Ergänzungen zur Vorlesung

Mehr zur Speziellen Relativitätstheorie

von DOMENICO GIULINI

Blatt 2

Vertiefendes zum Additionsgesetz für Geschwindigkeiten

1. Polarzerlegung und Komposition von Lorentztransformationen

Im Folgenden sei stets $\vec{\beta} = \vec{v}/c$, $\beta = \|\vec{\beta}\|$ und $\gamma(\vec{\beta}) = 1/\sqrt{1-\beta^2}$. Unter „Geschwindigkeit" wollen wir immer die Geschwindigkeit in Einheiten der Lichtgeschwindigkeit verstehen, also $\vec{\beta}$; $\vec{v}$ wird weiter nicht vorkommen. Die Menge der Geschwindigkeiten $\mathcal{B}$ ist also durch den dreidimensionalen Einheitsball gegeben:

$$\mathcal{B} := \{\beta \in \mathbb{R}^3 \, : \, \|\vec{\beta}\| = 1\}\,. \tag{1}$$

Die Polarzerlegung erlaubt es, jeder Lorentztransformation L eindeutig einen „Boost" (eine Geschwindigkeitstransformation) und eine räumliche Drehung zuzuordnen. Dabei beschränken wir uns im Folgenden auf eigentliche (Determinante $= 1$), orthochrone (keine Zeitspiegelungen) Lorentztransformationen, d.h. $L \in SO^{\uparrow}(1{,}3)$. Polar zerlegt hat man mit $\vec{\beta} \in \mathcal{B}$ und $\mathbf{D} \in SO(3)$:

$$L(\vec{\beta}, \mathbf{D}) := B(\vec{\beta}) \cdot R(\mathbf{D})\,, \tag{2}$$

wobei Boost und Drehung gegeben sind durch:

$$B(\vec{\beta}) = \begin{pmatrix} \gamma & \gamma\vec{\beta}^{\top} \\ \gamma\vec{\beta} & \mathbf{1}_3 + (\gamma-1)\vec{n}\otimes\vec{n}^{\top} \end{pmatrix} \quad \text{bzw.} \quad R(\mathbf{D}) = \begin{pmatrix} 1 & \vec{0}^{\top} \\ \vec{0}^{\top} & \mathbf{D} \end{pmatrix}\,. \tag{3}$$

Die Eigenwerte der Matrix $B(\vec{\beta})$ sind

$$\lambda_1 = \sqrt{\frac{1+\beta}{1-\beta}}\,, \qquad \lambda_2 = \sqrt{\frac{1-\beta}{1+\beta}}\,, \qquad \lambda_3 = 1\,, \qquad \lambda_4 = 1\,, \tag{4}$$

was die Positivität zeigt. Es parametrisieren $(\vec{\beta}, \mathbf{D}) \in \mathcal{B} \times SO(3)$ also die Gruppe $SO^{\uparrow}(1{,}3)$. Diese Parametrisierung ist eineindeutig. Es gilt offensichtlich die

Äquivarianzeigenschaft

$$\mathsf{R}(\mathbf{D}) \cdot \mathsf{B}(\vec{\beta}) \cdot \mathsf{R}(\mathbf{D}^\top) = \mathsf{R}(\mathbf{D} \cdot \vec{\beta}) \,. \tag{5}$$

Die Polarzerlegung der Komposition zweier Boosts ist gegeben durch

$$\mathsf{B}(\vec{\beta}_1) \cdot \mathsf{B}(\vec{\beta}_2) =: \mathsf{B}(\vec{\beta}_1 \star \vec{\beta}_2) \cdot \mathsf{R}\big(\mathbf{T}[\vec{\beta}_1, \vec{\beta}_2]\big) \,. \tag{6}$$

Hierin bezeichnet $\star$ das *relativistische Additionsgesetz für Geschwindigkeiten* und $\mathbf{T}$ die *Thomasdrehung*. Wir werden beide in den nächsten Abschnitten genauer beschreiben.

Als erstes ziehen wir zwei wichtige Konsequenzen aus (6): Da die linke Seite invariant unter Transposition und gleichzeitigem Austausch $\vec{\beta}_1 \leftrightarrow \vec{\beta}_2$ ist, gilt dies auch für die rechte Seite. Mit Hilfe von (5) folgt daraus (wir schreiben $\mathbf{T}^\top[\vec{\beta}_1, \vec{\beta}_2]$ für $(\mathbf{T}[\vec{\beta}_1, \vec{\beta}_2])^\top$)

$$\begin{aligned} \mathsf{B}(\vec{\beta}_1 \star \vec{\beta}_2) \cdot \mathsf{R}(\mathbf{T}[\vec{\beta}_1, \vec{\beta}_2]) &= \mathsf{R}(\mathbf{T}^\top[\vec{\beta}_2, \vec{\beta}_1]) \cdot \mathsf{B}(\vec{\beta}_2 \star \vec{\beta}_1) \\ &= \mathsf{B}\big(\mathbf{T}^\top[\vec{\beta}_2, \vec{\beta}_1] \cdot \vec{\beta}_2 \star \vec{\beta}_1\big) \cdot \mathsf{R}(\mathbf{T}^\top[\vec{\beta}_2, \vec{\beta}_1]) \,. \end{aligned} \tag{7}$$

Linke und rechte Seite dieser Gleichung sind in polar zerlegter Form. Deren Eindeutigkeit liefert sofort

$$\mathbf{T}^\top[\vec{\beta}_1, \vec{\beta}_2] = \mathbf{T}^{-1}[\vec{\beta}_1, \vec{\beta}_2] = \mathbf{T}[\vec{\beta}_2, \vec{\beta}_1] \tag{8}$$

und

$$\vec{\beta}_2 \star \vec{\beta}_1 = \mathbf{T}[\vec{\beta}_2, \vec{\beta}_1] \cdot (\vec{\beta}_1 \star \vec{\beta}_1) \,. \tag{9}$$

Als zweites folgern wir aus (2) mit Hilfe von (5) und (6) das Gesetz für die Komposition zweier allgemeiner Lorentztransformationen:

$$\begin{aligned} \mathsf{L}(\vec{\beta}_1, \mathbf{D}_1) \cdot \mathsf{L}(\vec{\beta}_2, \mathbf{D}_2) &= \mathsf{B}(\vec{\beta}_1) \cdot \mathsf{R}(\mathbf{D}_1) \cdot \mathsf{B}(\vec{\beta}_2) \cdot \mathsf{R}(\mathbf{D}_2) \\ &= \mathsf{B}(\vec{\beta}_1) \cdot \mathsf{B}(\mathbf{D}_1 \cdot \vec{\beta}_2) \cdot \mathsf{R}(\mathbf{D}_1 \cdot \mathbf{D}_2) \\ &= \mathsf{L}(\vec{\beta}_1 \star \mathbf{D}_1 \cdot \vec{\beta}_2\,,\, \mathbf{T}[\vec{\beta}_1\,,\, \mathbf{D}_1 \cdot \vec{\beta}_2] \cdot \mathbf{D}_1 \cdot \mathbf{D}_2) \,. \end{aligned} \tag{10}$$

Dies definiert eine Gruppenstruktur auf dem Parameterraum $\mathcal{B} \times \mathsf{SO}(3)$:

$$(\vec{\beta}_1, \mathbf{D}_1)(\vec{\beta}_2, \mathbf{D}_2) = (\vec{\beta}_1 \star \mathbf{D}_1 \cdot \vec{\beta}_2\,,\, \mathbf{T}[\vec{\beta}_1\,,\, \mathbf{D}_1 \cdot \vec{\beta}_2] \cdot \mathbf{D}_1 \cdot \mathbf{D}_2) \,. \tag{11}$$

Daraus folgt der Ausdruck für das Inverse, wenn man benutzt, dass $\mathbf{T}[\vec{\beta}_1, \vec{\beta}_2] = \mathbf{1}_3$ wenn $\{\vec{\beta}_1, \vec{\beta}_2\}$ linear abhängig ist (siehe (19) unten):

$$(\vec{\beta}, \mathbf{D})^{-1} = (-\mathbf{D}^{-1} \cdot \vec{\beta}\,,\, \mathbf{D}^{-1})\,. \tag{12}$$

Das Gesetz (11) muss als Verallgemeinerung der semi-direkten Produktstruktur verstanden werden, die in der Galileigruppe gegeben ist. Diese ist aber unverträglich mit einem beschränkten Bereich $\mathcal{B} \subset \mathbb{R}^3$ von Geschwindigkeiten. In der (homogenen) Galileigruppe bilden die reinen Boosts eine invariante Untergruppe, in der Lorentzgruppe nach (5) zwar eine invariante Menge, aber keine Untergruppe mehr. Die Lorentzgruppe ist tatsächlich *einfach*, d.h. besitzt als invariante Untergruppen nur die trivialen (d.h. die Identität und die ganze Gruppe).

2. Das Additionsgesetz

Die „relativistische Geschwindigkeitsaddition" ist eine Abbildung $\mathcal{B} \times \mathcal{B} \to \mathbb{R}^3$, $(\vec{\beta}_1, \vec{\beta}_2) \mapsto \vec{\beta}_1 \star \vec{\beta}_2$, die wie folgt definiert ist:

$$\vec{\beta}_1 \star \vec{\beta}_2 := \frac{\vec{\beta}_1 + \vec{\beta}_{2\parallel} + \gamma^{-1}(\beta_1)\vec{\beta}_{2\perp}}{1 + \vec{\beta}_1 \cdot \vec{\beta}_2}\,. \tag{13}$$

Hier bezeichnen $\parallel$ und $\perp$ die Projektionen parallel und orthogonal zu $\vec{\beta}_1$, also $\vec{\beta}_{2\parallel} := \vec{n}_1(\vec{n}_1 \cdot \vec{\beta}_2)$ und $\vec{\beta}_{2\perp} := \vec{\beta}_2 - \vec{\beta}_{2\parallel}$, wobei $\vec{n}_1 := \vec{\beta}_1/\beta_1$. Für diese Abbildung gilt aber

$$\gamma(\vec{\beta}_1 \star \vec{\beta}_2) = \gamma(\vec{\beta}_1)\gamma(\vec{\beta}_2)\bigl(1 + \vec{\beta}_1 \cdot \vec{\beta}_2\bigr)\,, \tag{14}$$

woraus sofort hervorgeht, dass ihr Bild in $\mathcal{B} \subset \mathbb{R}^3$ liegt, es sich also um eine Abbildung $\mathcal{B} \times \mathcal{B} \to \mathcal{B}$ handelt. Somit macht $\star$ die Menge $\mathcal{B}$ zu einem *Grupoid*. Dieses besitzt mit $\vec{0}$ ein beidseitiges Identitätselement, denn aus (13) folgt sofort

$$\vec{0} \star \vec{\beta} = \vec{\beta} \star \vec{0} = \vec{\beta}\,. \tag{15}$$

Ebenso direkt folgt, dass jedes $\vec{\beta} \in \mathcal{B}$ mit $-\vec{\beta}$ ein beidseitiges Inverses besitzt:

$$(-\vec{\beta}) \star \vec{\beta} = \vec{\beta} \star (-\vec{\beta}) = \vec{0}\,. \tag{16}$$

Damit fehlt zur Gruppeneigenschaft nur die Assoziativität, die auch tatsächlich nicht gilt (siehe unten).

Mit Hilfe der Assoziativität zeigt man normalerweise die eindeutige Auflösbarkeit von Gleichungen der Form $a = bc$ nach entweder b oder c, denn es folgt z.B. $b^{-1}a = b^{-1}(bc) = (b^{-1}b)c = c$. Ein Wegfallen der Assoziativität lässt also befürchten, dass diese Auflösbarkeit ebenfalls verloren geht. Der Witz im vorliegenden Fall ist aber, dass sie trotzdem gilt. Man nennt einen Grupoid eine *Quasigruppe*, wenn es zu jedem Paar von Elementen a, b ein eindeutig bestimmtes Elementenpaar x, y gibt, so dass $xa = b$ und $ay = b$. (Nicht zu verwechseln mit „Halbgruppe", was einen assoziatives Grupoid bezeichnet.) Ein *Loop* ist eine Quasigruppe mit Identität, woraus sofort die Existenz von Inversen folgt. Die Identität und die inversen Elemente sind (per Definition des Loops) eindeutig. Wie wir unten beweisen werden, macht das Additionsgesetz (13) die Menge (1) zu einem Loop.

Aus (13) folgt auch sofort die Äquivarianz unter Drehungen:

$$(\mathbf{D}\cdot\vec{\beta}_1)\star(\mathbf{D}\cdot\vec{\beta}_2) = \mathbf{D}\cdot(\vec{\beta}_1\star\vec{\beta}_2)\,,\quad \forall\mathbf{D}\in\mathrm{SO}(3)\,. \tag{17}$$

3. Die Thomasdrehung

Die Thomasdrehung ist eine Abbildung $\mathbf{T} : \mathcal{B}\times\mathcal{B} \to \mathrm{SO}(3)$; den Wert auf einem Paar $(\vec{\beta}_1, \vec{\beta}_2)$ bezeichnen wir mit $\mathbf{T}[\vec{\beta}_1, \vec{\beta}_2]$. Dieser ist gegeben durch eine Drehung in der $\vec{\beta}_1\vec{\beta}_2$–Ebene um einen Winkel θ, der folgender Gleichung genügt:

$$\cos\theta = 1 - \frac{(\gamma_1-1)(\gamma_2-1)}{\gamma_\star+1}\sin^2\varphi\,. \tag{18}$$

Dabei haben wir $\gamma_\star = \gamma(\vec{\beta}_1\star\vec{\beta}_2)$ und $\gamma_i = \gamma(\vec{\beta}_i)$ für $i = 1, 2$ gesetzt und mit φ den Winkel zwischen $\vec{\beta}_1$ und $\vec{\beta}_2$ bezeichnet. Daraus und mit (9) sieht man sofort, dass

$$\vec{\beta}_1\star\vec{\beta}_2 = \vec{\beta}_2\star\vec{\beta}_1 \Leftrightarrow \mathbf{T}[\vec{\beta}_1, \vec{\beta}_2] = \mathbf{1}_3 \Leftrightarrow \{\vec{\beta}_1, \vec{\beta}_2\}\quad \text{ist linear abhängig}\,. \tag{19}$$

Mit Hilfe von (14) lässt sich aus (18) entweder $\gamma_\star$ zu Gunsten von φ, γ_1 und γ_2, oder φ zu Gunsten von $\gamma_\star$, γ_1 und γ_2 eliminieren. Man erhält dann

$$\cos\theta = 1 - \frac{(\gamma_1-1)(\gamma_2-1)\sin^2\varphi}{1+\gamma_1\gamma_2+\sqrt{(\gamma_1^2-1)(\gamma_2^2-1)}\cos\varphi}\,, \tag{20}$$

beziehungsweise den in $\gamma_\star$, γ_1 und γ_2 symmetrischen Ausdruck

$$\cos(\theta/2) = \frac{1+\gamma_\star+\gamma_1+\gamma_2}{\sqrt{2(1+\gamma_\star)(1+\gamma_1)(1+\gamma_2)}}\,. \tag{21}$$

Die Drehrichtung von $\mathbf{T}[\vec{\beta}_1, \vec{\beta}_2]$ ist in der natürlich orientierten $\vec{\beta}_1\vec{\beta}_2$–Ebene im negativen Sinn (d.h. $\vec{\beta}_2$ wird zu $\vec{\beta}_1$ hin gedreht).

4. Nichtkommutative, nichtassoziative Loop-Struktur

In (9) ist bereits qualitativ und quantitativ die Nichtkommutativität von $\star$ ausgedrückt. Wir ziehen weitere Folgerungen.

4.1 Aus (17) und (9) ergibt sich folgende Äquivarianzeigenschaft von $\mathbf{T}$:

$$\mathbf{T}[\mathbf{D}\cdot\vec{\beta}_1\,,\,\mathbf{D}\cdot\vec{\beta}_2] = \mathbf{D}\cdot\mathbf{T}[\vec{\beta}_1, \vec{\beta}_2]\cdot\mathbf{D}^{-1}\,, \quad \forall\mathbf{D}\in\mathsf{SO}(3)\,. \tag{22}$$

4.2 Als nächstes betrachten wir das Produkt dreier Boosts: $\mathsf{B}(\vec{\beta}_1)\cdot\mathsf{B}(\vec{\beta}_2)\cdot\mathsf{B}(\vec{\beta}_1)$. Dieses ist wieder ein reiner Boost, d.h. positiv definit und symmetrisch. Die Symmetrie ist trivial und die Positivität folgt ebenfalls sofort aus der Symmetrie und Positivität der einzelnen $\mathsf{B}(\vec{\beta}_i)$. Nun ist aber

$$\begin{aligned}\mathsf{B}(\vec{\beta}_1)\cdot\mathsf{B}(\vec{\beta}_2)\cdot\mathsf{B}(\vec{\beta}_1) &= \mathsf{B}(\vec{\beta}_1)\cdot\big(\mathsf{B}(\vec{\beta}_2)\cdot\mathsf{B}(\vec{\beta}_1)\big)\\ &= \mathsf{B}(\vec{\beta}_1)\cdot\mathsf{B}(\vec{\beta}_2\star\vec{\beta}_1)\cdot\mathbf{T}[\vec{\beta}_2, \vec{\beta}_1]\\ &= \mathsf{B}\big(\vec{\beta}_1\star(\vec{\beta}_2\star\vec{\beta}_1)\big)\cdot\mathbf{T}[\vec{\beta}_1\,,\,\vec{\beta}_2\star\vec{\beta}_1]\cdot\mathbf{T}[\vec{\beta}_2, \vec{\beta}_1]\,.\end{aligned} \tag{23}$$

Beide Seiten dieser Gleichung sind in polar zerlegter Form. Deren Eindeutigkeit liefert also insbesondere, dass die Drehung auf der rechten Seite trivial ist, was wegen (8) äquivalent ist zu:

$$\mathbf{T}[\vec{\beta}_1, \vec{\beta}_2] = \mathbf{T}[\vec{\beta}_1\,,\,\vec{\beta}_2\star\vec{\beta}_1] = \mathbf{T}[\vec{\beta}_1\star\vec{\beta}_2\,,\,\vec{\beta}_2]\,, \tag{24}$$

wobei die zweite Gleichheit aus der ersten durch Vertauschen von $\vec{\beta}_1$ mit $\vec{\beta}_2$ und Invertieren hervorgeht. Man nennt dies die „Loop Eigenschaft" von $\mathbf{T}$.

4.3 Ein Maß für die Nichtassoziativität erhalten wir wie folgt: Wir betrachten drei Boosts und nutzen die Assoziativität der Gruppenmultiplikation:

$$\mathsf{B}(\vec{\beta}_1)\cdot\big(\mathsf{B}(\vec{\beta}_2)\cdot\mathsf{B}(\vec{\beta}_3)\big) = \big(\mathsf{B}(\vec{\beta}_1)\cdot\mathsf{B}(\vec{\beta}_2)\big)\cdot\mathsf{B}(\vec{\beta}_3)\,. \tag{25}$$

Die linke Seite ist gleich (sukzessives Anwenden von (6))

$$\begin{aligned}&\mathsf{B}(\vec{\beta}_1)\cdot\mathsf{B}(\vec{\beta}_2\star\vec{\beta}_3)\cdot\mathsf{R}(\mathbf{T}[\vec{\beta}_2, \vec{\beta}_3])\\ =\;&\mathsf{B}\big(\vec{\beta}_1\star(\vec{\beta}_2\star\vec{\beta}_3)\big)\cdot\mathsf{R}(\mathbf{T}[\vec{\beta}_1, \vec{\beta}_2\star\vec{\beta}_3]\cdot\mathbf{T}[\vec{\beta}_2, \vec{\beta}_3])\end{aligned} \tag{26}$$

und die rechte, unter Benutzung von (5), gleich

$$\begin{aligned}
& \mathsf{B}(\vec\beta_1 \star \vec\beta_2) \cdot \mathsf{R}(\mathbf{T}[\vec\beta_1, \vec\beta_2]) \cdot \mathsf{B}(\vec\beta_3) \\
= \; & \mathsf{B}(\vec\beta_1 \star \vec\beta_2) \cdot \mathsf{B}(\mathbf{T}[\vec\beta_1, \vec\beta_2] \cdot \vec\beta_3) \cdot \mathsf{R}(\mathbf{T}[\vec\beta_1, \vec\beta_2]) \\
= \; & \mathsf{B}\big((\vec\beta_1 \star \vec\beta_2) \star (\mathbf{T}[\vec\beta_1, \vec\beta_2] \cdot \vec\beta_3)\big) \cdot \mathsf{R}(\mathbf{T}[\vec\beta_1 \star \vec\beta_2\,,\, \mathbf{T}[\vec\beta_1, \vec\beta_2] \cdot \vec\beta_3] \cdot \mathbf{T}[\vec\beta_1, \vec\beta_2])\,.
\end{aligned} \tag{27}$$

Nun sind (26) und (27) wieder in polar zerlegter Form, so dass deren Eindeutigkeit die Gleichheit der entsprechenden Faktoren zur Folge hat. Für den Boost-Anteil folgt:

$$\vec\beta_1 \star (\vec\beta_2 \star \vec\beta_3) \;=\; (\vec\beta_1 \star \vec\beta_2) \star (\mathbf{T}[\vec\beta_1, \vec\beta_2] \cdot \vec\beta_3)\,, \tag{28}$$

$$(\vec\beta_1 \star \vec\beta_2) \star \vec\beta_3 \;=\; \vec\beta_1 \star (\vec\beta_2 \star (\mathbf{T}[\vec\beta_2, \vec\beta_1] \cdot \vec\beta_3))\,, \tag{29}$$

wobei sich (29) aus (28) ergibt, indem man dort $\mathbf{T}[\vec\beta_1, \vec\beta_2] \cdot \vec\beta_3 = \vec\beta'_3$ setzt, die Gleichung auf $\vec\beta_1$, $\vec\beta_2$ und $\vec\beta'_3$ umschreibt und am Ende den Strich an $\vec\beta'_3$ wieder weg lässt. Damit ist die Nichtassoziativität qualitativ und quantitativ beschrieben.

Vergleich der Rotationsanteile von (26) und (27) ergibt

$$\mathbf{T}[\vec\beta_1, \vec\beta_2 \star \vec\beta_3] \cdot \mathbf{T}[\vec\beta_2, \vec\beta_3] = \mathbf{T}[\vec\beta_1 \star \vec\beta_2\,,\, \mathbf{T}[\vec\beta_1, \vec\beta_2] \cdot \vec\beta_3] \cdot \mathbf{T}[\vec\beta_1, \vec\beta_2]\,. \tag{30}$$

Setzt man darin $\vec\beta_1 = -\vec\beta_2$ so wird die rechte Seite zur Identitätsmatrix. Benennt man dann noch um: $\vec\beta_2 \to \vec\beta_1$ und $\vec\beta_3 \to \vec\beta_2$, so folgt:

$$\mathbf{T}[\vec\beta_1, \vec\beta_2] = \mathbf{T}[\vec\beta_1 \star \vec\beta_2, -\vec\beta_1] = \mathbf{T}[-\vec\beta_2, \vec\beta_2 \star \vec\beta_1] \tag{31}$$

wobei die zweite Gleichheit wieder aus der ersten durch Inversenbildung und gleichzeitiges Vertauschen $\vec\beta_1 \leftrightarrow \vec\beta_2$ folgt. Diese ist in gewisser Analogie zur Loop Eigenschaft (24) zu sehen.

4.4 Schließlich zeigen wir noch die Quasigruppeneigenschaft, d.h die eindeutige Auflösbarkeit von

$$\vec\beta_1 \star \vec\beta_2 = \vec\beta_3 \tag{32}$$

nach $\vec\beta_2$ und $\vec\beta_3$.

Satz 1. *Die Gleichung (32) impliziert*

$$\vec\beta_2 \;=\; (-\vec\beta_2) \star \vec\beta_3\,, \tag{33}$$

$$\vec\beta_1 \;=\; \vec\beta_3 \star (-\mathbf{T}[\vec\beta_3, \vec\beta_2] \cdot \vec\beta_2)\,. \tag{34}$$

Beweis. Operiert man auf (32) von links mit $-\vec{\beta}_1\star$ und benutzt (28) so ergibt sich wegen $\mathbf{T}[\vec{\beta}_1, -\vec{\beta}_1] = \mathbf{1}_3$ sofort (33). Eine analoge Anwendung von (29) bringt hinsichtlich der Lösung nach $\vec{\beta}_1$ nichts, da $\vec{\beta}_1$ ausschließlich als Funktion von $\vec{\beta}_2$ und $\vec{\beta}_3$ geschrieben werden soll, die rechte Seite von (29) im Argument von $\mathbf{T}$ aber $\vec{\beta}_1$ (statt $\vec{\beta}_3$) enthält. Wir müssen deshalb zum Beweis von (34) weiter ausholen, indem wir auf die Gruppenstruktur (11-12) des Parameterraum zurückgreifen. Wir betrachten nun die Gleichung

$$(\vec{\beta}_3, \mathbf{D}_3) = (\vec{\beta}_1, \mathbf{D}_1)(\vec{\beta}_2, \mathbf{D}_2)\,, \tag{35}$$

die nach (11) äquivalent ist zu

$$\begin{aligned} \vec{\beta}_3 &= \vec{\beta}_1 \star \mathbf{D}_1 \vec{\beta}_2\,, && (36)\\ \mathbf{D}_3 &= \mathbf{T}[\vec{\beta}_1, \mathbf{D}_1 \cdot \vec{\beta}_2] \cdot \mathbf{D}_1 \cdot \mathbf{D}_2\,. && (37) \end{aligned}$$

Andererseits kann (35) innerhalb der Gruppenstruktur aber eindeutig nach $\vec{\beta}_1$ und $\mathbf{D}_1$ gelöst werden. Mit (12) folgt

$$(\vec{\beta}_1, \mathbf{D}_1) = (\vec{\beta}_3, \mathbf{D}_3)(-\mathbf{D}_2^\top \vec{\beta}_2, \mathbf{D}_2^\top)\,, \tag{38}$$

was nach (11) wieder äquivalent ist zu

$$\begin{aligned} \vec{\beta}_1 &= \vec{\beta}_3 \star (-\mathbf{D}_3 \cdot \mathbf{D}_2^\top \cdot \vec{\beta}_2)\,, && (39)\\ \mathbf{D}_1 &= \mathbf{T}[\vec{\beta}_3, -\mathbf{D}_3 \cdot \mathbf{D}_2^\top \cdot \vec{\beta}_2] \cdot \mathbf{D}_3 \cdot \mathbf{D}_2^\top\,. && (40) \end{aligned}$$

Wegen der Gruppenstruktur sind (36-37) und (39-40) äquivalente Gleichungssysteme. Insbesondere dürfen wir (39-40) als Folge von (36-37) ansehen.

Wir spezialisieren nun auf $\mathbf{D}_1 = \mathbf{1}_3$, im dem (36) gerade zu (32) wird. Aus (37) wird dann

$$\begin{aligned} \mathbf{D}_3 \cdot \mathbf{D}_2^\top &= \mathbf{T}[\vec{\beta}_1, \vec{\beta}_2]\\ &= \mathbf{T}[\vec{\beta}_1 \star \vec{\beta}_2\,,\ \vec{\beta}_2] \quad \text{(wegen (24))}\\ &= \mathbf{T}[\vec{\beta}_3, \vec{\beta}_2] \quad \text{(mit Benutzung von (37))} \end{aligned} \tag{41}$$

Einsetzen von (41) in (39) ergibt dann (34). Somit ist (34) als Folge von (36-37) nachgewiesen. □

3. Zur Eindeutigkeit von Gruppenstrukturen auf Intervallen

Für parallele Geschwindigkeiten vereinfacht sich das speziell-relativistische Additionsgesetz (13) zu

$$\beta_1 \star \beta_2 = \frac{\beta_1 + \beta_2}{1 + \beta_1\beta_2}\,, \tag{42}$$

wobei $\beta_{1,2} \in (-1, 1)$. Setzt man $\beta_{1,2} = \tanh(\rho_{1,2})$ so entspricht (42) gerade der üblichen Addition $(+)$ auf der reellen Achse für den Parameter ρ, den man die „Rapidität“ nennt. Mit anderen Worten, ist $f : (-1, 1) \to \mathbb{R}$, $x \mapsto f(x) := \tanh^{-1}(x)$, dann

$$\beta_1 \star \beta_2 = f^{-1}\big(f(\beta_1) + f(\beta_2)\big)\,. \tag{43}$$

Die Bijektion $f : (-1, 1) \to \mathbb{R}$ vermittelt also einen Gruppenisomorphismus von $\{(-1, 1), \star\}$ nach $\{\mathbb{R}, +\}$. Das speziell-relativistische Gesetz (42) ist also nichts anderes als das übliche Additionsgesetz auf $\mathbb{R}$ in gestauchten Koordinaten.

Ganz allgemein kann man jedes offene Intervall $I \subset \mathbb{R}$ zu einer kommutativen Gruppe machen, indem man die Multiplikation $\star$ durch (43) definiert, wobei $f : I \to \mathbb{R}$ eine Bijektion ist. Die Frage ist nun, ob dadurch *alle* Gruppenstrukturen auf I erfasst werden. Diese Fragen wollen wir hier für den Fall positiv beantworten, in dem die Gruppenmultiplikation C^1 (also stetig differenzierbar) ist.

Sei nun $I \subset \mathbb{R}$ ein offenes Intervall und $F : I \times I \to I$ eine nicht als abelsch vorausgesetzte Gruppenmultiplikation. Das bedingt folgendes:

G1 Existenz eines neutralen Elements: Es existiert ein Element $e \in I$, so dass $F(e, x) = F(x, e) = x, \quad \forall x \in I$. Beachte: Die Existenz impliziert bereits die Eindeutigkeit, denn gilt $F(e, x) = F(x, e) = F(e', x) = F(x, e') = x$ für alle x, so setzt man zuerst $x = e$ und dann $x = e'$ und sieht sofort $e = e'$.

G2 Assoziativität: Für alle $x, y, z \in I$ gilt

$$F\big(x, F(y, z)\big) = F\big(F(x, y), z\big)\,. \tag{44}$$

G3 Existenz inverser Elemente: Zu jedem $x \in I$ existiert ein $x^{-1} \in I$, so dass $F(x, x^{-1}) = F(x^{-1}, x) = e$. Wieder impliziert die Existenz, dass jedem x ein eindeutiges inverses Element zukommt. Gilt nämlich $e = F(x, h) = F(h, x) = F(x, k) = F(k, x)$, so ergibt Linksmultiplikation mit k unter Benutzung der Assoziativität: $k = F(k, e) = F(k, F(x, h)) = F(F(k, x), h) = F(e, h) = h$. Somit existiert eine Inversenabbildung $i : I \to I$, $x \mapsto x^{-1}$.

> Diese ist involutiv (selbstinvers, $i \circ i = \mathrm{id}_I$) und damit notwendig bijektiv. Die Involutivität folgt daraus, dass x invers zu $i(x)$ ist, also wegen der Eindeutigkeit $x = i(i(x))$ gilt. (Zur Injektivität: Aus $i(x) = i(y)$ folgt $x = y$ durch Abbilden mit i. Zur Surjektivität: x ist Bild von $i(x)$.)

Die Gruppe $\{I, F, i\}$ ist eine topologische Gruppe falls die Abbildungen F und i stetig sind. Es wäre interessant zu wissen, ob folgende Vermutung richtig ist:

Hoffnung: Ist $I \subset \mathbb{R}$ ein Intervall und $\{I, F, i\}$ topologische Gruppe (bezüglich der induzierten Topologie von I). Dann existiert ein Homöomorphismus (Bijektion die in beiden Richtungen stetig ist) $f : I \to \mathbb{R}$, so dass $f(e) = 0$, $f \circ i = -f$ und $F(x, y) = f^{-1}\big(f(x) + f(y)\big)$.

Beweisen wollen wir folgende schwächere Version:

Theorem: Ist $I \subset \mathbb{R}$ ein Intervall und $\{I, F, i\}$ topologische Gruppe (bezüglich der induzierten Topologie von I) in der die Abbildungen F und i C^1 (einmal stetig differenzierbar) sind. Dann existiert ein C^1 Diffeomorphismus $f : I \to \mathbb{R}$, so dass $f(e) = 0$, $f \circ i = -f$ und $F(x, y) = f^{-1}\big(f(x) + f(y)\big)$.

Beweis: Existiert f, so gilt $f(F(x, y)) = f(x) + f(y)$. Partielle Differentiation nach x ergibt $f'(F(x, y))F_1(x, y) = f'(x)$. (Hier und in Folgenden bezeichnet F_1 die partielle Ableitung von F nach dem ersten Argument und f' wie gewöhnlich die Ableitung von f.) Ausgewertet bei $x = e$ ergibt dies $f'(y)F_1(e, y) = f'(e)$, oder

$$f(y) = \int_e^y \frac{f'(e)}{F_1(e, t)}\, dt\,. \tag{45}$$

Dies erfüllt $f(e) = 0$ und ist tatsächlich Konsistent, denn Differentiation nach y und Auswertung bei $y = e$ ergibt $f'(e) = f(e)/F_1(e, e)f$. Nun folgt aus $F(x, e) = x$ durch Differentiation nach x aber $F_1(x, e) := 1$, insbesondere also

$$F_1(e, e) := 1\,. \tag{46}$$

Somit kann f durch eine explizite Formel mit Hilfe der Ableitung von F und einer frei wählbaren Konstante, $f'(e)$, ausgedrückt werden. Die Idee ist nun, diese Formel selbst zur Konstruktion von f zu verwenden.

Dazu ist zunächst zu zeigen, dass der Integrand in (45) stetig ist (denn f soll C^1 sein). In der Tat ist $F_1(e, t) > 0\,\forall t \in I$. Dies folgt aus der Differentiation der Assoziativitätsbedingung (44) nach x

$$F_1\big(x, F(y, z)\big) = F_1\big(F(x, y), z\big)F_1(x, y) \tag{47}$$

und anschließender Auswertung bei $x = e$:

$$F_1\big(e, F(y,z)\big) = F_1(y,z)F_1(e,y)\,. \tag{48}$$

Wäre $F_1(e,y) = 0$ für ein $y \in I$, dann auch $F_1\big(e, F(y,z)\big) = 0$, und zwar für alle $z \in I$, insbesondere für $z = i(y)$. Also wäre insbesondere $F_1(e,e) = 0$, in Widerspruch zu (46). Also ist $F_1(e,y) \neq 0$ für alle y, und wegen $F_1(e,e) = 1$ und der Stetigkeit von F_1 ist $F_1(e,y)$ überall positiv. Wir setzen die in (45) unbestimmte Konstante $f'(e) = 1$ und erhalten so als Definitionsgleichung für f

$$f(x) := \int_e^x \frac{dt}{F_1(e,t)}\,dt\,. \tag{49}$$

Wir müssen nun zeigen, dass das so definierte f die gewünschten Eigenschaften hat. Zunächst folgt aus (49)

$$f'(x) := 1/F_1(e,x) > 0 \tag{50}$$

und auch $f'(e) = 1$, wie auch oben bereits gesehen. Um die Eigenschaft der Homomorphie zu zeigen, betrachten wir die C^1 Funktion $T : I \times I \to \mathbb{R}$, $T(x,y) := f(F(x,y)) - f(x) - f(y)$. Differentiation nach x ergibt

$$\begin{aligned} T_1(x,y) &= f'(F(x,y))F_1(x,y) - f'(x) \\ &\overset{(50)}{=} \frac{F_1(x,y)}{F_1(e,F(x,y))} - \frac{1}{F_1(e,x)} \\ &= \frac{F_1(x,y)F_1(e,x) - F_1(e,F(x,y))}{F_1(e,x)F_1(e,F(x,y))} = 0 \end{aligned} \tag{51}$$

wobei das Verschwinden (letzte Gleichung) aus der differenzierten Assoziativitätsbedingung (48) folgt, indem man dort $y \mapsto x$ und $z \mapsto y$ ersetzt. Also gilt $T(x,y) = g(y)$ und $g(y) = T(e,y) = f(y) - f(0) - f(y) = 0$, so dass $T \equiv 0$. Somit ist f ein Gruppenhomomorphismus. Insbesondere gilt $f \circ i = -f$

Aus (50) folgt die Injektivität der Abbildung f. Wegen der Stetigkeit ist f eine offene Abbildung (Satz von der Gebietstreue), bildet also offene Mengen in offene Mengen ab und ist somit ein Homöomorphismus auf sein Bild $f(I) \subseteq \mathbb{R}$. Da das Bild einer offenen zusammenhängenden Menge unter einer offenen stetigen Abbildung selbst offen und zusammenhängend ist, muss $f(I)$ entweder ein offenes Intervall, eine offene Halbgerade oder ganz $\mathbb{R}$ sein. Nur letzteres kommt aber in

Frage. Denn sei $x \neq e$, so dass auch $f(x) \neq 0 \in \mathbb{R}$, dann ist (wir schreiben jetzt $x \star y$ statt $F(x, y)$)

$$\begin{aligned} f(\underbrace{x \star \cdots \star x}_{n \text{ Faktoren}}) &= n\, f(x) \\ f(\underbrace{i(x) \star \cdots \star i(x)}_{n \text{ Faktoren}}) &= -n\, f(x)\,, \end{aligned} \tag{52}$$

so dass das Bild von f weder auf der positiven noch auf der negativen reellen Achse beschränkt sein kann. Also ist f eine stetig differenzierbare Isomorphie der Gruppen (I, F, i) und $(\mathbb{R}, +, -)$.