derselbe.

Jetzt wissen wir wie die Komponenten eines Vektors unter der Transformation sich verhalten, wie schaut's mit den Basisvektoren selbst aus? Seien  $E_l^a$  und  $\tilde{E}_l^a$  Basen für das jeweilige Koordinatensystem. Dann lautet das Transformationsgesetz für Basisvektoren:

$$\tilde{E}_l^a = \frac{dx^k}{d\tilde{x}^l} E_k^a. \quad (4.15)$$

<sup>9</sup>Und sie ist nichtmal abhängig von der gewählten Kurve, jede Kurve für die  $x^i(t=0) = x_0^i$  gilt, ist benutzbar.

<sup>10</sup>Nur der Wert  $x^i(t=0) = x_0^i$  ist wichtig!

Objekte, die so transformieren heißen *kovariant*. Diese Beziehung kann man verifizieren, indem man erkennt, dass der Vektor  $v^a$  in jedem Koordinatensystem gleich sein muss;  $v^a = \tilde{v}^i \tilde{E}_i^a = v^k E_k^a$ :

$$v^a = \tilde{v}^i \tilde{E}_i^a = \frac{d\tilde{x}^i}{dx^k} \frac{dx^m}{d\tilde{x}^i} v^k E_m^a = \delta_k^m v^k E_m^a = v^k E_k^a. \quad (4.16)$$

Der nächste Schritt wird nicht sehr schlüssig wirken. Denn wir wollen nun neue Basisvektoren, die das Transformationsgesetz zwischen Kartenabbildungen schon von Natur aus haben. Wie soll das gehen, eine Basis ist doch willkürlich, aber die Transformationseigenschaft ändert sich ja nicht. Deshalb müssen wir auf ganz andere Objekte zurückgreifen. Unsere neuen Basisvektoren sollen nun *Ableitungsoperatoren* sein. Was? Wie soll ein Ableitungsoperator ein Vektor sein? Diese Frage wird zunächst ignoriert und wir betrachten das Transformationsverhalten von  $\frac{\partial}{\partial x^i} = \partial_i$ :

$$\tilde{\partial}_i = \frac{\partial}{\partial \tilde{x}^i} = \frac{\partial x^k}{\partial \tilde{x}^i} \frac{\partial}{\partial x^k} = \frac{dx^k}{d\tilde{x}^i} \partial_k \quad (4.17)$$

Man erkennt, die Ableitungsoperatoren haben dasselbe Transformationsverhalten, wie unsere Basisvektoren. Man kann auch beweisen, dass die Ableitungsoperatoren einen Vektorraum aufspannen und dadurch eine Basis besitzen. Deshalb werden wir von nun an schreiben:

$$v^a = v^i \partial_i^a = \tilde{v}^i \tilde{\partial}_i^a. \quad (4.18)$$

Nun können wir den *Tangentenraum*  $T_p M$  definieren:  $T_p M: \{v^a | v^a = v^i \partial_i^a \text{ in einer Karte die } p \text{ beinhaltet}\}$ . Das heißt im Tangentenraum befinden sich alle Tangentenvektoren  $v^a$  von Kurven, welche tangential im Punkt  $p$  liegen. Man kann sich vorstellen, eine Ebene an einen Punkt auf einer Kugel und normal zur Oberfläche zu platzieren. Diese Ebene wird durch die Tangentialvektoren am Punkt aufgespannt. Abb. 4.12 illustriert diese Idee.

Wenn man einen Vektorraum hat, dann bekommt man immer was gratis dazu: den Dualen-Raum  $(T_p M)^*$ . Seine Elemente sind lineare Funktionen deren Input die Tangentialvektoren des  $T_p M$  sind. Das Transformationsverhalten für die Komponenten der Dualen Vektoren  $\omega_k$  (und ihren Basisvektoren  $e_a^k$ ) lautet wie folgt:

$$\tilde{\omega}_i = \frac{\partial x^k}{\partial \tilde{x}^i} \omega_k, \quad (4.19)$$

$$\tilde{e}_a^i = \frac{\partial \tilde{x}^k}{\partial x^i} e_a^k. \quad (4.20)$$

Man kann erkennen, dass sich die dualen Komponenten wie die normalen Basisvektoren transformieren und die dualen Basisvektoren wie die Komponenten der normalen Basisvektoren. D.h. die dualen Komponenten sind kovariant und die dualen Basisvektoren sind kontravariant. Wie zuvor wollen wir aber wieder Objekte, die dieses Transformationsverhalten haben, als unsere Basisvektoren. Dies erhält

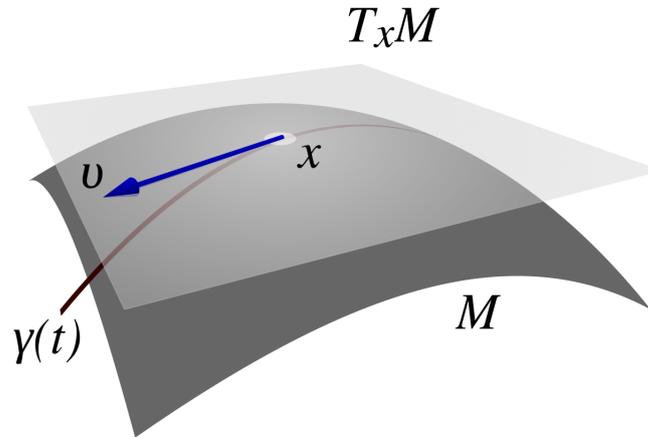


Abbildung 4.12: Zur Illustration des Tangentenraums.

| Transformationsverhalten | Objekte                                     | Transformationsgesetz  |
|--------------------------|---|--|
| kovariant                | (Basis-)Vektoren, duale Komponenten         | $\tilde{\partial}_i^a = \frac{dx^k}{d\tilde{x}^i} \partial_k^a; \tilde{\omega}_i = \frac{\partial x^k}{\partial \tilde{x}^i} \omega_k$ |
| kontravariant            | normale Komponenten, duale (Basis-)Vektoren | $\tilde{v}^i = \frac{d\tilde{x}^i}{dx^k} v^k; d\tilde{x}_a^i = \frac{\partial \tilde{x}^i}{\partial x^k} dx_a^k$                       |

Tabelle 4.1: Zu den verschiedenen Transformationsgesetzen.

man mit absoluten Differentialen:

$$d\tilde{x}^i = \frac{\partial \tilde{x}^i}{\partial x^k} dx^k. \quad (4.21)$$

Dadurch werden diese absoluten Differentiale unseren neuen dualen Basisvektoren darstellen:

$$\omega_a = \omega_i dx_a^i = \tilde{\omega}_i d\tilde{x}_a^i. \quad (4.22)$$

Diese Differentiale bilden auch einen Vektorraum, deshalb ist dieses Vorgehen valide. Weil diese Überlegungen recht verwirrend sind, werden hier nochmals die Transformationsgesetze von Vektoren und dualen Vektoren in Tabelle 4.1 aufgelistet.

Trivialer Weise gilt folgende Beziehung:

$$dx_a^i \partial_j^a = \delta_j^i. \quad (4.23)$$

Mit Vektoren und dualen Vektoren können wir nun allgemeine *Tensoren* auf ihrem Transformationsverhalten analysieren. Ein Tensor  $t_{b_1, \dots, b_s}^{a_1, \dots, a_r}$  ist ein Element des Tensor-Tangentenraums  $\otimes_s^r T_p M$ .

In der neuen Notation lautet die Basisdarstellung:

$$t_{b_1, \dots, b_s}^{a_1, \dots, a_r} = t_{j_1, \dots, j_s}^{i_1, \dots, i_r} \partial_{i_1}^{a_1} \dots \partial_{i_r}^{a_r} dx_{b_1}^{j_1} \dots dx_{b_s}^{j_s} \quad (4.24)$$

Als Zusammenfassung: Über Kurven, die einen gewissen Punkt gemeinsam haben, konnten wir den Tangentialvektor definieren. Die Ableitung nach dem Kurvenparameter ist linear, wodurch Addition und Skalarmultiplikation gegeben ist. Die Verbindung zur Mannigfaltigkeit entsteht dadurch, dass wir die Kurve im  $\mathbb{R}^n$  als ein Bild unter der Kartenabbildung einer Kurve auf der Mannigfaltigkeit ansehen. Durch Kartenwechselfunktionen konnten wir danach das Transformationsverhalten unserer Tangentenvektoren analysieren, wodurch wir den Tangentenraum um den Punkt  $p \in M$  erhielten. Daraus folgten die Dualen Vektoren, sowie die Tensoren. Nun werden wir uns einem ähnlichen Thema mit einem Beispiel anschauen: Vektorfelder auf Mannigfaltigkeiten.

Ein Vektorfeld  $v^a$  auf  $M$  sei wie folgt definiert:  $v^a : p \in M \rightarrow v^a(p) \in T_p M, \forall p \in M$ . Also statt nur ein bestimmtes  $p$  auf der Mannigfaltigkeit zu betrachten, werden hier alle Punkte miteinbezogen und jedem ein Vektor im zugehörigen Tangentenraum zugeordnet. Man kann das auch für Tensorfelder verallgemeinern:  $t_{b_1, \dots, b_s}^{a_1, \dots, a_r} : p \in M \rightarrow t_{b_1, \dots, b_s}^{a_1, \dots, a_r}(p) \in \otimes_s T_p M$ . Also ein Vektor/Tensorfeld ordnet jedem Punkt auf der Mannigfaltigkeit einen Vektor/Tensor im entsprechenden Tangenten(Tensor)raum über dem Punkt zu. Um das zu illustrieren wird nun ein kleines Rechenbeispiel durchgerechnet.

### Mathematical Interlude: Strömungen auf der Kugel

Wir betrachten wieder die Kugel als Mannigfaltigkeit an, mit unseren bekannten Kartenabbildungen. Die Koordinatensysteme sollen wie folgt mit  $x^i$  und  $\tilde{x}^i$  bezeichnet. Die Kartenwechselabbildung lautet bekannterweise:  $\tilde{x}^i = \tilde{x}^i(x^m) = \frac{x^i}{x^k x^k}$ . Nun betrachten wir ein Vektorfeld, ein recht einfaches, aber selbst damit kann man schon vieles, was eine Mannigfaltigkeit ausmacht, erkennen. Es gilt für unser Vektorfeld:

$$v^a = x^i \partial_i^a. \quad (4.25)$$

In Abb. 4.13 wird dieses Vektorfeld illustriert. Es zeigt vom Ursprung weg, und der Betrag der Vektoren wird immer größer, je weiter weg man sich vom Ursprung befindet. Nun sind wir daran interessiert, wie das Vektorfeld im anderen Koordinatensystem aussieht. Dafür müssen wir nur das Transformationsgesetz für die Basisvektoren und die Kartenwechselfunktion für die Komponenten anwenden und schauen was herauskommt:

$$v^a = x^i \partial_i^a \quad (4.26)$$

$$v^a = \frac{\tilde{x}^i}{\tilde{x}^k \tilde{x}^k} \frac{d\tilde{x}^j}{dx^i} \tilde{\partial}_j^a \quad (4.27)$$

Nun müssen wir noch die Ableitungen der Kartenwechselfunktion  $\frac{d\tilde{x}^j}{dx^i}$  berechnen:

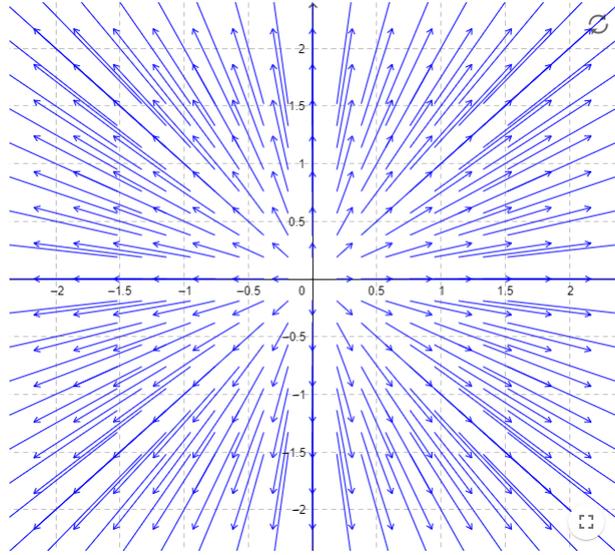


Abbildung 4.13: Zur Illustration des Vektorfeldes  $v^a$ .

$$\frac{d\tilde{x}^j}{dx^i} = \frac{d\tilde{x}^j(x^m)}{dx^i} \quad (4.28)$$

$$\frac{d}{dx^i} \frac{x^j}{x^k x^k} = \frac{\delta_i^j}{x^k x^k} - \frac{2x^i x^j}{(x^k x^k)^2} = \quad (4.29)$$

$$\frac{d\tilde{x}^j}{dx^i} = \frac{1}{x^k x^k} (\delta_i^j - \frac{2x^i x^j}{x^k x^k}) \quad (4.30)$$

Nun müssen wir diese Ableitung noch in das Vektorfeld einsetzen:

$$v^a = \frac{\tilde{x}^i}{\tilde{x}^k \tilde{x}^k} \frac{d\tilde{x}^j}{dx^i} \tilde{\partial}_j^a \quad (4.31)$$

$$v^a = \frac{\tilde{x}^i}{\tilde{x}^k \tilde{x}^k} \frac{1}{x^k x^k} (\delta_i^j - \frac{2x^i x^j}{x^k x^k}) \tilde{\partial}_j^a \quad (4.32)$$

$$v^a = \frac{\tilde{x}^i}{\tilde{x}^k \tilde{x}^k} \frac{\tilde{x}^k \tilde{x}^k}{1} (\delta_i^j - \frac{2x^i x^j}{x^k x^k}) \tilde{\partial}_j^a \quad (4.33)$$

$$v^a = \tilde{x}^i (\delta_i^j - \frac{2x^i x^j}{x^k x^k}) \tilde{\partial}_j^a \quad (4.34)$$

$$v^a = (\tilde{x}^j - \frac{2\tilde{x}^i x^i x^j}{x^k x^k}) \tilde{\partial}_j^a \quad (4.35)$$

$$v^a = (\tilde{x}^j - 2\tilde{x}^i \frac{x^i}{x^k x^k} \frac{\tilde{x}^j}{\tilde{x}^k \tilde{x}^k}) \tilde{\partial}_j^a \quad (4.36)$$

$$v^a = (\tilde{x}^j - 2\tilde{x}^i x^i \tilde{x}^j) \tilde{\partial}_j^a \quad (4.37)$$

$$v^a = -\tilde{x}^j \tilde{\partial}_j^a. \quad (4.38)$$

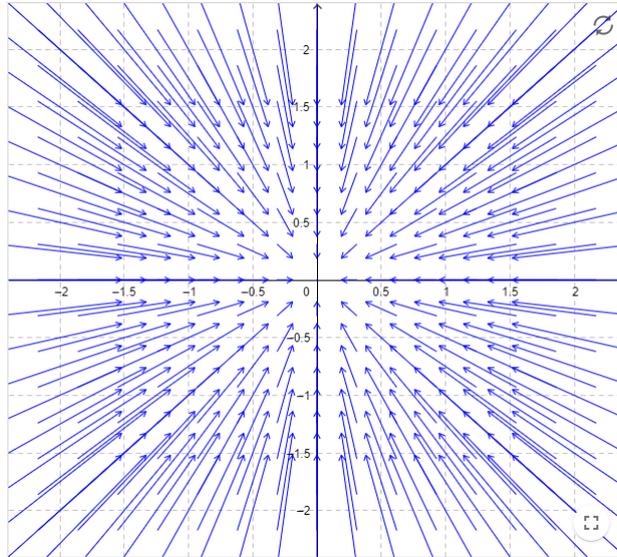


Abbildung 4.14: Zur Illustration des Vektorfeldes  $v^a$  im  $\tilde{x}^i$ -Koordinatensystem.

wobei wir im letzten Schritt  $\tilde{x}^i x^i = \tilde{x}^i \frac{\tilde{x}^i}{\tilde{x}^k \tilde{x}^k} = 1$  und bei (4.34)  $\frac{1}{x^k x^k} \frac{1}{\tilde{x}^k \tilde{x}^k} = 1$  ausnutzten. Nun haben wir mit der Kartenwechselfunktion und dem Transformationsverhalten des Basisvektors das Vektorfeld im anderen Koordinatensystem erhalten! In Abb. 4.14 ist es illustriert. Wie kann man sich das jetzt auf der Kugel vorstellen? Man muss sich überlegen, dass die jeweiligen Ursprünge in Abb.4.13 und Abb.4.14 der Nord/Südpol der Kugel darstellen. Dann kann man sich vorstellen, dass die Vektoren am Südpol beginnen<sup>11</sup> entlang der Kugel bis zum Äquator größer werden, dann aber kleiner im Betrag im Nordpol enden<sup>12</sup>.

<sup>11</sup>Wie die elektrischen Feldstärkevektoren eines geladenen Teilchens.

<sup>12</sup>Wie bei einem negativ geladenen Teilchen.

## 4.4.2 Parallelverschiebung und kovariante Ableitung

Für die Allgemeine Relativitätstheorie sind aber nicht irgendwelche Kurven und Vektorfelder interessant, sondern welche die *parallel* zueinander sind. Also müssen wir nun besprechen, was parallele Vektorfelder entlang Kurven sind. Man stelle sich eine Kurve  $\gamma$  im  $\mathbb{R}^{n13}$  vor. Ihr Tangentenvektor ist für jeden Wert des Kurvenparameters  $t$  als  $u^a(t) = \dot{x}^i(t)\partial_i^a$  gegeben. Seien nun  $v^i(x^m)$  die Komponenten eines Vektorfeldes  $v^a$  bzgl. einer Basis im  $\mathbb{R}^n$  gegeben. Nun wollen wir fordern, dass  $v^i(x^m)$  entlang der Kurve  $\gamma$  parallel ist. D.h. das Vektorfeld ändert sich weder in Richtung noch in Betrag entlang der Kurve. Mathematisch heißt das:

$$\frac{d}{dt}v^i(x^m(t)) = 0. \quad (4.39)$$

Wendet man nun die Kettenregel an mit  $\partial_l = \frac{\partial}{\partial x^l}$ , erhält man:

$$\dot{x}^l(t)\partial_l v^i(x^m(t)) = 0. \quad (4.40)$$

Das ist das Skalarprodukt zwischen dem Tangentenvektor und dem Gradienten von  $v^i$  bzgl. einer bestimmten Basis. Um diese Bedingung in Basisdarstellung zu schreiben, müssen wir betrachten wie der Gradient von  $v^a$  aussieht. Es gilt:

$$v^a = v^i \partial_i^a \quad (4.41)$$

$$\partial_b v^a := \partial_j v^i dx_b^j \partial_i^a = \partial_b(v^i)\partial_i^a. \quad (4.42)$$

Das stellt einen Tensor der Stufe (1,1) dar, wobei  $dx_b^j \partial_i^a$  die Basis und  $\partial_j v^i$  der tensorielle Gradient von  $v^i$  ist. Mit den Basisvektoren schreibt man:

$$u^a \partial_a v^b = \dot{x}^i(t)\partial_i^a \partial_j dx_a^j v^k \partial_k^b = \delta_i^j \dot{x}^i(t)\partial_j v^k \partial_k^b = (\dot{x}^j(t)\partial_j v^k)\partial_k^b = 0 \rightarrow \dot{x}^j(t)\partial_j v^k = 0, \quad (4.43)$$

wobei  $\partial_a = \partial_j dx_a^j$  und  $\partial_i^a dx_a^j = \delta_i^j$  gilt. Um diese Parallelitätsbedingung weiter zu analysieren, müssen wir anschauen, was der Gradient von skalaren Funktionen ist. Sei  $f$  eine Skalare Funktion im  $x^i$ -Koordinatensystem. Da der Wert der Funktion an der Stelle  $x^i$  in jedem Koordinatensystem gleich sein muss, gilt  $f(x^j) = \tilde{f}(\tilde{x}^i)$ . Dabei ist nicht gemeint, dass die Funktion bei  $x^i = (1, 2)$  und  $\tilde{x}^i = (1, 2)$  gleich sind, denn allg. beschreiben diese zwei Punkte nicht dieselben Punkt auf der Mannigfaltigkeit. Ergo müssen diese zwei Punkte nicht in der Funktion übereinstimmen. Nur wenn sie denselben Punkt  $p$  als Urbild besitzen, also denselben Punkt auf der Mannigfaltigkeit beschreiben, gilt die Gleichheit. Betrachten wir nun den Gradienten der Funktion:

$$\tilde{\partial}_a f = \tilde{\partial}_i \tilde{f} d\tilde{x}_a^i = \tilde{\partial}_i x^j \partial_j f d\tilde{x}_a^i = \tilde{\partial}_i x^j d\tilde{x}_a^i \partial_j f = dx_a^j \partial_j f = \partial_a f, \quad (4.44)$$

---

<sup>13</sup>Der Autor tut sich schwer eine n-dimensionale Kurve sich vorzustellen, deshalb reicht der  $\mathbb{R}^3$  für's Verständnis.

wobei wir im vorletzten Schritt das Transformationsverhalten der Differentiale ausgenutzt haben:  $dx_a^j = \tilde{\partial}_i x^j d\tilde{x}_a^i$ . Also ist der Gradient einer Skalaren Funktion immer gleich, egal welches Koordinatensystem wir verwenden! Diese Ableitung nennt man *koordinaten-kovariante Ableitung*; „koordinaten“, weil wir nach ganz bestimmten Koordinaten ableiten, „kovariant“, da sie aus einem Tensor der Stufe (r,s) einen Tensor der Stufe (r,s+1) macht. Betrachten wir nochmals den Gradienten eines Vektorfeldes, diesmal wenden wir aber die Kettenregel auf  $v^i \partial_i^b$  an:

$$\partial_a v^b = \partial_a(v^i \partial_i^b) = \partial_a(v^i) \partial_i^b + v^i \partial_a(\partial_i^b) \quad (4.45)$$

Doch in Gleichung (4.42) hatten wir keinen zweiten Term, d.h.  $v^i \partial_a(\partial_i^b) = 0$  bzw.  $\partial_a(\partial_i^b) = 0$ . Also sind die Basisvektoren bzgl. der koordinaten-kovarianten Ableitung wie Konstanten. Das gilt aber nur für die entsprechenden Basisvektoren! Das zeigt, dass die Ableitung und die Basisvektoren fest miteinander verknüpft sind.

Aber wer hält uns davon ab,  $v^b$  im  $\tilde{x}^i$ -Koordinatensystem darzustellen? Schauen wir mal an, was passiert, wenn wir den Gradienten auf das Vektorfeld wirken lassen, wobei wir die Basisdarstellung im  $\tilde{x}^i$ -Koordinatensystem verwenden:

$$\partial_a v^b = \partial_a(\tilde{v}^i \tilde{\partial}_i^b) = \partial_a(\tilde{v}^i) \tilde{\partial}_i^b + \tilde{v}^i \partial_a(\tilde{\partial}_i^b) = \tilde{\partial}_a(\tilde{v}^i) \tilde{\partial}_i^b + \tilde{v}^i \partial_a(\tilde{\partial}_i^b) = \quad (4.46)$$

$$\tilde{\partial}_a(\tilde{v}^i \tilde{\partial}_i^b) + \tilde{v}^i \partial_a(\tilde{\partial}_i^b) = \tilde{\partial}_a v^b + \tilde{v}^i \partial_a(\tilde{\partial}_i^b) = \tilde{\partial}_a v^b + \tilde{v}^j \partial_a \tilde{\partial}_j x^k \partial_k^b = \quad (4.47)$$

$$\partial_a v^b = \tilde{\partial}_a v^b + \tilde{v}^j d\tilde{x}_a^l \tilde{\partial}_l \tilde{\partial}_j x^k \partial_k \tilde{x}^p \tilde{\partial}_p^b \quad (4.48)$$

$$\partial_a v^b = \tilde{\partial}_a v^b + \tilde{v}^j d\tilde{x}_a^l \tilde{\partial}_p^b (\tilde{\partial}_l \tilde{\partial}_j x^k \partial_k \tilde{x}^p) \quad (4.49)$$

$$\partial_a v^b = \tilde{\partial}_a v^b + v^c d\tilde{x}_c^j d\tilde{x}_a^l \tilde{\partial}_p^b (\tilde{\partial}_l \tilde{\partial}_j x^k \partial_k \tilde{x}^p) \quad (4.50)$$

$$\partial_a v^b = \tilde{\partial}_a v^b + \Gamma_{ac}^b v^c \quad (4.51)$$

Gehen wir diese Umformungen Schritt für Schritt durch. Als erstes wurde die Kettenregel<sup>14</sup> angewandt. Dann nutzten wir den Fakt aus, dass Funktionen nach jeder koordinaten-kovariante Ableitung abgeleitet gleich sind, ergo  $\partial_a(\tilde{v}^i) = \tilde{\partial}_a(\tilde{v}^i)$ . Danach haben wir die Beziehung  $\tilde{\partial}_a \tilde{\partial}_i^b = 0$  verwendet um  $\tilde{\partial}_a(\tilde{v}^i \tilde{\partial}_i^b)$  zu erhalten. Da  $v^b = \tilde{v}^i \tilde{\partial}_i^b$  gilt, fassten wir  $\tilde{\partial}_a(\tilde{v}^i \tilde{\partial}_i^b)$  zu  $\tilde{\partial}_a v^b$  zusammen. Damit waren wir mit dem ersten Term fertig und haben uns dann mit dem zweiten zugewandt. Da haben wir als erstes das Transformationsgesetz für  $\tilde{\partial}_i^b$  angewandt. Danach haben wir die koordinaten-kovariante Ableitung in der Basisdarstellung geschrieben und im selben Schritt transformierten wir  $\partial_k^b$  wieder zurück ins  $\tilde{x}^i$ -Koordinatensystem. Bei (4.49) formten wir nichts um, sondern änderten nur die Anordnung der einzelnen Terme. Im nächsten Schritt nutzten wir aus, dass man mit dualen Vektoren die Komponenten eines Vektor extrahieren kann, wie in Gl.(3.18), also wir substituierten  $v^j = v^c dx_c^j$ . Dann fassten wir diesen komischen Term zusammen zu einem Tensor mit Stufe (1,2):  $\Gamma_{ac}^b = d\tilde{x}_c^j d\tilde{x}_a^l \tilde{\partial}_p^b (\tilde{\partial}_l \tilde{\partial}_j x^k \partial_k \tilde{x}^p)$ . Dieser Tensor heißt *Differenzentensor* oder *Christoffel-Symbol* und er ist enorm wichtig, denn mit ihm, kann man eine koordinaten-kovariante Ableitung in Abhängigkeit

<sup>14</sup>Auch als Leibniz-Regel bekannt!

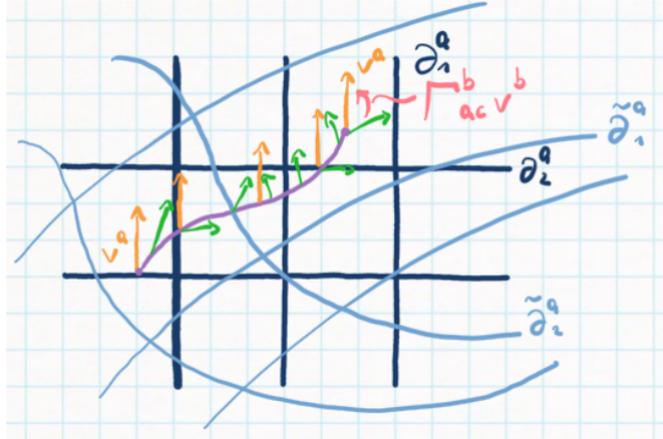


Abbildung 4.15: Zur Illustration der Parallelverschiebung entlang verschiedener Koordinatensysteme.

einer anderen anschreiben. Der Grund warum er Differenzentensor heißt, wird mit einer trivialen Umformung klar:

$$(\partial_a - \tilde{\partial}_a)v^b = \Gamma^b_{ac}v^c. \quad (4.52)$$

Also gibt das Christoffel-Symbol den Unterschied zwischen zwei koordinaten-kovarianten Ableitungen an. Man kann ihn auch anders, geometrisch, interpretieren. Man stelle sich normale kartesische Koordinaten vor. Darauf stellt man sich ein beliebig gekrümmtes Koordinatensystem vor. Wenn man nun einen Vektor entlang des kartesischen Koordinatensystems parallelverschiebt, bleibt seine Länge und seine Richtung erhalten. Wollte man den Vektor jetzt entlang den gekrümmten Koordinaten parallelverschieben, würde seine Länge und Richtung im allg. nicht gleich bleiben. Doch am Ende der Kurve, braucht man nur eine lineare Transformation, also Streckung, Stauchung, Drehung, ..., um den Vektor so zu präparieren, sodass er so aussieht, als hätte man ihn entlang der kartesischen Koordinaten verschoben. Diese lineare Transformation ist das Christoffel-Symbol  $\Gamma^b_{ac}$ . In Abb. 4.15 wird dieses Vorgehen illustriert. Man beachte, im Christoffel-Symbol kommt eine zweite Ableitung nach den Koordinaten vor  $\tilde{\partial}_l \tilde{\partial}_j x^k(\tilde{x}^m)$  vor, d.h. wenn diese Ableitung 0 ist, heißt das, dass die zwei Koordinatensysteme linear miteinander in Verbindung stehen. Dann wird das Christoffel-Symbol 0, die Ableitungen sind gleich und die Parallelverschiebungen stimmen überein.

Nun wollen wir eine allgemeinere Ableitung definieren, welche global auf der Mannigfaltigkeit existiert. Die koordinaten-kovariante Ableitung ist beschränkt auf die Koordinatensysteme, welche durch die Kartenabbildungen aufgespannt werden. Wir definieren nun die **kovariante Ableitung**  $\nabla_a$  auf der Mannigfaltigkeit  $M$ ;  $\nabla_a : \mathcal{T}_q^p \rightarrow \mathcal{T}_{q+1}^p$ . Also sie erzeugt aus einem Tensor der Stufe  $(p,q)$  einen Tensor der Stufe  $(p,q+1)$ , wie die koordinaten-kovariante Ableitung. Da wie immer unser Leben öde wäre, müssen wir nun vier Bedingungen aufstellen, die die kovariante Ableitung erfüllen muss, wobei  $t_{b_1, \dots, b_q}^{a_1, \dots, a_p} = t_{\mathcal{B}}^{\mathcal{A}}$  gilt:

1.  $\nabla_a(t_{\mathcal{B}}^A + s_{\mathcal{B}}^A) = \nabla_a t_{\mathcal{B}}^A + \nabla_a s_{\mathcal{B}}^A | \nabla_a(\lambda t_{\mathcal{B}}^A) = \lambda \nabla_a t_{\mathcal{B}}^A$
2.  $\nabla_a(t_{\mathcal{B}}^A s_{\mathcal{D}}^C) = s_{\mathcal{D}}^C \nabla_a t_{\mathcal{B}}^A + t_{\mathcal{B}}^A \nabla_a s_{\mathcal{D}}^C$
3.  $\nabla_a f = \partial_a f$
4.  $\nabla_a \nabla_b f = \nabla_b \nabla_a f$

Also wir wollen, dass 1. die kovariante Ableitung linear ist. 2. wollen wir, dass sie die Leibnitz-Regel erfüllt. 3. fordern wir, dass die kovariante Ableitung auf skalare Funktionen gleich der koordinatenkovarianten Ableitung ist, bzw. daraus folgt, dass jede kovariante Ableitung  $(\nabla_a, \tilde{\nabla}_a, \dots)$  auf skalare Funktionen gleich sind, da jede koordinatenkovariante Ableitung auf skalare Funktionen gleich sind. Zu guter Letzt fordern wir den Satz von Schwarz, d.h. dass die Reihenfolge von zwei kovarianten Ableitung keinen Unterschied macht. Eine kleine Umformung zeigt, dass der Kommutator zwischen zwei kovarianten Ableitungen, angewandt auf eine skalare Funktionen, verschwindet.

Nun wollen wir schauen, wie diese Ableitung, definiert nur durch 4 Bedingungen, auf Vektoren und duale Vektoren wirkt. Betrachten wir als erstes zwei kovariante Ableitungen, die beide diese 4 Bedingungen erfüllen:

$$(\tilde{\nabla}_a - \nabla_a)v^b = (\tilde{\nabla}_a - \nabla_a)(v^i \partial_i^b) = (\tilde{\nabla}_a - \nabla_a)(v^i) \partial_i^b + v^i (\tilde{\nabla}_a - \nabla_a)(\partial_i^b) = \quad (4.53)$$

$$v^c dx_c^i (\tilde{\nabla}_a - \nabla_a)(\partial_i^b) = C_{ca}^b v^c \quad (4.54)$$

Dieser Differenzentensor ist global auf der Mannigfaltigkeit definiert, denn es gilt:

$$\hat{C}_{ca}^b = d\hat{x}_c^i (\tilde{\nabla}_a - \nabla_a) \hat{\partial}_i^b \quad (4.55)$$

$$\hat{C}_{ca}^b = d\hat{x}_c^i \partial_i x^j (\tilde{\nabla}_a - \nabla_a) \partial_j^b \quad (4.56)$$

$$\hat{C}_{ca}^b = dx_c^i (\tilde{\nabla}_a - \nabla_a) \partial_j^b = C_{ca}^b, \quad (4.57)$$

wobei  $\hat{C}_{ca}^b$  der Differenzentensor in einem  $\hat{x}^i$ -Koordinatensystem ist. Wir erkennen, dass der Tensor unabhängig vom Koordinatensystem ist, wodurch er eine globale Größe auf der Mannigfaltigkeit ist! Nun betrachten wir diese Differenz zwischen zwei kovarianten Ableitungen wirkend auf duale Vektoren:

$$(\tilde{\nabla}_a - \nabla_a)(\omega_b v^b) = 0 = (\tilde{\nabla}_a - \nabla_a)(\omega_b) v^b + \omega_b (\tilde{\nabla}_a - \nabla_a)(v^b) = \quad (4.58)$$

$$(\tilde{\nabla}_a - \nabla_a)(\omega_b) v^b + \omega_b C_{ca}^b v^c \quad (4.59)$$

$$(\tilde{\nabla}_a - \nabla_a)(\omega_b) v^b = -C_{ca}^b \omega_b v^c \quad (4.60)$$

$$(\tilde{\nabla}_a - \nabla_a)\omega_b = -C_{ca}^b \omega_b \quad (4.61)$$

Also hat der Differenzentensor  $C_{ac}^b$  ein negatives Vorzeichen für duale Vektoren. Aus Bedingung 4

erhalten wir  $\nabla_a \nabla_b f = \nabla_b \nabla_a f$ , aber  $\nabla_b(f)$  ist ein Tensor der Stufe  $(0,1)$ , also können wir schreiben:

$$\tilde{\nabla}_a \tilde{\nabla}_b(f) = \tilde{\nabla}_a(\nabla_b f) = \nabla_a \nabla_b f - C_{ba}^c \nabla_c f \quad (4.62)$$

$$\tilde{\nabla}_b \tilde{\nabla}_a(f) = \nabla_b \nabla_a f - C_{ab}^c \nabla_c f \quad (4.63)$$

Diese zwei Gleichungen können wir voneinander subtrahieren und wir wissen (Bed.4), dass das Ergebnis 0 sein soll:

$$(C_{ba}^c - C_{ab}^c) \nabla_c f = 0, \quad (4.64)$$

Also muss  $(C_{ba}^c - C_{ab}^c) = 0$  gelten, was bedeutet, dass der Differenzentensor für die kovariante Ableitung symmetrisch bzgl. den zwei unteren Indizes ist. Nun wollen wir noch eine Größe, einen Tensor, definieren, nämlich den *Riemann'schen Krümmungstensor*:

$$[\nabla_a, \nabla_b]v^c = R_{dac}^c v^d, \quad (4.65)$$

diese Größe wird die Krümmung unserer Mannigfaltigkeit bestimmen. Zur Zusammenfassung dieses letzten Abschnittes der kovarianten Ableitung: Die 3. Bedingung bedingt die Existenz des Differenzentensor  $C_{ab}^c v^b = (\tilde{\nabla}_a - \nabla_a)v^c$  und die 4. Bedingung bedingt die Existenz des Riemann-Tensor  $R_{dab}^c v^d = [\nabla_a, \nabla_b]v^c$ . Doch nun ein kleiner Einschub in Algebras und Derivationen.

### Mathematical Interlude: Algebra und Derivationen

Eine *Algebra* ist ein Vektorraum, auf dem bilineare Abbildungen definiert sind. Man sagt das Paar  $(A, \cdot)$  ist eine Algebra. Für  $\cdot$  gilt:

1.  $\cdot : A \otimes A \rightarrow A$
2.  $(a, b) = a \cdot b \in A$
3.  $(\lambda a + \mu b, \nu c + \epsilon d) = \lambda \nu a \cdot c + \lambda \epsilon a \cdot d + \mu \nu b \cdot c + \mu \epsilon b \cdot d$

Eine bestimmte bilineare Abbildung, eine *Derivation*, hat folgende Eigenschaft:

$$\mathcal{D}(A) = \{D : A \rightarrow A | D \text{ ist bilinear und } D(A \cdot B) = D(A) \cdot B + A \cdot D(B)\}.$$

Also erfüllen sie die Leibniz-Regel! D.h. eine Derivation ist eine bestimmte bilineare Abbildung mit Leibniz-Regel. Nun betrachten wir eine zweifache Anwendung von zwei Derivationen  $D_1$  und  $D_2$  auf ein inneres Produkt  $(a \cdot b)$ :

$$D_1 D_2(a \cdot b) = D_1(D_2(a) \cdot b + a \cdot D_2(b)) = \quad (4.66)$$

$$D_1(D_2(a)) \cdot b + D_2(a) \cdot D_1(b) + D_1(a) \cdot D_2(b) + D_1(D_2(b)) \cdot a \quad (4.67)$$

Berechnet man nun  $D_2 D_1(a \cdot b)$  und subtrahiert man davon Gl. (4.67) erhält man:

$$[D_1, D_2](a \cdot b) = [D_1, D_2](a) \cdot b + [D_1, D_2](b) \cdot a. \quad (4.68)$$

D.h. der Kommutator von zwei Derivationen ist selbst auch eine Derivation, da er auch die Leibniz-Regel erfüllt! Nun kann man eine Algebra bilden, nämlich eine deren Vektorraum der Raum der Derivationen  $\mathcal{D}(A)$  ist. Die bilineare Abbildung ist dann ein Produkt:  $(\mathcal{D}(A), *)$  mit  $* : D * D' := [D, D']$ , also einfach der Kommutator. Nach weiterer Überlegung kommt man zum Schluss, dass eine Derivation existiert, die durch die Derivationen von  $\mathcal{D}(A)$  erzeugt wird:  $\mathfrak{D} : \mathcal{D}(A) \rightarrow \mathcal{D}(A)$  und diese Derivation lautet trivialerweise wie folgt:  $\mathfrak{D}(D) = [D_1, D]$ , wobei  $D_1$  eine beliebige feste Derivation ist. Diese Derivation erfüllt natürlich auch die Leibniz-Regel:

$$\mathfrak{D}(D_2 * D_3) = \mathfrak{D}(D_2) * D_3 + D_2 * \mathfrak{D}(D_3) = [D_1, D_2] * D_3 + D_2 * [D_1, D_3] \quad (4.69)$$

$$\mathfrak{D}(D_2 * D_3) = [D_1, [D_2, D_3]] = [[D_1, D_2], D_3] + [D_2, [D_1, D_3]], \quad (4.70)$$

wobei diese Beziehung als Jacobi-Identität bekannt ist. Wir erkennen, im Kontext der Derivationen ist das nichts Anderes als die Leibniz-Regel.

Nun können wir uns wieder dem Riemann'scher Krümmungstensor. Da  $[\nabla_a, \nabla_b]v^c = R_{dac}^c v^d$  die Definition des Krümmungstensor ist, schreiben wir ihn schnell in  $[\nabla_a, \nabla_b](v^i \partial_i^c) = v^d dx_d^i [\nabla_a, \nabla_b] \partial_i^c = R_{dac}^c v^d$  um. Wir können überprüfen, ob dieser Tensor global auf unserer Mannigfaltigkeit existiert, so wie wir es beim Differenzentensor  $C_{ab}^b$  gemacht haben:

$$\hat{R}_{dab}^c = d\hat{x}_d^i [\nabla_a, \nabla_b] \hat{\partial}_i^c \quad (4.71)$$

$$\hat{R}_{dab}^c = d\hat{x}_d^i \hat{\partial}_i x^j [\nabla_a, \nabla_b] \partial_j^c \quad (4.72)$$

$$\hat{R}_{dab}^c = dx_d^j [\nabla_a, \nabla_b] \partial_j^c = R_{dab}^c \quad (4.73)$$

Da offenbar die Gleichheit zwischen zwei willkürlichen Koordinatensystemen gilt, gilt die Gleichheit überall auf der Mannigfaltigkeit, womit der Krümmungstensor global definiert ist. Aber der Krümmungstensor ist dennoch abhängig von der jeweiligen kovarianten Ableitung, die im Kommutator steht. Wir haben schon gesehen, dass wir koordinaten-kovariante und kovariante Ableitungen in Abhängigkeit von anderen Ableitungen anschreiben können, aber geht das auch mit dem Krümmungstensor? Wenn schon so gefragt wird, dann heißt die Antwort ja:

$$\tilde{\nabla}_a \tilde{\nabla}_b v^c = \nabla_a (\tilde{\nabla}_b v^c) + C_{ma}^c (\tilde{\nabla}_b v^m) \quad (4.74)$$

$$\tilde{\nabla}_a \tilde{\nabla}_b v^c = \nabla_a \nabla_b v^c + \nabla_a (C_{mb}^c) v^m + C_{mb}^c \nabla_a (v^m) + C_{ma}^c (\nabla_b) v^m + C_{ma}^c C_{nb}^m v^n - C_{ba}^m \nabla_m v^c - C_{ba}^m C_{nm}^c v^n \quad (4.75)$$

Bildet man nun den Kommutator  $[\tilde{\nabla}_a, \tilde{\nabla}_b]$  erhält man:

$$\tilde{R}_{dab}^c v^d = [\tilde{\nabla}_a, \tilde{\nabla}_b]v^c = R_{dab}^c v^d + (\nabla_a C_{mb}^c - \nabla_b C_{ma}^c + C_{na}^c C_{mb}^n - C_{nb}^c C_{ma}^n)v^m \quad (4.76)$$

D.h. der Riemann-Tensor lautet:

$$\tilde{R}_{dab}^c = R_{dab}^c + \nabla_a C_{db}^c - \nabla_b C_{da}^c + C_{ma}^c C_{db}^m - C_{mb}^c C_{da}^m \quad (4.77)$$

Als kleine Nebenbemerkung: Der Riemann-Tensor ist für die normale Ableitung  $\partial_a$  0:

$$\partial_a \partial_b v^c = \partial_a (dx_b^\mu \partial_\mu v^\nu \partial_\nu^c) \quad (4.78)$$

$$\partial_a \partial_b v^c = dx_b^\mu dx_a^\rho \partial_\rho \partial_m u v^\nu \partial_\nu^c \quad (4.79)$$

$$\partial_a \partial_b v^c = dx_a^\rho dx_b^\mu \partial_m u \partial_\rho v^\nu \partial_\nu^c \quad (4.80)$$

$$\partial_a \partial_b v^c = \partial_b \partial_a v^c \rightarrow [\partial_a, \partial_b]v^c = 0. \quad (4.81)$$

Da wir uns intensiv mit einer globalen Ableitung für unsere Mannigfaltigkeiten auseinander gesetzt haben, können wir wieder zurück zu unserem eigentlichen Thema: Parallele Vektoren entlang Kurven. Zur kurzem Wiederholung, ein Vektorfeld  $v^a$  ist entlang einer Kurve  $\gamma$  parallel zu sich selbst, wenn gilt:

$$u^b \partial_b v^a = 0, \quad (4.82)$$

wobei  $u^b$  der Tangentenvektor der Kurve und  $\partial_b$  die koordinaten-kovariante Ableitung ist. Jetzt können wir aber diese Bedingung verallgemeinern und sagen, eine Kurve ist  $\nabla$ -parallel entlang  $\gamma$ , wenn gilt:

$$u^b \nabla_b v^a = 0. \quad (4.83)$$

Wir wissen ja, dass wir die kovariante Ableitung mit einem gut gewählten Differenzentensor in Abhängigkeit einer anderen Ableitung ausdrücken können. Und genau das machen wir jetzt auch: Sei  $\nabla_a = \partial_a + \Gamma_a$ . Wir erweitern also unsere koordinaten-kovariante Ableitung mit einem (noch unbekanntem) Differenzentensor. Schauen wir mal, was passiert, wenn wir diese Beziehung in Gl.

(4.83) einsetzen:

$$u^a \nabla_a v^b = u^a \partial_a v^b + u^a \Gamma_{ca}^b v^c = 0 \quad (4.84)$$

$$\frac{d}{dt} v^\nu \partial_\nu^b + \dot{x}^\mu \partial_\mu^a \Gamma_{\rho\sigma}^\nu v^\kappa \partial_\kappa^c \partial_\nu^b dx_c^\rho dx_a^\sigma = 0 \quad (4.85)$$

$$\frac{d}{dt} v^\nu \partial_\nu^b + \Gamma_{\rho\sigma}^\nu \delta_\mu^\sigma \delta_\kappa^\rho \dot{x}^\mu v^\kappa \partial_\nu^b = 0 \quad (4.86)$$

$$\frac{d}{dt} v^\nu \partial_\nu^b + \Gamma_{\rho\sigma}^\nu \dot{x}^\sigma v^\rho \partial_\nu^b = 0 \left( \frac{d}{dt} v^\nu + \Gamma_{\rho\sigma}^\nu \dot{x}^\sigma v^\rho \right) \partial_\nu^b = 0 \quad (4.87)$$

$$\frac{d}{dt} v^\nu(t) + [\Gamma_{\rho\sigma}^\nu(x^\mu(t)) \dot{x}^\sigma(t)] v^\rho(t) = 0 \quad (4.88)$$

$$\dot{v}^\nu(t) + H_\rho^\nu(t) v^\rho(t) = 0. \quad (4.89)$$

Um die Leserschaft nun völlig zu verwirren, wird nun behauptet, dass wir diese Gleichung kennen: Das ist die Schrödingergleichung. Erster Term: eine Zeitableitung, zweiter Term: ein zeitabhängiger Hamilton-Operator wirkend auf unser gesuchtes Vektorfeld. Bevor jeder dieses Skriptum frustriert zur Seite legt, seid unbesorgt, das ist nicht *die* Schrödingergleichung. Sie ist nur mathematisch ähnlich aufgebaut. Denn mit einem Zeitentwicklungsoperator  $U_\rho^\nu(t)$  kann man diese Gleichung lösen:

$$v^\nu(t) = U_\rho^\nu(t) v^\rho(0) \rightarrow \dot{U}_\rho^\nu + H_\sigma^\nu U_\rho^\sigma = 0, \quad (4.90)$$

mit folgender Anfangsbedingung:  $U_\rho^\nu(0) = \delta_\rho^\nu$ . Wenn wir nun wissen, wie das Vektorfeld am Anfang der Kurve aussieht ( $v^\rho(0)$ ), müssen wir nur diese „Schrödingergleichung“ für den Zeitentwicklungsoperator lösen und können dadurch das Vektorfeld an jeder Stelle der Kurve bestimmen. Hier muss ein großer Unterschied gemacht werden mit Gl. (4.83): Während wir dort eine Kurve hatten, mussten wir schauen ob irgendein Vektorfeld, welches global definiert war, parallel zur Kurve ist. In Gl. (4.90) brauchen wir nur die Kurve und die Anfangsbedingung kennen, um das Vektorfeld *auf der Kurve* zu bestimmen. Wie das Vektorfeld außerhalb der Kurve aussieht, können wir nicht sagen, nur wie es sich entlang der Kurve verhält.

## Kapitel 5

# Einstein'sche-Feldgleichungen

Nun können wir uns endlich etwas Physik zuwenden. In diesem letztem Kapitel werden wir ein Analogon des 2. Newton'schen Gesetzes herleiten, Ableitungen von Vektorfeldern unter Strömungen betrachten und den Einstein-Tensor postulieren.

### 5.1 Autoparallele Kurven

Im frühen 20. Jahrhundert dachte Albert Einstein einige Zeit über Gravitation und Raumzeit nach. Er saß so viel in seinem Büro, dass er eines Tages einen Bauarbeiter sah, wie dieser vom Gerüst fiel. Einstein, besorgt, lief natürlich hinaus auf die Straße, um nach dem Armen zu sehen. Dieser antwortete auf die Frage, wie sich der Fall anfühlte, „Ich fühlte mich ganz *frei*“. Das war der Moment in dem Einstein klar wurde, dass Gravitation und der freie Fall ein und dieselbe Sache sein müssen. Denn in einem dunklen beschleunigenden Aufzug, ohne äußere Sinneseindrücke, kann man nicht sagen, ob man sich in einem Gravitationsfeld oder im freien Fall befindet.

Wenn laut Einstein Gravitation keine Kraft ist, sondern die Krümmung der Raumzeit, wie würde dann die Bewegungsgleichung eines freien Körpers in einer gekrümmten Raumzeit aussehen? Das sind die sogenannten *Autoparallelen Kurven*.

Autoparallele Kurven sind jene, deren Tangentenvektorfeld parallel zu sich selbst ist, d.h. die Parallelverschiebung des Tangentenvektor ist entlang der Kurve 0. Symbolisch gilt:

$$(u^a \nabla_a) u^b = 0. \quad (5.1)$$

Wir erkennen, dass hier nur das willkürliche Vektorfeld  $v^b$  aus Gl. (4.83) mit dem Tangentenvektorfeld ausgetauscht wurde. Gleichmaßen gilt die „Schrödingergleichung“ für das Tangentenvektorfeld:

$$\ddot{x}^\nu(t) + \Gamma_{\rho\sigma}^\nu \dot{x}^\rho(t) \dot{x}^\sigma(t) = 0. \quad (5.2)$$

Mit einer elementaren Umformung und Multiplikation mit der Masse  $m$ , könnte man schon eine Verbindung mit dem Gravitationspotential  $\Phi_g$  machen:

$$m\ddot{x}^\nu(t) = -m\Gamma_{\rho\sigma}^\nu \dot{x}^\rho(t)\dot{x}^\sigma(t) = -m\partial_\nu\Phi_g, \quad (5.3)$$

aber man erkennt, dass diese Beziehung nicht ganz richtig sein kann, da während links ein kontravarianter Index der Beschleunigung ist, befindet sich rechts ein kovarianter Index des Gradienten. Also so einfach ist es nun doch nicht.

Die Gleichung (5.2) hat nun einige andere Eigenschaften als Gl. (4.83): Wir suchen hier *nicht* nach einem Vektorfeld, welches parallel ist, sondern *direkt* die Kurve, deren Tangentenvektoren immer parallel sind, bzw. die Gl. (5.2) erfüllt. Darüber hinaus ist sie nicht mehr linear, sie kann also nicht mehr so wie die Schrödingergleichung gelöst werden. Da sie dadurch ein völlig anderes Objekt darstellt, hat sie auch einen eigenen Namen: *Die Gleichung der Autoparallelen Kurven*. Sie beantwortet die Frage, was die Kurven von freien Teilchen in einer gekrümmten Raumzeit sind. Wir haben hier aber ein kleines Problem. Für den Tensor  $\Gamma_{\rho\sigma}^\nu$  brauchen wir die kovariante Ableitung, aber wie sieht diese genau aus? Wir wissen nur, dass sie die vier Bedingungen erfüllen muss, aber ansonst haben wir ihr nicht mehr Form gegeben. Im Folgenden werden wir eine Bedingung für die kovariante Ableitung aufstellen, wodurch wir von allen möglichen Ableitungen, nur *eine* aussuchen können und diese ist dann *eindeutig*.

Hierfür müssen wir die *Metrik* wieder ins Gedächtnis bringen. Im Mathematical Interlude zur abstrakten Indexschreibweise (Kapitel 3.5) haben wir sie schon kennengelernt. Hier ist die Metrik  $g_{ab}$  ein Tensorfeld der Stufe (0,2) und stellt eine Verallgemeinerung der Minkowski-Metrik der SRT da.  $g_{ab}$  ist symmetrisch  $g_{ab} = g_{ba}$  und für eine orthonormale Basis des Tangentenraums  $T_pM$  gilt:  $g_{ab}E_\alpha^a E_\beta^b = \eta_{\alpha\beta}$ , also für jeden Tangentenraum, stellt die Metrik die Minkowski-Metrik dar. In normalen Vektorräumen stellt die Metrik eine Art von Innerem Produkt dar: Mit der Metrik kann man beliebige Indizes hinauf- und hinabstellen, je nach dem ob man  $g_{ab}$  oder ihre Inverse  $g^{ab}$  benutzt. Dadurch kann man das Innere Produkt zwischen zwei Vektoren bilden:  $v^a w^b g_{ab} = v^a w_a$ . Mit einem Inneren Produkt können wir das Verhältnis von zwei Vektoren zueinander beschreiben. Seien zwei Vektorfelder  $v^b$  und  $w^b$  parallel entlang der Kurve  $\gamma$ , also für sie gilt jeweils:

$$(u^a \nabla_a) v^b = 0 \quad (5.4)$$

$$(u^a \nabla_a) w^b = 0. \quad (5.5)$$

Dadurch können wir ohne Fehler folgendes behaupten:

$$(u^a \nabla_a)(v^b w^c g_{bc}) = w^c g_{bc} u^a \nabla_a (v^b) + v^b g_{bc} u^a \nabla_a (w^c) + v^b w^c u^a \nabla_a (g_{bc}) = v^b w^c u^a \nabla_a (g_{bc}) = 0. \quad (5.6)$$

Wir verlangen also, dass das Innere Produkt der beiden willkürlichen Vektoren  $v^b w^c g_{bc}$  konstant

entlang der Kurve ist. Als Konsequenz bekommen wir folgende Bedingung für unsere Ableitung:

$$\nabla_a(g_{bc}) = 0. \quad (5.7)$$

Diese Bedingung gilt für beliebige Anfangsbedingungen der Vektorfelder  $v^b$  und  $w^c$ , sowie der Kurve mit  $x(t=0)^i = x_0^i$ , wodurch diese Bedingung global für unsere kovariante Ableitung gelten soll. Um mehr Information aus dieser Beziehung zu gewinnen, schreiben wir sie nun in Abhängigkeit einer anderen Ableitung um:

$$\nabla_a(g_{bc}) = 0 = \tilde{\nabla}_a g_{bc} - C_{ba}^m g_{mc} - C_{ca}^m g_{bm}. \quad (5.8)$$

Durch elementarer Umformung erhält man:

$$C_{ba}^m g_{mc} + C_{ca}^m g_{bm} = \tilde{\nabla}_a g_{bc}. \quad (5.9)$$

Nun wollen wir die Eigenschaft der Metrik ausnutzen, mit der wir Indizes herumschieben können:

$$C_{cba} + C_{bca} = \tilde{\nabla}_a g_{bc}. \quad (5.10)$$

Nun permutieren wir die Indizes, wodurch wir folgendes erhalten:

$$\mathbf{I} : C_{cba} + C_{bca} = \tilde{\nabla}_a g_{bc} \quad (5.11)$$

$$\mathbf{II} : C_{acb} + C_{cab} = \tilde{\nabla}_b g_{ca} \quad (5.12)$$

$$\mathbf{III} : C_{bac} + C_{abc} = \tilde{\nabla}_c g_{ab} \quad (5.13)$$

Nun subtrahieren wir Gl. von **I** Gl. **III** und addieren **II**. Durch die Symmetrie des Differenzentensors in den letzten beiden Indizes, kürzt sich jeweils  $C_{abc}$  mit  $C_{acb}$  und  $C_{bac}$  mit  $C_{bca}$ : Dadurch erhalten wir folgende Gleichung:

$$2C_{cab} = \tilde{\nabla}_a g_{cb} + \tilde{\nabla}_b g_{ca} - \tilde{\nabla}_c g_{ab}. \quad (5.14)$$

Nun nutzen wir wieder die Höherung oder Erniedrigung der Indizes mittels einer Metrik und erhalten:

$$C_{ab}^c = \frac{1}{2} g^{cm} (\tilde{\nabla}_a g_{mb} + \tilde{\nabla}_b g_{ma} - \tilde{\nabla}_m g_{ab}), \quad (5.15)$$

wodurch wir den Differenzentensor mittels der Ableitung definiert haben, wo die Metrik bei kovarianter Ableitung 0 wird. In der Literatur wird oft geschrieben, dass unsere kovariante Ableitung folgendermaßen aussieht:

$$\nabla_a = \partial_a + \Gamma_a, \quad (5.16)$$

wodurch das Christoffel-Symbol folgende Form hat:

$$\Gamma_{ab}^c = \frac{1}{2} g^{cm} (\partial_a g_{mb} + \partial_b g_{ma} - \partial_m g_{ab}). \quad (5.17)$$

Damit haben wir die sogenannte *Levi-Civita* Ableitung gefunden. Man sieht, dass die Metrik unsere Ableitung bestimmt und unsere Ableitung bestimmt die Krümmung (Riemann-Tensor). Somit wird die Metrik das zentrale Objekt der Allgemeinen Relativitätstheorie. Doch wie bestimmen wir  $g_{ab}$ ?

## 5.2 Lie-Ableitung

Um weitere Information zur Metrik zu bekommen, müssen wir uns noch schnell mit der *Newton-Deviationsgleichung* auseinandersetzen. Wie lautet diese Gleichung? Dafür betrachte man eine Schar an freifallenden Objekten, deren Masse  $m(s)$  und deren Aufenthaltspunkt  $x^i(t, s)$  lautet. Hier ist  $s$  der Selektorparameter, der das betrachtete Objekt bestimmt. Z.B. kann bei  $s=1$  ein Fahrrad sein,  $s=4$  ist ein Ball,  $s=2$  ein Hamster usw. Wenn diese Objekte in einem Gravitationsfeld fallen, lautet ihre Bewegungsgleichung:

$$m(s)\ddot{x}^i(t, s) = -m(s)\partial_i\Phi(x^m(t, s)), \quad (5.18)$$

wobei  $\Phi(x^m(t, s))$  das Gravitationspotential in Abhängigkeit des Ortes des betrachteten Objektes ist. Nun definieren wir einen Vektor, der von einem Objekt aufs nächste hinzeigt. Origineller Weise nennen wir ihn Zeiger und es gilt:

$$\xi^i(t) = \frac{\partial x^i(t, s)}{\partial s}, \quad (5.19)$$

also die Änderung der Kurve des Objektes in Richtung des nächsten Objektes gibt uns unseren Zeiger. Wir wollen auf eine Bewegungsgleichung des Zeigers kommen, deshalb differenzieren wir Gl. (5.18) nach dem Selektorparameter  $s$ :

$$\frac{\partial}{\partial s}\ddot{x}^i(t, s) = -\frac{\partial}{\partial s}\partial_i\Phi(x^m(t, s)). \quad (5.20)$$

Unter Verwendung der Linearität der Ableitung (Vertauschung von  $\frac{\partial}{\partial s}$  und  $\partial_i$ ) und der Kettenregel bei  $\Phi(x^m(t, s))$  erhalten wir:

$$\frac{d^2}{dt^2}\frac{\partial x^i(t, s)}{\partial s} + \partial_k\partial_i\Phi(x^m(t, s))\frac{\partial x^k(t, s)}{\partial s} = 0 \quad (5.21)$$

$$\ddot{\xi}^i + \partial_k\partial_i\Phi(x^m(t, s))\xi^k(t) = 0. \quad (5.22)$$

Gl. (5.22) ist die Newton-Deviationsgleichung. Sie beschreibt uns die Relativbewegung von Objekten, die sich im freien Fall des Gravitationspotentials befinden. Sie beantwortet die Frage: Wie fällt ein Objekt relativ zu mir? Wenn man sich nun vorstellt, man wäre in einem fallenden Fahrstuhl und man hat einen Ball in seiner Hand, kann man den Gradienten des Gravitationspotentials  $\partial_i\Phi(x^m(t, s))$  mit der Fallbeschleunigung substituieren und man erhält die homogene Bewegungsgleichung  $\ddot{\xi}^i = 0$ . Und das würde ja auch passieren, wenn man den Ball aus seiner Hand lässt, würde er sich, relativ zu uns, linear bewegen. Wenn man nun das System erweitert, vergrößert, dann kann man diese Vereinfachung nicht machen, und man erhält Abweichungen, die dem tensoriellen Gradienten

des Gravitationspotentials  $\partial_k \partial_i \Phi(x^m(t, s))$  entsprechen. Um nun ein Analogon für eine gekrümmte Raumzeit zu erhalten, müssen wir besprechen, was Lie-Ableitungen für Vektorfelder sind.

Sei  $\xi^a$  ein Vektorfeld auf unserer Mannigfaltigkeit. Dann kann man  $\xi^a$  als *Strömung* auffassen. Wir definieren nun die *Stromlinie* oder Integralkurve:  $\dot{x}^a(t) = \xi^a(x^b(t))$ , also der Tangentenvektor einer Kurve ist gleich der Strömung an dem Punkt des Tangentenvektors. Das ist einsichtig, wenn man sich einen Fluss vorstellt und einen Korken hineinwirft. Die Kurve, die der Korken macht, ist unser  $x^i(t)$  und an jedem Punkt der Kurve ist die Geschwindigkeit  $\dot{x}^a(t)$  (Tangentenvektor) gleich der Strömung an diesem Punkt  $\xi^a(x^b(t))$ . Nun definieren wir die *Flussabbildung*:

$$\Phi_t^i(x_0^m) := x^i(t). \quad (5.23)$$

Sie zeigt uns den Verlauf einer Kurve für einen gewissen Anfangswert  $x^i(t=0) = x_0^i$  aus einer Schar von Anfangswerten  $x_0^i \in B$  wobei B ein Bereich auf unserer Mannigfaltigkeit ist. Für die Flussabbildung gilt:

$$\Phi_0^i(x_0^m) = \mathbb{I}(x_0^i) = x_0^i \quad (5.24)$$

$$\Phi_t^i(\Phi_s^j(x_0^m)) = \Phi_{i+j}^i(x_0^m) \quad (5.25)$$

Also für den Zeitpunkt  $t=0$  fließt die Kurve entlang der Strömung nicht und die Flussabbildung gibt den Anfangswert zurück, wirkt also wie die Identität  $\mathbb{I}$ . Darüber hinaus soll die Zusammensetzung von zwei gleichen Flussabbildungen für verschiedene Zeiten gleich der Flussabbildung für die Summe der Zeiten sein. Das macht auch bildlich Sinn: Man betrachte wieder den Korken im Fluss. Als erstes beobachtet man ihn für  $t$  Sekunden, dabei fließt der entlang der Strömung. Dann ab dem Zeitpunkt  $t$  befindet sich der Korken bei  $x^i(t)$ . Dann betrachten wir den Korken abermals, dieses Mal für  $s$  Sekunden. Dann befindet er sich dieses Mal bei  $\Phi_s^j(\Phi_t^i(x_0^m))$ , weil hier der Anfangswert der Endwert der letzten Beobachtung war. Dieser Vorgang ist aber equivalent dazu, den Korken einfach eine Zeitdauer  $t+s$  zu beobachten.

Man kann sich schon irgendwie denken, dass man aus der Flussabbildung das Strömungsvektorfeld herausextrahieren kann. Betrachten wir die zeitliche Ableitung der Flussabbildung:

$$\frac{d}{dt} \Phi_t^i(x_0^i)|_{t=0} = \dot{x}^i(t=0) = \xi^i(x^m(t=0)) = \xi^i(x_0^m), \quad (5.26)$$

also die zeitliche Änderung der Flussabbildung zu einem gewissen Zeitpunkt  $t=0$ , der ja willkürlich ist, ist gleich dem Strömungsvektorfeld  $\xi^a$ .

Betrachten wir nun ein Skalarfeld auf unserer Mannigfaltigkeit unter dem Einfluss der Strömung  $\xi^a$ . Es sei  $f(x^m)$  ein Skalarfeld (z.B.: die Temperatur einer Flüssigkeit an einer Stelle  $x^m$ ). Dann ist  $\tilde{f}(x^m)$  das Skalarfeld, welches mit der Strömung entlangfließt. Dann gilt  $\tilde{f}(x^m) = f(\Phi_t^i(x^m))$ . Nun

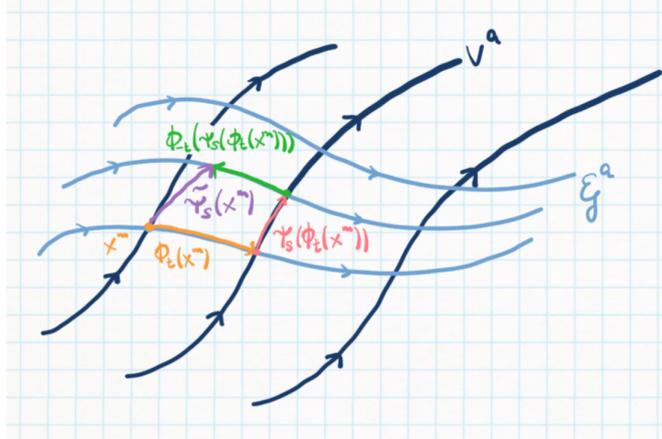


Abbildung 5.1: Zur Illustration der Herleitung der Flussabbildung  $\tilde{\psi}_s$ .

wollen wir die zeitliche Ableitung dieses strömenden Skalarfeldes berechnen:

$$\frac{d}{dt} \tilde{f}(x^m)|_{t=0} = \frac{d}{dt} f(\Phi_t^i(x^m))|_{t=0} = \partial_i f(x^m) \xi^l(x^m) = \delta_l^k \partial_k f(x^m) \xi^l(x^m) = \quad (5.27)$$

$$\partial_i^a dx_a^k \partial_k f(x^m) \xi^l(x^m) = \partial_i^a \xi^l(x^m) dx_a^k \partial_k f(x^m) = \xi^a \partial_a f(x^m) := (L_\xi f)(x^m). \quad (5.28)$$

Das ist die sogenannte *Lie-Ableitung* für Skalarfelder. Sie beschreibt die zeitliche Änderung eines Skalarfeldes  $f$  unter einer Strömung  $\xi^a$  in erster Ordnung. Die koordinaten-kovariante Ableitung ist aber nur lokal definiert, deshalb wollen wir unsere Levi-Cevitá Ableitung verwenden:

$$(L_\xi f)(x^m) = \xi^a \nabla_a f(x^m). \quad (5.29)$$

Nun kommt der nächste Schritt, wie sieht es aus für Tensoren höherer Stufe, bzw. was ist  $L_\xi v^a$ ? Dafür betrachten wir zwei Vektorfelder  $\xi^a$  und  $v^a$ . Beide haben eine Flussabbildung, die man wie in Gl.(5.26) herleiten kann. Für unser Vektorfeld  $v^a$  sei  $\psi_s$  die Flussabbildung. Nun wollen wir eine Flussabbildung für  $v^a$  bei der der Einfluss der Strömung von  $\xi^a$  konstruieren. Man betrachte dafür Abb. 5.1. Dort sind beide Strömungen visualisiert. Um eine Flussabbildung  $\tilde{\psi}_s$  zu konstruieren, die die Strömung  $\xi^a$  beinhaltet, gehen wir von einem Punkt  $x^m$  aus. Von dort strömen wir entlang der Strömung  $\xi^a$  ein bisschen. Dann strömen wir entlang unserer Strömung  $v^a$  und strömen dann in negativer Richtung (deshalb mit  $-t$  gekennzeichnet) wieder entlang  $\xi^a$ . Dabei wird man nicht zum selben Punkt gelangen, als wäre man von  $x^m$  entlang  $v^a$  geströmt, doch das wollen wir ja auch erzielen. Die Flussabbildung lautet:

$$\tilde{\psi}_s^k(x^m) := \Phi_{-t}^k(\psi_s^j(\Phi_t^i(x^m))). \quad (5.30)$$

Eine wichtige Eigenschaft der Flussabbildung ist Gl.(5.24), die auch  $\tilde{\psi}_s$  erfüllen muss. Betrachten

wir  $\psi_{s=0}^k(x^m)$ :

$$\psi_{s=0}^k(x^m) = \Phi_{-t}^k(\psi_{s=0}^j(\Phi_t^i(x^m))) = \Phi_{-t}^k(\Phi_t^i(x^m)) = \Phi_{-t+t}^k(x^m) = x^m, \quad (5.31)$$

also unsere Bedingung ist erfüllt. Wie wir in Gl.(5.26) das Vektorfeld aus der Flussabbildung extrahieren konnten, können wir auch hier das Vektorfeld  $\tilde{v}^a$  extrahieren, welches unter dem Einfluss der Strömung  $\xi^a$  strömt:

$$\tilde{v}^i(x^m) = \frac{d}{ds} \psi_s^i(x^m)|_{s=0} = \frac{\partial \Phi_{-t}^i}{\partial x^l}(\Phi_t^j(x^m)) \frac{d\psi_s^l}{ds}(\Phi_t^j(x^m)) = \frac{\partial \Phi_{-t}^i}{\partial x^l}(\Phi_t^j(x^m)) v^l(\Phi_t^j(x^m)), \quad (5.32)$$

wobei wir  $\frac{d\psi_s^l}{ds} = v^l$  genutzt haben. Wie wir in Gl.(5.27) die Lie-Ableitung für Skalarfelder herleiteten, können wir hier dasselbe für Tensorfelder der Stufe (1,0) machen:

$$\frac{d}{dt} \tilde{v}^i(x^m)|_{t=0} = -\frac{d}{dx^l} \frac{d\Phi_t^i}{dt} v^l + \delta_l^i \partial_k v^l \xi^k = \frac{d}{dx^l} \xi^i v^l + \partial_i v^l \xi^i = L_\xi v^i, \quad (5.33)$$

wobei bei jedem Vektorfeld die Abhängigkeit von  $x^m$  für Übersichtlichkeit unterschlagen worden ist. Das ist die Lie-Ableitung für die Komponenten des Vektorfeldes  $v^a$ . In Basisschreibweise:

$$L_\xi v^a = (-\partial_l \xi^i v^l + \partial_k v^i \xi^k) \partial_i^a \quad (5.34)$$

$$L_\xi v^a = (-\delta_l^k \partial_k \xi^i v^l + \delta_l^k \partial_k v^i \xi^l) \partial_i^a \quad (5.35)$$

$$L_\xi v^a = (-\partial_l^b dx_b^k \partial_k \xi^i v^l \partial_i^a + \partial_l^b dx_b^k \partial_k v^i \xi^l \partial_i^a) \quad (5.36)$$

$$L_\xi v^a = \xi^b \partial_b v^a - v^b \partial_b \xi^a, \quad (5.37)$$

wobei wir  $\delta_l^k = \partial_l^b dx_b^k$  und dann die Basisschreibweise der Vektoren (z.B.:  $v^a = v^i \partial_i^a$ ) ausnutzten. Wie vorher müssen wir die koordinaten-kovariante Ableitung mit der Levi-Cevitá Ableitung austauschen, um die Globalität der Lie-Ableitung zu gewährleisten:

$$L_\xi v^a = \xi^b \nabla_b v^a - v^b \nabla_b \xi^a. \quad (5.38)$$

Falls wir aber doch irgendwie eine andere Ableitung finden, die auch  $\tilde{\nabla}_a g_{bc} = 0$  erfüllt, müssen wir nicht besorgt sein, denn die Lie-Ableitung für Vektorfelder ist eindeutig:

$$\xi^b \tilde{\nabla}_b v^a - v^b \tilde{\nabla}_b \xi^a = \xi^b (\nabla_b v^a + C_{mb}^c v^m) - v^b (\nabla_b \xi^a + C_{mb}^c \xi^m) = \quad (5.39)$$

$$\xi^b \nabla_b v^a - v^b \nabla_b \xi^a + \xi^b C_{mb}^c v^m - v^b C_{mb}^c \xi^m = \xi^b \nabla_b v^a - v^b \nabla_b \xi^a + \xi^b C_{mb}^c v^m - v^b C_{mb}^c \xi^m = \quad (5.40)$$

$$\xi^b \nabla_b v^a - v^b \nabla_b \xi^a + \xi^b C_{mb}^c v^m - \xi^m C_{bm}^c v^b = \xi^b \nabla_b v^a - v^b \nabla_b \xi^a, \quad (5.41)$$

da der Differenzentensor  $C_{mb}^c$  symmetrisch in den unteren zwei Indizes ist. Nun wollen wir wissen, was die Lie-Ableitung für kovariante Vektorfelder  $\omega_a$  ist. Dafür betrachten wir die Lie-Ableitung von

$\omega_b v^b$ , unter Verwendung der Leibniz-Regel und Gl. (5.29) und Gl. (5.38):

$$L_\xi(\omega_b v^b) = L_\xi(\omega_b v^b) \quad (5.42)$$

$$L_\xi \omega_b v^b + \omega_b L_\xi v^b = (\xi^a \nabla_a)(\omega_b v^b) \quad (5.43)$$

$$L_\xi \omega_b v^b + \omega_b \xi^a \nabla_a v^b - \omega_b v^a \nabla_a \xi^b = \xi^a \nabla_a \omega_b v^b + \omega_b \xi^a \nabla_a v^b \quad (5.44)$$

$$L_\xi \omega_b v^b = \omega_b v^a \nabla_a \xi^b + \xi^a \nabla_a \omega_b v^b \quad (5.45)$$

$$L_\xi \omega_b = \xi^a \nabla_a \omega_b + \omega_b \nabla_a \xi^b. \quad (5.46)$$

Hier wurde schon in weise Voraussicht die Levi-Cevitá Ableitung eingefügt. Nun können wir uns der Herleitung der Einstein'schen Feldgleichungen zuwenden.

### 5.3 Der Einstein-Tensor

Der Einstein-Tensor verbindet die im System befindliche Energie/Materie mit der Krümmung der Raumzeit. Um eine derartige Gleichung aufzustellen, hat Einstein 1914 ein geometrisches Analogon der Newton-Deviationsgleichung hergeleitet und das machen wir hier heute auch.

Wir betrachten eine Schar an kräftefreie Objekte mit Geschwindigkeitsvektor  $u^a$ . Da keine Gravitation wirkt, aber die Raumzeit gekrümmt wird, sind die Kurven der Körper Autoparallele, d.h. sie genügen der Gl. (5.1) oder hier für nochmals wiederholt:

$$(u^a \nabla_a)u^b = 0. \quad (5.47)$$

Dabei repräsentiert  $u^a$  ein zeitartiges ( $u^2 = -1$ ), geodätisches Vektorfeld auf unserer Mannigfaltigkeit, da für jeden Punkt eine Kurve existiert und deren Tangentenvektor dort gebildet werden kann. Nun stellen wir uns eine Kurve  $\gamma$  vor, die in einer beliebigen Richtung zu den Autoparallelen verläuft. Dann nennen wir  $\Phi_t(\gamma)$  eine Kurve, bei der jeder Punkt von der ursprünglichen Kurve unter der Strömung  $u^a$  entstanden ist. Dafür kann man sich ein Seil in einem Fluss senkrecht zur Flussrichtung vorstellen: Zu jedem Zeitpunkt  $t$  hat das Seil die Form wie zum Zeitpunkt  $t=0$  mit dem Einfluss der Strömung, jeder Punkt auf dem Seil wird mit dem dort befindlichen Vektor des Vektorfeldes  $u^a$  nach vorne bewegt. Das Tangentenvektorfeld der Kurve  $\gamma$  existiert damit auf der Raumzeit und soll  $\xi^a$  heißen. Da wir in der Konstruktion des Vektorfeldes die Querströmung  $u^a$  miteinbezogen, gilt folgende Beziehung:

$$L_u \xi^a = 0. \quad (5.48)$$

Die Lie-Ableitung beschreibt ja die Änderung des Vektorfeldes unter dem Einfluss eines zusätzlichen Vektorfeldes. Wenn wir nun das Vektorfeld mit der Wirkung der Strömung herleiten, ist die Änderung bzgl. der Strömung 0. Explizit ausgeschrieben heißt Gl.(5.47):

$$L_u \xi^a = u^b \nabla_b \xi^a - \xi^b \nabla_b u^a = 0. \quad (5.49)$$

Dabei beschreibt  $u^b \nabla_b \xi^a$  die zeitliche Ableitung bezüglich dem Beobachter mit Tangentenvektor  $u^a$ . Sie zeigt die Änderung von  $\xi^a$  nach seiner Eigenzeit an. Eine (zweifache) zeitliche Ableitung kommt auch in der Newton-Deviationsgleichung vor, deshalb ist unser erster Verdacht zur *geodätischen Deviationsgleichung*, dass wir  $(u^a \nabla_a)^2 \xi^b$  bilden:

$$(u\nabla)^2 \xi^a = (u\nabla)(u^b \nabla_b) \xi^a \quad (5.50)$$

$$(u\nabla)^2 \xi^a = (u^c \nabla_c)(\xi^b \nabla_b u^a) \quad (5.51)$$

$$(u\nabla)^2 \xi^a = u^c \nabla_c (\xi^b) \nabla_b u^a + \xi^b u^c \nabla_c (\nabla_b u^a) \quad (5.52)$$

$$(u\nabla)^2 \xi^a = u^c \nabla_c (\xi^b) \nabla_b u^a + \xi^b u^c \nabla_b (\nabla_x u^a) + R_{dcb}^a u^d \xi^b u^c, \quad (5.53)$$

Sieht unnötig kompliziert aus. Wir müssen nun betrachten was  $\xi^b \nabla_b \cdot 0 = \xi^b \nabla_b (u^c \nabla_c u^a)$  ist:

$$\xi^b \nabla_b \cdot 0 = \xi^b \nabla_b (u^c \nabla_c u^a) = \xi^b \nabla_b (u^c) \nabla_c (u^a) + \xi^b u^c \nabla_b \nabla_c u^a = 0. \quad (5.54)$$

D.h. wir können  $\xi^b u^c \nabla_b (\nabla_c u^a)$  mit  $-\xi^b \nabla_b (u^c) \nabla_c (u^a)$  in Gl. (5.53) ersetzen:

$$(u\nabla)^2 \xi^a = u^c \nabla_c (\xi^b) \nabla_b u^a + \xi^b \nabla_b (u^c) \nabla_c (u^a) + R_{dcb}^a u^d \xi^b u^c \quad (5.55)$$

$$(u\nabla)^2 \xi^a = (u\nabla)(\xi^b) \nabla_b u^a - (\xi\nabla)u^b \nabla_b (u^a) + R_{dcb}^a u^d \xi^b u^c \quad (5.56)$$

$$(u\nabla)^2 \xi^a = ((u\nabla)(\xi^b) - (\xi\nabla)u^b) \nabla_b u^a + R_{dcb}^a u^d \xi^b u^c \quad (5.57)$$

$$(u\nabla)^2 \xi^a = L_u \xi^b \nabla_b u^c + R_{dcb}^a u^d \xi^b u^c \quad (5.58)$$

$$(u\nabla)^2 \xi^a = R_{dcb}^a u^d \xi^b u^c. \quad (5.59)$$

Das kann man unter Verwendung der Antisymmetrie des Riemann-Tensors in den zwei letzten unteren Indizes umschreiben zu:

$$(u\nabla)^2 \xi^a + R_{abc}^a u^d u^c \xi^b = 0. \quad (5.60)$$

Das ist die *Geodätische Deviationsgleichung*. Hier können wir eine Verbindung zwischen dem Riemann-Tensor und dem tensoriellen Gradienten des Gravitationspotential in Gl. (5.22) machen. Es gilt also:

$$R_{bcd}^a u^b u^d = \partial^a \partial_c \Phi, \quad (5.61)$$

für die Korrelation zwischen den Indizes wurde die Metrik auf der linken Seite verwendet. Da aber  $\partial^a$  Ableitungen in jeder Raumzeit-Koordinate darstellt und  $\Phi$  nur vom Ort abhängt macht die Metrikanwendung keinen Unterschied<sup>1</sup>. Das Gravitationspotential erfüllt eine Poisson-Gleichung  $\Delta \Phi = 4\pi \rho_m$ , wobei  $\rho_m$  die Materiedichte ist. Da der Laplace-Operator  $\Delta = \partial^a \partial_a$  ist, müssten wir Gl.(5.61) kon-

<sup>1</sup>Im Mathematical Interlude zur Abstrakten Indexschreibweise lernten wir die Metrik kennen. Dort sahen wir, dass die kovariante zeitliche Komponente eines Vektors ( $v_0 = v^\nu \eta_{0\nu}$ ) nur ein anderes Vorzeichen hat. Da die zeitliche Ableitung aber von sich aus 0 wäre, gilt  $\partial^a \Phi(x^m) = \partial_a \Phi(x^m)$ .

trahieren, um die Materiedichte in Beziehung zum Riemann-Tensor zu bringen:

$$R_{bd}u^b u^d = \Delta\Phi = 4\pi\rho_m = 4\pi T_{ab}u^a u^b, \quad (5.62)$$

wobei der letzte Schritt der Elektrodynamik verschuldet ist. Die Spur vom Riemann-Tensor  $R_{bad}^a = R_{bd}$  heißt Ricci-Tensor und  $T_{ab}$  ist der *Energie-Impuls-Tensor* bekannt aus der Elektrodynamik, der den Fluss von Energie und Impuls für ein System beschreibt. Einstein hat 1914 dann gefolgert, dass folgende Gleichung gelten muss:

$$R_{ab} = 4\pi T_{ab}, \quad (5.63)$$

doch leider vergebens, denn diese Beziehung ist falsch. Denn, wieder aus der Elektrodynamik, ist bekannt, dass der Energie-Impuls-Tensor divergenzfrei ist:  $\nabla^a T_{ab} = 0$ , was aber nicht für den Ricci-Tensor gilt.  $\nabla^a R_{ab} = 0$  gelte nicht allgemein, weshalb Gl. (5.63) nicht richtig sein kann. Um den Fehler zu beheben, brauchen wir die 2 Bianchi Identitäten, die allgemeine Aussagen über den Riemann-Tensor machen. Die Bianchi-Identitäten seien im Folgenden ohne expliziten Beweis<sup>2</sup> aufgelistet:

1.  $R_{abc}^m + R_{cab}^m + R_{bca}^m = 0$
2.  $\nabla_a R_{nbc}^d + \nabla_c R_{nab}^d + \nabla_b R_{nca}^d = 0$

Kontrahiert man nun in der zweiten Bianchi-Identität die Indizes a und d und wendet dann die Metrik  $g^{nb}$  an, erhält man:

$$\nabla_a R_{nbc}^a g^{nb} + \nabla_c R_{nbg}^{nb} - \nabla_b R_{nc} g^{nb} = 0 \quad (5.64)$$

$$-\nabla_a R_c^a + \nabla_c R - \nabla_a R_c^a = 0 \quad (5.65)$$

$$\nabla_a R_b^a = \frac{1}{2} \nabla_b R. \quad (5.66)$$

$R = R_{ab}g^{ab}$  ist der *Krümmungsskalar*. Damit haben wir also gezeigt, dass der Gradient des Ricci-Tensors allgemein nicht 0 ist. Durch Umformung mit der Metrik, kann man aber ein Gebilde erzeugen, was wie der Energie-Impuls-Tensor divergenzfrei ist:

$$\nabla_c R_b^c = \frac{1}{2} \nabla_b R \quad (5.67)$$

$$\nabla^a g_{ac} R_b^c = \frac{1}{2} \nabla^a g_{ab} R \quad (5.68)$$

$$\nabla^a R_{ab} = \frac{1}{2} \nabla^a g_{ab} R \quad (5.69)$$

$$\nabla^a (R_{ab} - \frac{1}{2} g_{ab} R) = 0. \quad (5.70)$$

Und hier sind wir, wo auch Einstein im Jahre 1915 war: Damit konnte er die Verbindung zwischen der Geometrie der Raumzeit mit der Materie ziehen. Durch Gleichsetzung von Gl.(5.70) und  $\nabla^a T_{ab} = 0$

<sup>2</sup>Für die jeweilige Bianchi-Identität muss man jeweils  $[\nabla_a, [\nabla_b, \nabla_c]]f = [[\nabla_a, \nabla_b], \nabla_c]f + [\nabla_b, [\nabla_a, \nabla_c]]f$  und  $[\nabla_a, [\nabla_b, \nabla_c]]v^d = [[\nabla_a, \nabla_b], \nabla_c]v^d + [\nabla_b, [\nabla_a, \nabla_c]]v^d$  betrachten, wobei f ein Skalarfeld und  $v^d$  ein Vektorfeld darstellt.

können wir die *Einstein'schen Feldgleichungen* aufstellen:

$$G_{ab} = R_{ab} - \frac{1}{2}g_{ab}R = 8\pi T_{ab}. \quad (5.71)$$

$G_{ab}$  heißt auch *Einstein-Tensor*. Der Faktor 2 auf der rechten Seite rührt daher, dass man hier einfache Materiedichte annimmt, eine der Form  $T_{ab} = \rho_m u_a u_b$ . Bildet man nun die Spur von  $G_{ab}$  erhält man  $G = R - 2R = -R = -8\pi\rho_m \rightarrow R = 8\pi\rho_m$ <sup>3</sup>. Kontrahiert man nun Gl.(5.71) mit  $u^a u^b$  erhält man:

$$G_{ab}u^a u^b = R_{ab}u^a u^b - \frac{1}{2}g_{ab}u^a u^b R = 8\pi T_{ab}u^a u^b = 8\pi\rho_m \quad (5.72)$$

$$4\pi\rho_m - \frac{1}{2}u_a u^b R = 4\pi\rho_m + \frac{1}{2}R = 4\pi\rho_m + \frac{1}{2}8\pi\rho_m = 8\pi\rho_m \quad (5.73)$$

$$8\pi\rho_m = 8\pi\rho_m, \quad (5.74)$$

was den Faktor 2 bestätigt.

Also haben wir nun die Einstein'schen Feldgleichungen postuliert, jetzt müssen wir sie noch lösen, also die Metrik finden und nach Hause gehen. Ganz so leicht ist es aber nicht. Denn um die Metrik zu finden, muss man  $\nabla^a T_{ab} = 0$  fordern. Aber die Ableitung kennen wir noch nicht, die wird erst durch die Metrik bestimmt, die kennen wir aber nicht, genau die suchen wir ja! Wenn das so ist, geben wir einfach eine Metrik vor und finden dazu dann den Energie-Impuls-Tensor. Das funktioniert auch nicht ganz so einfach, da meist ein negativer  $T_{ab}$  herauskommt, was unphysikalisch ist. Also was macht man?

Tatsächlich gab es schon 1916 eine exakte Lösung zur Gl. (5.71), nämlich  $T_{ab} = 0$ . Hört sich trivial an „Ja das ist doch der freie Raum, da gibt es keine Raumkrümmung!“. Da aber Gl.(5.71) *nicht-linear* ist, gibt es noch zusätzliche Lösungen. Welche die vorhersagen, dass Raumkrümmung *durch Raumkrümmung* entsteht. Raumkrümmung besitzt selber also Energie, und diese selbsterzeugende Raumkrümmung nennt man *Gravitationswellen*.

---

<sup>3</sup> $T = \rho_m u_a u^a = \rho_m(-1)$ , da wir einen zeitartigen Beobachter mit Geschwindigkeit  $u^a$  betrachten.

# Kapitel 6

## Schluss

Und hier, wo es erst so richtig anfängt wird, endet das Skriptum. Wie geht es nun weiter? Ganz einfach: In Geometrie und Gravitation I im Sommersemester. Ich bedanke mich an die Leserschaft und hoffe, dass dieses Skriptum hilfreich für das Verständnis des Stoffes. Ziel dieses Skriptums war es in Prof. Balasins meist verwirrendem Vorgehen etwas Einsicht zu bringen. Fehler bitte an [a.miethlinger@gmail.com](mailto:a.miethlinger@gmail.com) melden!

Ich wünsche den hier lesenden Studierenden alles Gute zur Prüfung!

-Alexander Miethlinger