Einführung in die Methoden der speziellen Relativitätstheorie oder „Wie rechnete Einstein?: Teil I"

In der speziellen, wie allgemeinen Relativitätstheorie sind einige Methoden und Schreibweisen gängig, die seltsam anmuten, wenn man sie das erste mal präsentiert bekommt. Es handelt sich hierbei jedoch um ganz allgemeine Prinzipien, so wie diese in der Differentialgeometrie eingeführt werden, jenem mathematischen Fachgebiet, das Methoden der Goemetrie und der Differentialrechnung vereinigt. Im Folgenden soll auf diese neuen Begrifflichkeiten anhand von unterschiedlichen Beispielen eingegangen werden. Die Notation ist so, dass Summenzeichen (meistens) weggelassen werden. Dann ist über doppelte Indizes zu summieren.

Als Aufgangspunkt betrachten wir das Skalarprodukt zweier Vektoren im dreidimensionalen (flachen) Raum, so wie wir es aus der Schule kennen:

$$\vec{a} \cdot \vec{b} = \sum_{\substack{i=1\\j=1}}^{3} a_i \mathbf{e}^i \cdot b_j \mathbf{e}^j = \sum_{\substack{i=1\\j=1}}^{3} (\mathbf{e}^i \mathbf{e}^j) a_i b_j = \sum_{\substack{i=1\\j=1}}^{3} g^{ij} a_i b_j \,,$$

mit der Definition

$$g^{ij} = \mathbf{e}^i \cdot \mathbf{e}^j \,.$$

Die definierte Größe g^{ij} ist ein Tensor. Dies folgt daraus, dass die Einheitsvektoren $\mathbf{e}^i$ und $\mathbf{e}^j$ selbst Tensoren (erster Stufe) sind. Sie transformieren sich nämlich nach den Gesetzen

$$\mathbf{e}'^i = A^i{}_k \mathbf{e}^k \,,$$

bzw.

$$\mathbf{e}'^j = A^j{}_l \mathbf{e}^l \,,$$

womit durch Einsetzen folgt:

$$g'^{ij} = \mathbf{e}'^i \cdot \mathbf{e}'^j = A^i{}_k \mathbf{e}^k \cdot A^j{}_l \mathbf{e}^l = A^i{}_k A^j{}_l (\mathbf{e}^k \cdot \mathbf{e}^l) = A^i{}_k A^j{}_l g^{kl} \,.$$

Hier steht gerade das Transformationsgesetz für einen Tensor zweiter Stufe! Der Tensor g^{ij} ist sehr wichtig, sowohl in der Geometrie als auch in der Physik. Man bezeichnet ihn als **metrischen Tensor**. In der Physik ist er Grundlage aller Berechnungen der Relativitätstheorie. Da g^{ij} ein Tensor zweiter Stufe ist, lässt sich dieser in Matrixschreibweise aufschreiben. Im Folgenden wird mit g^{ij} immer eine Komponente des Tensors bezeichnet, wobei (g^{ij}) für die Matrixdarstellung des Tensors steht. Im Falle des dreidimensionalen flachen Raumes gilt nun:

$$(g^{ij}) = \begin{pmatrix} \mathbf{e}^1 \cdot \mathbf{e}^1 & \mathbf{e}^1 \cdot \mathbf{e}^2 & \mathbf{e}^1 \cdot \mathbf{e}^3 \\ \mathbf{e}^2 \cdot \mathbf{e}^1 & \mathbf{e}^2 \cdot \mathbf{e}^2 & \mathbf{e}^2 \cdot \mathbf{e}^3 \\ \mathbf{e}^3 \cdot \mathbf{e}^1 & \mathbf{e}^3 \cdot \mathbf{e}^2 & \mathbf{e}^3 \cdot \mathbf{e}^3 \end{pmatrix} .$$

Die Einträge sind also gegeben durch Skalarprodukte der jeweiligen Einheitsvektoren. Mit der euklidischen Standardbasis

$$\mathbf{e}^1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} , \quad \mathbf{e}^2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} , \quad \mathbf{e}^3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} ,$$

folgt hieraus:

$$(g^{ij}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathbb{1}_3 ,$$

also gerade die dreidimensionale Einheitsmatrix. Es gilt also im $\mathbb{R}^3$ für das Skalarprodukt zweier Vektoren:

$$\mathbf{a} \cdot \mathbf{b} = \sum_{ij} g^{ij} a_i b_j = (a_1, a_2, a_3) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = a_1 b_1 + a_2 b_2 + a_3 b_3 ,$$

so wie wir es kennen.

Ein bisschen anders verhält sich das Ganze in anderen Koordinaten oder sogar gekrümmten Räumen. Wird ein solcher Raum parametrisiert durch eine Funktion $\mathbf{x}(u_1, u_2, \ldots)$, wobei u_1, u_2, … die Koordinaten des Raumes sind, so gilt für den metrischen Tensor:

$$g^{ij} = \frac{\partial \mathbf{x}}{\partial u_i} \cdot \frac{\partial \mathbf{x}}{\partial u_j} .$$

Betrachten wir hierzu als Beispiel zunächst einen flachen Raum, nämlich den $\mathbb{R}^2$. Parametrisiert kann dieser werden durch

$$\mathbf{x}(u, v) = \begin{pmatrix} u \\ v \end{pmatrix} ,$$

wobei $u \in (-\infty, \infty)$ und $v \in (-\infty, \infty)$ die Wertebereiche der Parameter u und v sind. Dann gilt:

$$\frac{\partial \mathbf{x}}{\partial u} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} , \quad \frac{\partial \mathbf{x}}{\partial v} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} .$$

Also ist unser metrischer Tensor gegeben durch

$$(g^{ij}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} .$$

Man kann den $\mathbb{R}^2$ auch mittels Polarkoordinaten parametrisieren:

$$\mathbf{x}(r, \varphi) = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix} ,$$

mit dem Wertebereich $r \in [0, \infty)$ und $\varphi \in [0, 2\pi)$. Dann erhalten wir

$$\frac{\partial \mathbf{x}}{\partial r} = \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} , \quad \frac{\partial \mathbf{x}}{\partial \varphi} = \begin{pmatrix} -r \sin \varphi \\ r \cos \varphi \end{pmatrix} ,$$

und es folgt der metrische Tensor:

$$(g^{ij}) = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix} .$$

In anderen Koordinatensystemen muss also der metrische Tensor nicht mehr unbedingt die Einheitsmatrix sein. Seine Aussagekraft ist beschränkt, hängt er doch von den zugrundeliegenden Koordinaten, die man verwendet, ab. (Größen, welche die innere Geometrie eines Raumes widerspiegeln sind die sogenannten Christoffel-Symbole oder der Riemannsche Krümmungstensor, was jedoch den Rahmen der kleinen Exkursion hier sprengen würde. Darauf soll deshalb nicht näher eingegangen werden.)

Betrachten wir als nächstes Beispiel eine Kugeloberfläche:

$$\mathbf{x} = \begin{pmatrix} r \sin \vartheta \cos \varphi \\ r \sin \vartheta \sin \varphi \\ r \cos \vartheta \end{pmatrix} .$$

Mit

$$\frac{\partial \mathbf{x}}{\partial r} = \begin{pmatrix} \sin \vartheta \cos \varphi \\ \sin \vartheta \sin \varphi \\ \cos \vartheta \end{pmatrix} , \quad \frac{\partial \mathbf{x}}{\partial \vartheta} = \begin{pmatrix} r \cos \vartheta \cos \varphi \\ r \cos \vartheta \sin \varphi \\ -r \sin \vartheta \end{pmatrix} , \quad \frac{\partial \mathbf{x}}{\partial \varphi} = \begin{pmatrix} -r \sin \vartheta \sin \varphi \\ r \sin \vartheta \cos \varphi \\ 0 \end{pmatrix} ,$$

folgt der metrische Tensor:

$$(g^{ij}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2 \vartheta \end{pmatrix} .$$

Somit können wir spaßeshalber das Skalarprodukt zweier infinitesimaler Vektoren $\mathrm{d}\mathbf{x}$ auf der Kugeloberfläche aufschreiben:

$$\begin{aligned}\mathrm{d}\mathbf{x}\cdot\mathrm{d}\mathbf{x} &= \sum_{i,j} g^{ij}(\mathrm{d}x)^i(\mathrm{d}x)^j = (\mathrm{d}r,\mathrm{d}\vartheta,\mathrm{d}\varphi)\cdot\begin{pmatrix}1 & 0 & 0\\ 0 & r^2 & 0\\ 0 & 0 & r^2\sin^2\vartheta\end{pmatrix}\cdot\begin{pmatrix}\mathrm{d}r\\ \mathrm{d}\vartheta\\ \mathrm{d}\varphi\end{pmatrix} = \\ &= \mathrm{d}r^2 + r^2(\mathrm{d}\vartheta^2 + \sin^2\vartheta\mathrm{d}\varphi^2)\,.\end{aligned}$$

Dieses sogenannte **Längenquadrat** ist eine wichtige Größe in der Relativitätstheorie.

Kommen wir nun nach diesen etlichen anschaulichen Beispielen zu derselbigen. Hier haben wir einen vierdimensionalen Raum („Raum-Zeit“), den man als Minkowski-Raum bezeichnet. Raum und Zeit bilden ein Gefüge und können nicht mehr unabhängig und getrennt voneinander betrachtet werden. Der Minkowski-Tensor, welcher diesen Raum beschreibt, ist gegeben durch:

$$(g^{\mu\nu}) = \begin{pmatrix}1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & -1 & 0\\ 0 & 0 & 0 & -1\end{pmatrix} = \mathrm{diag}(1,-1,-1,-1)\,.$$

In der Relativitätstheorie verwendet man außerdem keine lateinischen Indizes mehr, sondern griechische. Wir definieren den sogenannten **kontravarianten Vierervektor** in diesem Raum durch:

$$x^\mu = \begin{pmatrix}x^0\\ x^1\\ x^2\\ x^3\end{pmatrix}\,,$$

wobei immer bei Null mit der Zählung der Indizes angefangen wird. An diesem Vektor ist nichts Ungewöhnliches. Schauen wir uns nun das Skalarprodukt des Vektors mit sich selbst an.

$$g_{\mu\nu}x^\mu x^\nu = x^\mu g_{\mu\nu}x^\nu := x^\mu x_\mu = x_\mu x^\mu\,,\quad x_\mu = g_{\mu\nu}x^\nu\,.$$

Man benötigt also, wenn man das Skalarprodukt ohne metrischen Tensor $g_{\mu\nu}$ schreibt, einen zweiten Vektor mit einem unteren Index, der über $x_\mu = g_{\mu\nu}x^\nu$ definiert ist. Was bedeutet es, wenn der Index eines solchen Vektors unten steht? Betrachten wir dazu das Ganze in Matrixschreibweise:

$$(x^\mu)(x_\mu) = (x^0,x^1,x^2,x^3)\cdot\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & -1 & 0\\ 0 & 0 & 0 & -1\end{pmatrix}\cdot\begin{pmatrix}x^0\\ x^1\\ x^2\\ x^3\end{pmatrix} = (x^0)^2-(x^1)^2-(x^2)^2-(x^3)^2\,.$$

Dasselbe Ergebnis erhält man mit der Wahl

$$x_\mu = \begin{pmatrix} x^0 \\ -x^1 \\ -x^2 \\ -x^3 \end{pmatrix} ,$$

wenn man also die Ortskomponenten des **kontravarianten Vierervektors** mit Minuszeichen versieht. Dann gilt nämlich, wenn man das Skalarprodukt ohne metrischen Tensor berechnet, so wie wir es die ganze Zeit immer getan haben:

$$(x^\mu) \cdot (x_\mu) = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \cdot \begin{pmatrix} x^0 \\ -x^1 \\ -x^2 \\ -x^3 \end{pmatrix} = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 .$$

Den Vektor x_μ mit einem unteren Index nennt man **kovarianten Vierervektor**. Ko- und kontravariante Vektoren stehen mittels des metrischen Tensors miteinander in Verbindung:

$$x_\mu = g_{\mu\nu} x^\nu ,$$

oder in Matrixschreibweise:

$$(x_\mu) = \begin{pmatrix} x^0 \\ -x^1 \\ -x^2 \\ -x^3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} .$$

Durch Multiplikation des metrischen Tensors $g_{\mu\nu}$ mit dem kontravarianten Vektor x^μ werden nämlich genau die Minuszeichen in der Ortskomponente erzeugt, die für den kovarianten Vektor benötigt werden. Man kann also mit der Metrik $g_{\mu\nu}$ Indizes von kontravarianten Vektoren herunterziehen und aus ihnen kovariante Vektoren machen. Funktioniert dies auch umgekehrt? Die Antwort ist: „Ja!“. Man benötigt hierzu den metrischen Tensor, bei dem beide Indizes oben stehen, also $g^{\mu\nu}$. (Bis jetzt haben wir nämlich nur mit $g_{\mu\nu}$ gearbeitet.) Wie bekommt man diesen Tensor? Durch eine einfache, aber geniale Überlegung! Wendet man auf einen kontravarianten Vektor x^μ den metrischen Tensor $g_{\mu\nu}$ an, um dessen Index herunterzuziehen, so muss bei nochmaliger Anwendung, aber dieses mal von $g^{\mu\nu}$, wieder der ursprüngliche kontravariante Vektor herauskommen:

$$x^\mu = g^{\mu\nu} x_\nu = g^{\mu\nu} g_{\nu\sigma} x^\sigma \overset{!}{=} x^\mu \Rightarrow g^{\mu\nu} g_{\nu\sigma} = \delta^\mu{}_\sigma .$$

Dann gilt nämlich:

$$x^\mu = \delta^\mu{}_\sigma x^\sigma ,$$

oder in Matrixschreibweise:

$$(x^\mu) = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}^2 \cdot \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} .$$

${\delta^\mu}_\sigma$ ist das vierdimensionale Kronecker-Delta oder in Matrixform die vierdimensionale Einheitsmatrix $\mathbb{1}_4$. Also muss gelten:

$$(g^{\mu\nu})(g_{\mu\nu}) = \mathbb{1} ,$$

und dieses ist gerade die Bedingung dafür, dass $(g_{\mu\nu})$ die inverse Matrix von $(g^{\mu\nu})$ ist. Im Falle des Minkowski-Raumes ist die inverse Matrix von $(g^{\mu\nu})$ besonders einfach zu bestimmen, weil es sich um eine Diagonalmatrix handelt. Die Inverse erhält man dann durch Kehrwertbildung ihrer Diagonalelemente, woraus jedoch die ursprüngliche Matrix wieder folgt:

$$(g_{\mu\nu}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = (g^{\mu\nu}) .$$

Damit gibt es also im Minkowski-Raum keinen Unterschied zwischen $g_{\mu\nu}$ und $g^{\mu\nu}$. (In gekrümmten Räumen, so wie sie in der allgemeinen Relativitätstheorie behandelt werden, muss dies jedoch nicht gelten. Da kann sich $g_{\mu\nu}$ von $g^{\mu\nu}$ drastisch unterscheiden!) Mittels $g_{\mu\nu}$ können wir nun also aus einen kontravarianten Vektor einen kovarianten machen, dessen Index also herunterziehen:

$$(x_\mu) = (g_{\mu\nu})(x^\nu) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} x^0 \\ -x^1 \\ -x^2 \\ -x^3 \end{pmatrix} .$$