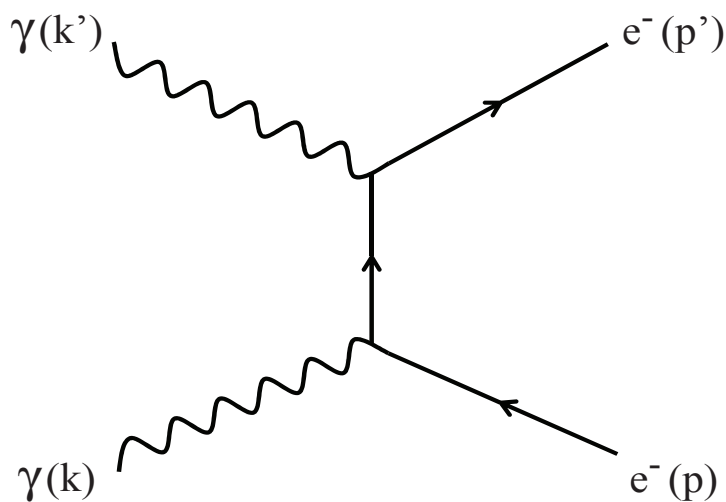


Notizen zur Vorlesung Mathematische Methoden der Teilchenphysik

Andreas Aste
Departement Physik der Universität Basel

Herbstsemester 2024

$\mathcal{L}_+^\uparrow = SO^+(1, 3)$	$F'^{\mu\nu}(x') = \Lambda^\mu_\alpha \Lambda^\nu_\beta F^{\alpha\beta}(x)$
$\partial_\mu F^{\mu\nu} = j^\nu$	$\square\varphi(x) + m^2\varphi(x) = 0$
$\psi'_L(x') = \epsilon A^* \epsilon^{-1} \psi_L(\Lambda^{-1}x')$	$\square A^\nu - \partial^\nu \partial_\mu A^\mu + m^2 A^\nu = 0$
$\underline{\partial} = \sigma_\mu \partial^\mu = \begin{pmatrix} \partial^0 + \partial^3 & \partial^1 - i\partial^2 \\ \partial^1 + i\partial^2 & \partial^0 - \partial^3 \end{pmatrix}$	$A^\mu(x) = \int_0^1 d\lambda \lambda F^{\mu\nu}(\lambda x) x_\nu$
$i\sigma_\mu \partial^\mu \psi_L(x) - \eta_L m_L (i\sigma_2) \psi_L(x)^* = 0$	
$(i\gamma_\mu \partial^\mu - m)\Psi(x) = 0$	$T^{\mu\nu}(x) = \varphi(x)g^{\mu\nu} + F^{\mu\nu}(x) + H^{\mu\nu}(x)$



Algebraische Strukturen und vollständige Vektorräume

Definition 1 Sei G eine Menge mit einer zweistelligen inneren Verknüpfung $\circ : G \times G \rightarrow G$. (G, \circ) heisst Gruppe, falls folgende Eigenschaften gelten:

- (M) $a, b \in G \Rightarrow \circ(a, b) =: a \circ b \in G$
- (A) $a \circ (b \circ c) = (a \circ b) \circ c \quad \forall a, b, c \in G$ (Assoziativität)
- (N) $\exists_1 n \in G, n \circ a = a \circ n = a \quad \forall a \in G$ (Existenz des eindeutigen neutralen Elements)
- (I) $a \in G \Rightarrow \exists_1 a^{-1} \in G, a^{-1} \circ a = a \circ a^{-1} = n$ (Inverses Element)

Gilt für G die Eigenschaft (M), so heisst (G, \circ) Magma. Beispiele: $(\mathbb{Z}, -)$, $(\mathbb{N}, a \circ b = a^b)$.

Gelten in G die Eigenschaften (M) und (A), so ist (G, \circ) eine Halbgruppe. Beispiel: $(\mathbb{N}, +)$.

Ist (G, \circ) eine Halbgruppe mit neutralem Element gemäss (N), nennt man (G, \circ) ein Monoid. Beispiel: (\mathbb{N}, \cdot) .

Eine Gruppe (G, \circ) heisst abelsch, falls $a \circ b = b \circ a$ für alle $a, b \in G$ gilt.

Definition 2 Sei K eine Menge mit zwei zweistelligen Verknüpfungen $\oplus : K \times K \rightarrow K$ und $\odot : K \times K \rightarrow K$. (K, \oplus, \odot) heisst Körper, falls gilt:

- (K1) (K, \oplus) ist eine abelsche Gruppe mit neutralem Element 0.
- (K2) $(K \setminus \{0\}, \odot)$ ist eine abelsche Gruppe mit neutralem Element 1.
- (K3) $\forall a, b, c \in K$ gilt: $a \odot (b \oplus c) = a \odot b \oplus a \odot c$, $(a \oplus b) \odot c = a \odot c \oplus b \odot c$ (Distributivgesetz).

Notationsregel: Die zweite Verknüpfung \odot wird vor der ersten Verknüpfung \oplus ausgeführt ('Mal vor Plus').

Beispiele: Die rationalen Zahlen $(\mathbb{Q}, +, \cdot)$, die reellen Zahlen $(\mathbb{R}, +, \cdot)$, die komplexen Zahlen $(\mathbb{C}, +, \cdot)$, meromorphe Funktionen auf einem Gebiet in der komplexen Ebene mit der üblichen Addition und Multiplikation.

Definition 3 Ein Vektorraum V_K über einem Körper (K, \oplus, \odot) ist eine additive abelsche Gruppe $(V_K, +)$, auf der zusätzlich eine Multiplikation $\star : K \times V_K \rightarrow V_K$ mit einem Skalar aus K erklärt ist.

Für alle Vektoren $u, v \in V_K$ und Skalare $\alpha, \beta \in K$ gelten die folgenden Definitions-Axiome (V1-4):

- (V1) $\alpha \star (\beta \star v) = (\alpha \odot \beta) \star v$
- (V2) $\alpha \star (u + v) = \alpha \star u + \alpha \star v$
- (V3) $(\alpha \oplus \beta) \star v = \alpha \star v + \beta \star v$
- (V4) $1 \star v = v$, wobei 1 das Einselement des Körpers K darstellt.

In der Quantenmechanik spielen reelle und komplexe Vektorräume ($\mathbb{K} = \mathbb{R}$ oder \mathbb{C}) eine wichtige Rolle.

Beispiele: $\mathbb{R}_{\mathbb{R}}^n, \mathbb{C}_{\mathbb{C}}^n, \mathbb{C}_{\mathbb{R}}^n, C^k(I) = \{f : I \rightarrow \mathbb{R} \text{ oder } \mathbb{C} \mid f \text{ k-mal stetig diff.-bar auf dem Intervall } I \subseteq \mathbb{R}\}$.

Bemerkung: Wird ein Vektorraum mit einer zusätzlichen vektorwertigen zweiten Verknüpfung ('Vektormultiplikation') zwischen je zwei Vektoren versehen, so spricht man oft von einer Algebra. Mengen mit Verknüpfungsstrukturen nennt man ganz allgemein algebraische Strukturen.

Beispiele: Dreidimensionaler euklidischer Vektorraum mit Vektoraddition und Vektorprodukt ('Kreuzprodukt'), Quaternionenalgebra $\mathbb{H}_{\mathbb{R}}$, Oktonionen $\mathbb{O}_{\mathbb{R}}$, komplexe $n \times n$ -Matrizen $Mat(n, \mathbb{C})$ mit entspr. Verknüpfungen.

Bemerkung: Ein Skalar ist also ein Element aus dem Grundkörper K eines Vektorraumes, ein Vektor ein Element aus V_K . Der Begriff Skalar wird in der Physik auch in einem anderen Sinne für eine vom Inertialsystem eines Beobachters unabhängige Grösse verwendet (z.B. die Ruhemasse eines Teilchens).

Definition 4 n Vektoren $\{v_1, \dots, v_n\} \subset V_K$ heissen linear unabhängig, falls für $\{\lambda_1, \dots, \lambda_n\} \subset K$

$$\lambda_1 \star v_1 + \dots + \lambda_n \star v_n = \mathbf{0} \quad \Leftrightarrow \quad \lambda_1 = \dots = \lambda_n = 0$$

gilt. Dabei ist der Nullvektor $\mathbf{0}$ das neutrale Element in $(V_K, +)$.

Definition 5 n linear unabhängige Vektoren $\{v_1, \dots, v_n\} \subset V_K$ sind eine Basis von V_K , falls sich jeder Vektor $v \in V_K$ als folgende Linearkombination darstellen lässt:

$$v = \mu_1 \star v_1 + \dots + \mu_n \star v_n, \quad \{\mu_1, \dots, \mu_n\} \subset K.$$

$n = \dim(V_K)$ ist die Dimension von V_K . Besteht der Vektorraum nur aus dem Element $\mathbf{0}$, so ist $\dim(\mathbf{0}) = 0$.

Beispiele: $\dim(\mathbb{R}_\mathbb{R}^n) = n$, $\dim(\mathbb{C}_\mathbb{C}^n) = n$, $\dim(\mathbb{C}_\mathbb{R}^n) = 2n$, $\dim(C^0([0, 1])) = \infty$.

Definition 6 Eine Norm ist eine Abbildung $\|\cdot\| : V_K \rightarrow \mathbb{R}_0^+$ von einem Vektorraum über dem Körper \mathbb{K} der reellen oder komplexen Zahlen in die nicht-negativen reellen Zahlen \mathbb{R}_0^+ , welche für alle Vektoren $x, y \in V_K$ und alle Skalare $\lambda \in \mathbb{K}$ folgende Definitions-Axiome erfüllt:

(N1) $\|x\| = 0 \Leftrightarrow x = 0$ (Definitheit)

(N2) $\|\lambda x\| = |\lambda| \cdot \|x\|$, $|\lambda| = \sqrt{\bar{\lambda} \cdot \lambda} \geq 0$ (Homogenität)

(N3) $\|x + y\| \leq \|x\| + \|y\|$ (Δ -Ungleichung oder Subadditivität)

Wie oben werden in der Folge die Symbole $+$ und \cdot sowohl für skalare wie auch vektorielle Addition und Multiplikation verwendet. $\bar{\lambda} = \lambda^* = \lambda_1 - i \cdot \lambda_2$ bezeichnet den komplex konjugierten Wert von $\lambda = \lambda_1 + i\lambda_2$, $\lambda_{1,2} \in \mathbb{R}$. Ein Vektorraum mit Norm heisst *normierter Vektorraum*.

Definition 7 Ein Banach-Raum \mathcal{B} ist ein vollständiger normierter Vektorraum, d.h. in einem solchen Raum besitzt jede Cauchy-Folge $\{a_n\}_{n \in \mathbb{N}} \subset \mathcal{B}$ (mit der Eigenschaft dass $\forall \epsilon > 0 \exists N(\epsilon)$ so, dass $\|a_n - a_m\| < \epsilon$ für alle $n, m \geq N(\epsilon)$) einen (eindeutigen) Grenzwert $a \in \mathcal{B}$; es gilt dann also für dieses a dass $\forall \epsilon > 0 \exists N(\epsilon)$ so, dass $\|a_m - a\| < \epsilon$ für alle $m \geq N(\epsilon)$. Cauchy-Folgen werden auch als Fundamentalfolgen, konzentrierte Folgen oder in sich konvergente Folgen bezeichnet.

Die erste systematische Untersuchung von Banachräumen findet sich in der Dissertation von Stefan Banach (1922).

Definition 8 Es sei V_K ein Vektorraum über den reellen oder komplexen Zahlen. Ein Skalarprodukt oder inneres Produkt ist eine positiv definite hermitesche Sesquilinearform ('eineinhalb-linear'), d.h. eine Abbildung $\langle \cdot, \cdot \rangle : V_K \times V_K \rightarrow \mathbb{K}$, welche für alle $x, y, z \in V_K$ und $\lambda \in \mathbb{K} = \mathbb{R}$ oder \mathbb{C} folgende Forderungen erfüllt:

(I1) $\langle x, x \rangle \geq 0$ (positive Definitheit)

(I2) $\langle x, x \rangle = 0 \Leftrightarrow x = 0$ (Definitheit)

(I3) $\langle x, y \rangle = \overline{\langle y, x \rangle}$ (Symmetrie, Hermitizität)

(I4) $\langle x, \lambda y + z \rangle = \lambda \langle x, y \rangle + \langle x, z \rangle$ (Linearität)

Aus diesen Axiomen folgt sofort die Antilinearität im ersten Argument $\langle \lambda x, y \rangle = \bar{\lambda} \langle x, y \rangle$.

Mathematiker fordern in (I4) die Linearität oft im linken statt im rechten Argument.

Definition 9 Ein Vektorraum mit innerem Produkt $(V_K, \langle \cdot, \cdot \rangle)$ heisst Prä-Hilbertraum. Im reellen Fall spricht man auch von einem euklidischen Raum $V_\mathbb{R}$, im komplexen Fall von einem unitären Raum $V_\mathbb{C}$. Diese Räume sind mit der induzierten Norm $\|x\| = \sqrt{\langle x, x \rangle} \geq 0$ auch normierte Vektorräume.

Definition 10 Ein Hilbertraum ist ein bezüglich der induzierten Norm vollständiger euklidischer oder unitärer Vektorraum mit Skalarprodukt. Oft wird nur der komplexe (unitäre) Fall betrachtet. Ein unvollständiger Prähilbertraum lässt sich immer eindeutig (bis auf Isomorphie) zu einem Hilbertraum erweitern, indem fehlende Grenzwerte in Form von Äquivalenzklassen entsprechender Cauchy-Folgen als abstrakte Elemente zum Prähilbertraum hinzugefügt werden.

Quanten, Skalen, Konstanten und Fluktuationen

“Es ist falsch zu glauben, daß es Aufgabe der Physik sei, herauszufinden, wie die Natur ist. Die Physik beschäftigt sich mit dem, was wir über die Natur sagen können.”

Niels Bohr (1885-1962).

Unschärfe

Aus den Postulaten der Quantenmechanik leitet sich ab, dass bei der simultanen Messung zweier durch selbst-adjungierte lineare Operatoren \hat{A} und \hat{B} repräsentierten Observablen an einem durch einen Zustandsvektor ψ in einem das quantenmechanische System repräsentierenden separablen unitären Hilbertraum \mathcal{H} für die Standardabweichungen $\Delta\hat{A}$ und $\Delta\hat{B}$ die *Unschärferelation* (UR)

$$\Delta\hat{A}\Delta\hat{B} \geq \frac{1}{2}|\langle\psi|[\hat{A}, \hat{B}]\psi\rangle| \quad (1)$$

gilt. Die Standardabweichung $\Delta\hat{A} > 0$ beschreibt eine mittlere Abweichung der Messwerte vom Erwartungswert $\langle\hat{A}\rangle = \langle\psi|\hat{A}\psi\rangle$ und ist durch die Varianz $\Delta\hat{A}^2 = \langle\psi|(\hat{A} - \langle\hat{A}\rangle)^2\psi\rangle$ gegeben. In ihrer ursprünglichen und prominentesten Form wurde die Unschärferelation für die Unbestimmtheit des Aufenthaltsortes und des Impulses eines Teilchens (in einer räumlichen Dimension) formuliert¹:

$$\Delta\hat{q}\Delta\hat{p} \geq \hbar/2. \quad (2)$$

Offensichtlich folgt diese Unbestimmtheit aus der in der Heisenberg-Algebra geforderten Kommutatorbeziehung $[\hat{q}, \hat{p}] = i\hbar$ für den Ortsoperator \hat{q} und den Impulsoperator \hat{p} . Die Relation $[\hat{q}, \hat{p}] = i\hbar$ kann übrigens nicht in endlichdimensionalen unitären Räumen erfüllt werden, da im n -dimensionalen Raum für die Spur (englisch: *trace*) von $[\hat{q}, \hat{p}]$ sofort der Widerspruch $\text{tr}[q, p] = \text{tr}(qp) - \text{tr}(pq) = 0 \neq i\hbar \text{tr id} = i\hbar n$ folgen würde.

Oft wird in der Literatur eine analoge Energie-Zeit-Unschärferelation der Form $\Delta E\Delta t \gtrsim \hbar$ konstruiert, welche für gewisse Betrachtungen durchaus praktisch, aber mit Vorsicht zu geniessen ist. So wird auch argumentiert, dass im physikalischen Raum für kurze Zeiten Δt der Energieerhaltungssatz um Energiebeträge der Größenordnung $\Delta E \simeq \hbar/\Delta t$ verletzt werden könne, was zur spontanen Bildung virtueller Teilchen als 'Vakuumfluktuationen' führe, welche zudem innerhalb der Zeit Δt über eine durch die Vakuum-Lichtgeschwindigkeit $c = 299792458\text{m/s}$ assoziierte Strecke der Größenordnung $\Delta x \simeq c\Delta t$ propagieren können. Dieses Bild ist zu relativieren; tatsächlich gilt in der Quantenfeldtheorie, welche Teilchenerzeugungs- und Teilchenvernichtungsvorgänge beschreibt, immer *strikt* der Energieerhaltungssatz². Es ist aber so, dass bei der Beobachtung eines kleinen Raumvolumens oder einem kurz andauernden Messvorgang Effekte eine Rolle spielen können, die sich als 'Teilchen-Fluktuationen' interpretieren lassen.

Der Grund für solche Effekte liegt in der Tatsache begründet, dass gewisse Messgrößen in der Quantenmechanik nicht mit dem Hamiltonoperator des untersuchten Systems kommutieren; entsprechend treten Energie-Unschärfen durch den Beobachtungsprozess auf. So bewegt sich das Elektron im Grundzustand eines ruhenden Wasserstoffatoms nicht; der Zustand des Elektrons wird durch seine Wellenfunktion beschrieben, welche die gesamte uns zugängliche Information über den Elektronenzustand repräsentiert. Will man aber den Ort oder die Geschwindigkeit des Elektrons bestimmen, so wird man bei wiederholter Messung verschiedene, also fluktuierende Werte beobachten, was auf den Messvorgang am zeitlich stationären System zurückzuführen ist.

¹... mit dem Planckschen Wirkungsquantum $h = 6.62607015 \cdot 10^{-34}\text{Js} = 2\pi\hbar = 6.2831853071\dots \cdot 1.0545718176\dots \cdot 10^{-34}\text{Js}$. \hbar wird als *reduziertes* Plancksches Wirkungsquantum bezeichnet.

²Allerdings ist das Konzept der Energieerhaltung auf gekrümmten Raumzeiten in der allgemeinen Relativitätstheorie problematisch.

Längenskalen

In den gängigen kosmologischen Modellen unseres Universums wird davon ausgegangen, dass vor 13.8 Milliarden Jahren ein als 'Urknall' bezeichneter Schöpfungsprozess stattfand. Zwar legt Licht im flachen Raum in 13.8 Milliarden Jahren eine Strecke von 13.8 Milliarden Lichtjahren zurück, doch wegen der Expansion des Universums dürften sich unterdessen die entferntesten Objekte, die sich prinzipiell noch von der Erde aus in einer frühen Phase ihrer Entstehung nach dem Urknall beobachten liessen, weiter als 13.8 Milliarden Lichtjahre von uns entfernt befinden. Man setzt den Radius des beobachtbaren Universums im Rahmen des Standard-Urknallmodells zu etwa 46.6 Milliarden Lichtjahren oder $4.41 \cdot 10^{26}$ m an.

Es sind aber auch die Naturgesetze im Kleinsten, welche das Verhalten des Universums im Grossen bestimmen. So erweisen sich bei genauerem Hinsehen viele Quantitäten in der Natur als quantisiert, und im Falle der elektrischen Ladung stellt sich heraus, dass diese strikt portioniert auftritt. Diese Tatsache harret bislang einer fundierten Erklärung, auch wenn gewisse Konsistenzbetrachtungen im Rahmen erweiterter quantenfeldtheoretischer Theorien der Elementarteilchen und ihrer Wechselwirkungen Hinweise auf eine mögliche Lösung des Problems liefern. So konnten bislang nur Teilchen beobachtet werden, deren elektrische Ladung ein ganzzahliges Vielfaches (eines Drittels³) der Elementarladung $e_0 = 1.602176634$ As betrug. Fällt ein Teilchen durch eine elektrische Potentialdifferenz von einem Volt, so gelangt man auch schon zu einer für die Atom-, Kern-, Astro- und Teilchen-Physik praktischen Energieeinheit, dem Elektron(en)volt (eV) mit $1\text{eV} := 1.602176634 \cdot 10^{-19}$ J. Auch Energien sind bisweilen 'gequantelt', aber nicht zwingend so gleichmässig wie die elektrische Ladung.

Die uns bekannten Lebensformen halten sich auf einer logarithmischen Skala betrachtet im Mittelfeld zwischen den grössten und den kleinsten uns durch Beobachtung zugänglichen Objekte auf. Wo schliesslich die Grenze zwischen klassischer Physik und Quantenmechanik anzusetzen ist, ist kaum klar zu beantworten, zumal der Begriff der klassischen Physik heutzutage für eine Theorie steht, welche sich als Grenzfall aus der fundamentalen Naturbeschreibung durch die Quantenmechanik ergibt. Makroskopische Quantenphänomene wie die Levitation von Supraleitern über Magneten (Meissner-Ochsenfeld-Effekt) oder die Suprafluidität von Heliumisotopen sind wohlbekannt. Da der Mensch selbst aus Molekülen besteht, deren typische thermische Energie sich aus der Boltzmann-Konstante $k_B = 1.380649 \cdot 10^{-23} \text{J} \cdot \text{K}^{-1}$ (exakt!) und der Körpertemperatur $T \simeq 310\text{K}$ zu etwa $k_B T \simeq 0.027\text{eV}$ errechnet, gehen aber Energiequanten in diesem Energiebereich im thermischen Rauschen eines menschlichen Körpers als 'Detektor' unter und sind daher *einzel*n nicht wahrnehmbar.

Betrachtet man Quanten etwas höherer Energie - beispielsweise optische Photonen mit Energien weniger eV - so gelangt man in den Bereich der Auflösung konventioneller Lichtmikroskope, welche mit der (reduzierten) Wellenlänge λ ($\lambda = \lambda/(2\pi)$) oder der (Kreis-)Frequenz ν ($\omega = 2\pi \cdot \nu$) des Lichts als elektromagnetische Welle und der Energie $E_\gamma = h\nu = \hbar\omega = hc/\lambda = \hbar c/\lambda$ der mit der Welle assoziierten Photonen (γ -Quanten) sowie deren Impuls $p_\gamma = \hbar/\lambda = \hbar/\lambda = E_\gamma/c$ verknüpft ist. Mit $p_\gamma \lambda = \hbar$ beobachtet man mit der Impuls-Wellenlängen-Beziehung des Lichts eine an die Orts-Impuls-Unschärferelation erinnernde Beziehung.

Das durchschnittlich normal sehende Auge reagiert bei Tageslicht am sensibelsten auf Photonen einer Wellenlänge von $555\text{nm} = 5.55 \cdot 10^{-7}\text{m}$. Daraus berechnet sich eine Frequenz von $\nu \simeq 540\text{THz} = 5.4 \cdot 10^{14}\text{s}^{-1}$ und eine Photonenenergie von $E_\gamma \simeq 2.23\text{eV} \simeq 3.58 \cdot 10^{-19}\text{J}$. Solche Energien liegen im Bereich molekularer Bindungsenergien; die Wellenlänge übertrifft die Ausdehnung von Atomen aber noch um etwa 3 Zehnerpotenzen.

Teilchenphysiker[innen] kennen ja mit Vorteil den Wert des folgenden Produkts zweier Naturkonstanten⁴

$$\hbar c = 197.33\text{MeV} \cdot \text{fm} \simeq 0.2\text{GeV} \cdot \text{fm}, \quad (3)$$

³Man vergesse nicht die Quarks. Die Ladung des Up-Quarks beträgt $\frac{2}{3}e_0$, die des Down-Quarks $-e_0/3$.

⁴1fm = 10^{-15}m = 1 Femtometer = 1 Fermi. Aus dem Dänischen 'femten' = 'fünfzehn' bzw. nach Enrico Fermi (1901-1954), Nobelpreis für Physik 1938.

welcher sofort eine Abschätzung der in der Kern- und Teilchen-Physik auftretenden typischen Längenskalen bei gegebener Energieskala und umgekehrt zulässt. So beträgt die räumliche Ausdehnung eines Protons etwa 1fm, während die Energien der sich im Proton aufhaltenden Valenz-Quarks im Bereich von Bruchteilen eines GeV liegen .

Dringt man also weiter vor in die mikroskopische, durch Quanteneffekte dominierte Welt, so ist zudem eine kurze Diskussion der Feinstrukturkonstante

$$\alpha = \frac{e_0^2}{4\pi\epsilon_0\hbar c} \simeq 0.0072973525664(17) = \frac{1}{137.035999139(31)} \quad (4)$$

(mit der elektrischen Feldkonstante $\epsilon_0 = 8.8541878128(13)\text{As/(Vm)}$ und der Vakuum-Lichtgeschwindigkeit $c = 299792458\text{m/s}$) durchaus interessant. Das Besondere an α ist die Tatsache, dass α eine *Zahl* ist; der Wert der Feinstrukturkonstante ist unabhängig vom verwendeten Einheitensystem. In diesem Sinne ist α fundamental für die Physik in unserem Universum, und α hat in den vergangenen Jahrmilliarden keine bislang nachweisbare Änderung erfahren.

Damit ergibt sich folgende Liste wichtiger Längen- und Energie-Skalen in der Quantenmechanik, wobei die untenstehende hierarchische Gliederung zunächst durch die Feinstrukturkonstante α vorgegeben ist:

- Bohrscher Radius

$$r_B = \frac{4\pi\epsilon_0\hbar^2}{m_e e_0^2} = 5.2917721067(12) \cdot 10^{-11}\text{m} . \quad (5)$$

Der Bohrradius ist ein Mass für die Ausdehnung des Wasserstoffatoms im Grundzustand. Die Bindungsenergie wie auch die kinetische Energie des Elektrons der Masse $m_e = 9.10938356(11) \cdot 10^{-31}\text{kg} \simeq 511\text{keV}/c^2$ in diesem Zustand betragen rund 13.6eV; das Elektron bewegt sich in einer *klassischen* Betrachtung (Bohrmodell) mit der Geschwindigkeit $v_{e^-} \simeq \alpha c \simeq c/137$.

- Reduzierte Comptonwellenlänge des Elektrons

$$\lambda_c = \alpha r_B = \frac{\lambda_c}{2\pi} = \frac{\hbar}{m_e c} = 3.86159267533(18)\text{m} . \quad (6)$$

Auf dieser mit der Ruheenergie des Elektrons von 511keV assoziierten Längenskala treten in der λ -Umgebung eines Elektrons geringfügige Abweichungen vom $1/r^2$ -Verhalten des Coulombfeldes auf, die in einem klassischen Bild durch die Bildung virtueller Elektron-Positron-Paare erklärt werden ('Vakuumpolarisation').

- Klassischer Elektronenradius

$$r_e = \alpha \lambda_c = \alpha^2 r_B = \frac{1}{4\pi\epsilon_0} \frac{e_0^2}{m_e c^2} = 2.8179403227(19) \cdot 10^{-15}\text{m} . \quad (7)$$

Stellt man sich ein Elektron als eine Ladungswolke vor und berechnet im Rahmen dieser naiven Modellvorstellung die ungefähre räumliche Ausdehnung dieser Wolke, bei der die potenzielle elektrostatische Energie der Ladungsverteilung der Ruheenergie $m_e c^2$ des Elektrons entspricht, so gelangt man schliesslich zu Werten in der Grössenordnung des klassischen Elektronenradius.

- Elektroschwache Längenskala

$$l_W \simeq \lambda_{c,Z} = \frac{\hbar}{m_Z c} = 2.163967(50) \cdot 10^{-18}\text{m} . \quad (8)$$

Die reduzierte Comptonwellenlänge $\lambda_{c,Z}$ des Z -Bosons (oder der etwas leichteren W^\pm -Bosonen) als Vermittler der schwachen Wechselwirkung ist zugleich ein Mass für die Reichweite dieser Kraft. Zerfällt beim β^- -Zerfall des Neutrons ein u -Quark in ein d -Quark, so geschieht dies in einer l_W -Umgebung der betreffenden Quarks. Das elementare Z -Boson ist mit einer Ruhemasse von $m_Z = 91.1876(21)\text{GeV}/c^2$ etwa 100mal schwerer als das Proton ($m_P = 938.2720813(58)\text{MeV}/c^2 = 1836.15267389(17)m_e$).

- Planck-Länge

$$l_P = \sqrt{\frac{\hbar G}{c^3}} = 1.616229(38) \cdot 10^{-35} \text{m.} \quad (9)$$

Auf dieser Längenskala verliert die Raumzeit wahrscheinlich alle klassischen Eigenschaften und muss in einem quantenmechanischen Rahmen untersucht werden. Stellt man sich vor, dass 'Vakuum-Massenfluktuationen' im Allerkleinsten so gross werden, dass sie gar der Bildung virtueller schwarzer Löcher Vorschub leisten, so lässt sich aus dem berühmten Resultat der allgemeinen Relativitätstheorie (ART) für den Schwarzschildradius r_S eines nicht rotierenden, ungeladenen schwarzen Loches der Masse M_{BH}

$$r_S = \frac{2G}{c^2} M_{BH} \quad (10)$$

mit einer Ortsunschärfe $\Delta x \simeq l_P \simeq r_S$ und Massenunschärfe $\Delta m \simeq m_P \simeq M_{BH}$ auf Planck-Skala mit

$$\Delta x \stackrel{UR}{\simeq} \frac{\hbar}{2m_P c} \quad \text{und} \quad \Delta x \stackrel{ART}{\simeq} \frac{2G}{c^2} m_P \quad \text{sofort} \quad \Delta x^2 \simeq \frac{G\hbar}{c^3} = l_P^2 \quad (11)$$

folgern. Im Kleinsten entpuppt sich die Raumzeit wohl als 'Quantenschaum' und verliert ihre in der Elementargeometrie suggerierten Eigenschaften. Natürlich ist $G = 6.67430(15) \cdot 10^{-11} \text{kg}^{-1} \cdot \text{s}^{-2} \cdot \text{m}^3$ die Gravitationskonstante.

Fundamentale Konstanten

Man mag sich fragen, ob 'Naturkonstanten' wie die Plancksche Konstante oder die Lichtgeschwindigkeit *zeitabhängig* sind; in einem verallgemeinerten Rahmen könnte die Plancksche Konstante gar als Raumzeit-Funktion $h(t, \vec{x})$ angesetzt werden.

Da aber Messungen auf einem Vergleichsprozess beruhen, macht es tatsächlich wenig Sinn, bei einer solchen Fragestellung einheitenbehaftete Grössen vorbehaltlos zu diskutieren. So wird beispielsweise davon ausgegangen, dass sich das Universum in den vergangenen Jahrmilliarden ausgedehnt hat; diese Aussage ist aber nur dann sinnvoll, wenn klar ist, dass die mit ihren verschiedenen Eigenschaften als Massstäbe dienenden Elementarteilchen als unveränderlich vorausgesetzt werden können.

Daher sind *dimensionslose* (einheitenunabhängige) Grössen besser geeignet, um als 'fundamentale Konstanten' im Rahmen theoretischer Betrachtungen untersucht zu werden. Ob sich der Wert der Feinstrukturkonstante α oder die Massenverhältnisse diverser Elementarteilchen ($\mu = m_P/m_e \simeq 6\pi^5$) als eher zufällig oder als ähnlich fundamental wie die Kreiszahl $\pi = 3.14159265\dots$ erweisen werden, ist eine interessante unbeantwortete Frage. Experimentelle Messungen beschränken die Zeitabhängigkeit der oben erwähnten Verhältnisse auf

$$\frac{|\dot{\alpha}|}{\alpha} < 10^{-17} a^{-1}, \quad \frac{|\dot{\mu}|}{\mu} < 10^{-16} a^{-1}. \quad (12)$$

Natürlich definiert jedes Verhältnis der Ruhemassen zweier (Elementar-)Teilchen eine fundamentale dimensionslose Grösse. Eine weitere dimensionslose Grösse stellt das Verhältnis R der elektrostatischen und gravitostatischen Anziehungskraft zwischen einem Elektron-Positron-Paar

$$F_{el}^{e^\pm} = \frac{e_0^2}{4\pi\epsilon_0 r^2}, \quad F_{grav}^{e^\pm} = G \frac{m_e^2}{r^2} \quad \rightarrow \quad R = F_{el}^{e^\pm} / F_{grav}^{e^\pm} = \frac{e_0^2}{4\pi\epsilon_0 G m_e^2} \simeq 4.1656 \cdot 10^{42} \quad (13)$$

dar.

Weitere dimensionslose Parameter tauchen im Standardmodell der Teilchenphysik in Form der starken und elektroschwachen Kopplungskonstanten und in der Cabibbo-Kobayashi-Maskawa-Matrix und der Pontecorvo-Maki-Nakagawa-Sakata-Matrix auf.

Einheiten in der Teilchenphysik

SI-Einheiten

Physikalische Grössen werden in der Experimentalphysik gemessen und in der theoretischen Physik berechnet. Der Wert G einer skalaren physikalischen Grösse \mathcal{G} als eine an einem Objekt der Physik quantitativ bestimmbare Eigenschaft eines Vorgangs oder Zustands wird üblicherweise als Produkt aus einem Zahlenwert $\{G\}$ und einer Masseinheit $[G]$ angegeben:

$$G = \{G\}[G]. \quad (14)$$

Der Begriff der *skalaren* Grösse wird hier natürlich im Sinne einer durch einen einzelnen Zahlenwert und die Masseinheit charakterisierbaren Grösse verwendet und hat hier keinen expliziten gruppentheoretischen Hintergrund. G kann also auch eine Komponente einer tensoriellen Grösse sein. Üblicherweise wird die Einheit $[G]$ durch ein Einheitenzeichen abgekürzt notiert (Meter \rightarrow m), und Einheiten können durch Potenzierung und Produktbildung zu neuen Einheiten abgeleitet oder durch Präfixe (welche oft ganzzahligen Potenzen von 10^3 entsprechen) so reskaliert werden, dass der Zahlenwert $\{G\}$ in einem wünschenswerten numerischen Bereich liegt (m \rightarrow km = 10^3 m).

Nach der Internationalen Organisation für Normung (ISO) sind Grössensymbole (Formelzeichen) in kursiver Schrift zu schreiben, Einheitenzeichen in aufrechter Schrift. Zwischen Zahlenwert und Einheit steht kein Multiplikationszeichen.

Im heutzutage am weitesten verbreiteten *Internationalen Einheitensystem* oder *SI-System* (französisch *Système international d'unités*) wurden nach praktischen Gesichtspunkten 7 Basiseinheiten zur Angabe physikalischer Grössen definiert. Die sieben Basiseinheiten heissen *Meter* (abgekürzt m), *Sekunde* (s), *Kilogramm* (kg), *Ampere* (A), *Kelvin* (K), *Mol* (mol) und *Candela* (cd). Dabei ist zu bemerken, dass die Einheit Ampere im Gegensatz zu ihrem Namensgeber André-Marie Ampère ohne accent grave geschrieben wird. Die Einheit Candela ist mit dem optischen Sinneseindruck einer Leuchte auf Auge und Gehirn eines Menschen verknüpft und soll wegen ihrer geringen Relevanz für die Teilchenphysik hier nicht weiter diskutiert werden.

Die historischen Hintergründe der heute gebräuchlichen SI-Einheiten stellen ein umfangreiches und interessantes Thema dar - so war die Sekunde bis in die 1950er Jahre als der 86400ste Teil des mittleren Sonnentages definiert, der Meter sollte den zehnmillionsten Teil des Erdquadranten des vom Nordpol ausgehenden und Paris berührenden Meridians betragen und wurde entsprechend einst durch zwei Markierungen auf Metallstäben, den *Urmetern*, definiert. Zudem ist ein Sekundenpendel recht präzise einen Meter lang. Doch auch diese Betrachtungen würden hier zu weit führen.

Seit dem 20. Mai 2019 sind sämtliche SI-Einheiten mit Hilfe physikalischer Konstanten definiert und somit nicht länger von menschengemachten Artefakten wie beispielsweise dem seit 1889 in Sèvres bei Paris aufbewahrten *Urkilogramm* oder einem Urmeter abhängig. Aktuell gilt:

Die Sekunde ist seit 1967 als das 9192631770fache der Periodendauer der dem Übergang zwischen den beiden Hyperfeinstruktur-niveaus des Grundzustandes von Atomen des Nuklids ^{133}Cs entsprechenden Strahlung definiert, d.h. die entsprechende Frequenz beträgt genau 9192631770s^{-1} . Diese Definition der Sekunde bezieht sich auf Cäsiumatome im Ruhezustand, bei einer Temperatur von 0K. Die Übergangslinienbreite ist sehr gering; zudem ist die Tatsache von Vorteil, dass ^{133}Cs das einzige stabile Isotop des Cäsiums darstellt, wodurch messtechnisch nachteilige Isotopeneffekte ausgeschlossen werden können.

Durch die Definition der Sekunde wird mit der Festlegung der Vakuum-Lichtgeschwindigkeit im Jahre 1983 auf den exakten Wert $c = 299792458\text{m}\cdot\text{s}^{-1}$ der Meter als die Strecke definiert, welche das Licht in einer 299792458stel Sekunde zurücklegt. Innerhalb der heutigen experimentellen Messgenauigkeit kann für die

Vakuum-Lichtgeschwindigkeit keine Dispersion, d.h. eine Abhängigkeit der Lichtgeschwindigkeit von der Frequenz einer elektromagnetischen ebenen Welle oder der Energie eines Photons, festgestellt werden.

Seit 2019 ist damit auch das Kilogramm als schwere und träge Masseneinheit durch die Festlegung der Planckschen Konstante auf $h = 6.62607015 \text{ kg} \cdot \text{m}^2 \cdot \text{s}^{-1}$ festgelegt. Dabei gilt die Masseneinheit Kilogramm (kg) trotz des Präfixes k = 1000, welches das Kilogramm die tausendfache Masse des Gramms (g) notiert, als SI-Basiseinheit. Aus den SI-Basiseinheiten s, m und kg bildet man die abgeleitete SI-Energieeinheit *Joule* ($1 \text{ J} = 1 \text{ kg} \cdot \text{m}^2 \cdot \text{s}^{-2}$).

Nach wie vor gilt das Ampere als SI-Basiseinheit und nicht die abgeleitete SI-Ladungseinheit Coulomb (C), auch wenn das Ampere seit 2019 durch die Festlegung der Elementarladung (bzw. der Protonladung) auf den exakten Wert $e_0 = 1.602176634 \cdot 10^{-19} \text{ C}$ mit $1 \text{ C} = 1 \text{ A} \cdot \text{s}$ erklärt ist.

Die Temperatureinheit Kelvin (bzw. Grad Celsius) wird neu durch die Festlegung der Boltzmann-Konstante auf den exakten Wert $k_B = 1.380649 \cdot 10^{-23} \text{ J/K}$ definiert. Früher war das Kelvin als der 273.16te Teil der thermodynamischen Temperatur des Tripelpunktes des Wassers definiert, wobei bereits die Definition des Wassers als Isotopengemisch eine gewisse Herausforderung darstellt. Bekanntlich ist die mittlere Energie der voll-elastischen punktförmigen Teilchen eines idealen Gases der Temperatur T gegeben durch $\bar{E} = 3k_B T/2$, der Energiebegriff ist also unmittelbar mit dem Temperaturbegriff verknüpft.

Das Mol als SI-Basiseinheit der Stoffmenge bezeichnete ursprünglich die Zahl der Atome in einer Probe aus 12g reinem ^{12}C . Seit 2019 enthält 1mol eines Stoffes genau $6.02214076 \cdot 10^{23}$ Teilchen, also rund 600 Trilliarden. In gewissem Sinne ist das Mol der Name einer grossen Zahl, die dem Zählen der Teilchen in einer makroskopischen Stoffmenge dienlich ist. Es ist klar, dass auf diese SI-Basiseinheit ohne grössere Not verzichtet werden könnte.

Einheiten in der Teilchenphysik

Im SI-Einheitensystem gilt also für das Beispiel des reduzierten Planckschen Wirkungsquantums \hbar exakt

$$\hbar = \{\hbar\}[\hbar], \quad \{\hbar\} = 6.62607015 \cdot 10^{-34} / (2 \cdot \pi) = 1.0545718176 \dots \cdot 10^{-34}, \quad [\hbar] = \text{kg} \cdot \text{m}^2 \cdot \text{s}^{-1} = \text{Js}, \quad (15)$$

wobei die Auslassung eines Punktes in der Einheit Js = J·s keine Notationsmissverständnisse verursachen kann.

In der Teilchenphysik wird ein *natürliches* Einheitensystem verwendet, in welchem

$$\hbar = c = k_B = \epsilon_0 = 1 \quad (16)$$

gilt. Der Wert der elektrischen Feldkonstante (auch Permittivität des Vakuums, Dielektrizitätskonstante des Vakuums oder Influenzkonstante genannt) beträgt $\epsilon_0 = 8.8541878128(13) \cdot 10^{-12} \text{ A}^2 \text{ s}^4 \text{ kg}^{-1} \text{ m}^{-3}$, wobei die Klammerzahl (13) die aktuelle Unsicherheit in den beiden letzten Dezimalen des experimentell bestimmten Zahlenwerts von ϵ_0 angibt. Vor dem 20. Mai 2019 galt noch exakt $\epsilon_0 = 10^7 / (4\pi \cdot 299792458^2) \text{ As} / (\text{Vm})$.

Generell zeichnen sich natürliche Einheitensysteme dadurch aus, dass Naturkonstanten (bis auf möglichst einfache numerische Faktoren) als Basiseinheiten zugrunde gelegt werden. Durch die vier Bedingungen in (16) reduziert sich die Anzahl der 5 in der Praxis der Teilchenphysik relevanten SI-Einheiten s, m, kg, C (oder A) und K auf eine einzige festzulegende Basiseinheit, mit welcher der Wert aller relevanten Grössen in der Teilchenphysik ausgedrückt werden kann. Oft bedient man sich hierzu der Energieeinheit Elektron(en)volt (eV), welche durch $1 \text{ eV} = 1.602176634 \cdot 10^{-19} \text{ J}$ definiert ist. In der Kernphysik tauchen oft Energien der Grössenordnung Megaelektronenvolt und Gigaelektronenvolt ($1 \text{ MeV} = 10^6 \text{ eV}$, $1 \text{ GeV} = 10^9 \text{ eV} = 1.602176634 \cdot 10^{-10} \text{ J}$) auf, in der Hochenergiephysik arbeitet man zur Zeit im TeV-Bereich (Teraelektronenvolt, $1 \text{ TeV} = 10^{12} \text{ eV}$).

Im SI-System ist ja $\hbar c = 197.3269982 \dots \text{MeV} \cdot \text{fm} \simeq 200 \text{MeV} \cdot \text{fm} = 0.2 \text{GeV} \cdot \text{fm}$. In der Teilchenphysik gilt aber $\hbar c = 1$. Deshalb kann beispielsweise eine Distanz d von einem Femtometer wegen $0.2 \text{GeV} \cdot \text{fm} \simeq 1$ in der Teilchenphysik mit Hilfe inverser Energieeinheiten alternativ durch $d = 1 \text{fm} \simeq 1/(0.2 \text{GeV}) = 5 \text{GeV}^{-1}$ angegeben werden.

Im allgemeinen Fall lässt sich der Wert G_{TP} einer in der Teilchenphysik durch Potenzen der gewählten Basisenergieeinheit ausgedrückten physikalischen Grösse in ihren Wert G_{SI} in SI-Einheiten umrechnen, indem man G_{TP} mit Potenzen der Konstanten \hbar , c , k_B und ϵ_0 im SI-System derart multipliziert, dass G_{SI} die korrekten Einheiten im SI-System erhält:

$$G_{SI} = G_{TP} \cdot \hbar_{SI}^\alpha \cdot c_{SI}^\beta \cdot k_{B,SI}^\gamma \cdot \epsilon_{0,SI}^\delta \quad \text{für geeignete } \alpha, \beta, \gamma, \delta. \quad (17)$$

Das Coulomb-Gesetz für die Kraft F zwischen zwei Ladungen q und Q im Abstand r lautet im SI-System

$$F = \frac{1}{4\pi\epsilon_0} \frac{qQ}{r^2}, \quad (18)$$

in der Teilchenphysik

$$F = \frac{1}{4\pi} \frac{qQ}{r^2}, \quad (19)$$

während im oben nicht besprochenen Gaußschen Einheitensystem

$$F = \frac{qQ}{r^2} \quad (20)$$

gilt.

Vorbemerkung: Raum-Zeit-Strukturen

Von Newton zu Einstein

Als Vorbemerkung zur nachfolgenden mathematisch-formalen Behandlung der relativistischen Raumzeitstruktur soll in diesem Unterkapitel zuerst die Grundidee der speziellen Relativitätstheorie in stark geraffter Form rekapituliert werden.

In der vor-relativistischen Physik stellte man sich den physikalischen Raum als dreidimensionalen euklidischen affinen Raum V_E^3 vor, welcher in Anlehnung an die Ideen Isaac Newtons die Bühne für die einer absoluten, gleichförmig und ohne Beziehung auf irgendwelche Gegenstände verfließenden Zeit unterworfenen Vorgänge liefert. Diese für gewisse Betrachtungen nach wie vor nützliche Vorstellung basiert auf der alltäglichen Erfahrung, dass man zwischen zwei Orten im physikalischen Raum - im scharf lokalisierten Idealfall zwischen zwei Punkten - einen Abstand messen kann. Dabei ist der physikalische Raum bei Abwesenheit starker Gravitation in sehr guter Näherung homogen und isotrop (d.h. strukturell richtungsunabhängig), ein 'Ursprung' oder 'Raum-Mittelpunkt' lässt sich nicht finden, was das Attribut der Affinität in der mathematischen Raumbezeichnung begründet. Zudem lässt die Beobachtung unserer räumlichen Umwelt die Einführung von Koordinatensystemen zu, welche jeden Raumpunkt durch drei Koordinaten (x, y, z) charakterisieren. Liegt schliesslich ein solches räumliches und durch einen denkenden Beobachter definiertes Gebilde vor, das erforderlich ist, um das zeitliche Verhalten ortsabhängiger physikalischer Grössen eindeutig und vollständig zu beschreiben, so spricht man von einem Bezugssystem.

Die in Bezugssystemen verwendeten Koordinatensysteme können allerdings 'krummlinig' oder beschleunigt sein und aus praktischer Sicht somit recht unangenehm. Daher mag man vorerst und mit Vorteil sogenannte *Inertialsysteme* betrachten, in welchen *alle* drei Newtonschen Gesetze uneingeschränkt gelten sollen; das erste dieser Gesetze nennt die geradlinig-gleichförmige Bewegung von kräftefreien Körpern als wesentliche Eigenschaft eines Inertialsystems, wobei man bereits hier über die Natur der Kräfte ins Philosophieren verfallen könnte. Das zweite Newtonsche Gesetz definiert Kräfte durch die von ihnen verursachten Beschleunigungen. Das dritte Newtonsche Gesetz schliesslich verlangt, dass es zu jeder Kraft eine Gegenkraft geben muss, wobei ausschliesslich Kräfte gemeint sind, die auf Wechselwirkungen zwischen Körpern zurückgehen. In solchen Inertialsystemen, welche unpräzise auch als 'unbeschleunigte' Bezugssysteme bezeichnet werden, können spezielle Koordinatensysteme in Übereinstimmung mit der idealisiert angenommenen Raumstruktur in homogener und isotroper Weise so eingeführt werden, dass der Abstand $d(P_1, P_2) > 0$ zwischen zwei Punkten P_1 und P_2 mit Koordinaten (x_1, y_1, z_1) und (x_2, y_2, z_3) durch den Satz des Pythagoras

$$d(P_1, P_2)^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 \geq 0 \quad (21)$$

berechnet werden kann. Abbildung 1 stellt zwei solche Punkte im V_E^3 mit ihren als Spaltenvektoren notierten Koordinatentripeln mittels einer zweidimensionalen Grafik dar; die Koordinatenachsen stehen dabei *senkrecht* zueinander und die Koordinatenwerte sind so geeicht, dass der Satz des Pythagoras auch tatsächlich immer in der Form von Gleichung (21) anwendbar ist. Ein solches Koordinatensystem nennt man *kartesisch*.

Tatsächlich impliziert die Existenz einer Distanzfunktion d gemäss Gleichung (21) auch die Existenz aller gängigen metrischen Alltagsbegriffe wie beispielsweise dem des Winkels, so dass im Falle eines kartesischen Koordinatensystems von senkrechten Koordinatenachsen gesprochen werden kann. Auch wenn verschiedene Beobachter in unterschiedlichen Inertialsystemen verschiedene kartesische Koordinatensysteme zur Beschreibung physikalischer Vorgänge spontan auswählen, so kommt dem Abstandsbegriff anscheinend eine fundamentalere, von der Wahl eines kartesischen Koordinatensystems unabhängige physikalische Bedeutung zu. $d(P_1, P_2)$ ist eine in der nicht-relativistischen Physik beobachterunabhängige Grösse. Sind die Koordinaten der Punkte $P_{1,2}$ im gestrichelten und gestrichelten kartesischen Koordinatensystem in Abbildung 1 alternativ durch $(x'_{1,2}, y'_{1,2}, z'_{1,2})$ gegeben, so gilt bei geeigneter Eichung wiederum der Satz des Pythagoras in einem weiteren

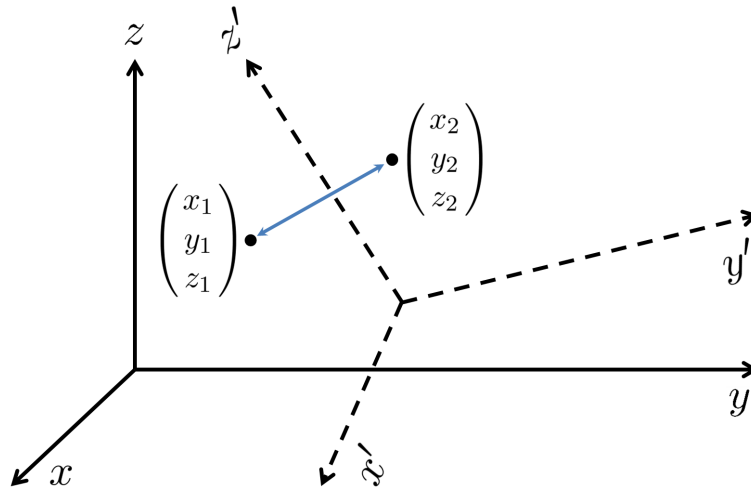


Abbildung 1: Kartesische Koordinatensysteme. Der skalare Abstand zweier Punkte ist unabhängig vom Beobachterzustand, im Gegensatz zu den einzelnen zugewiesenen (kartesischen) Koordinaten.

Inertialsystem

$$d(P_1, P_2)^2 = (x'_1 - x'_2)^2 + (y'_1 - y'_2)^2 + (z'_1 - z'_2)^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2. \quad (22)$$

Diese Beobachtung ist Ausdruck einer in der Mathematik mit dem Gruppensymbol $ISO(3)$ bezeichneten *Symmetrie* der euklidischen affinen Raumes V_E^3 , nämlich der Isometriengruppe oder *Bewegungsgruppe*.

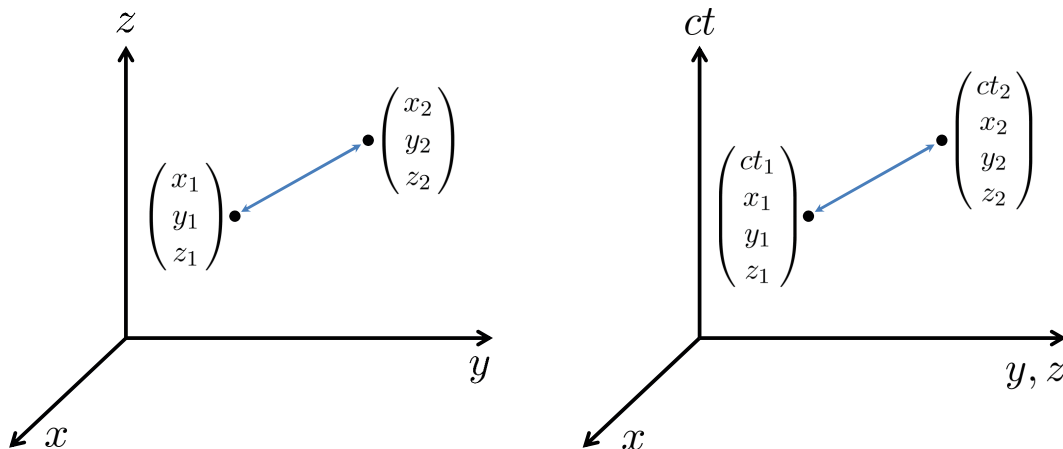


Abbildung 2: Vom dreidimensionalen euklidischen Raum zur vierdimensionalen Minkowski-Raumzeit.

Fasst man nun wie einst Albert Einstein die Zeit als eine vierte Dimension auf und betrachtet neu die flache vierdimensionale *Raumzeit*, deren Punkte als *Ereignisse* zu einem Zeitpunkt t an einem Ort mit Koordinaten (x, y, z) interpretierbar sind, dann wird die grafische Darstellung der Begebenheiten in diesem Raum zur Herausforderung. Es sind daher in Abbildung 2 zwei räumliche Achsen zu *einer* Achse zusammengefasst. Zudem misst man die Zeit traditionell in einer anderen Einheit als räumliche Abstände. Wenn dieser Umstand stört, kann die Zeit t mit einer Geschwindigkeit c multipliziert und somit mit Hilfe von Längeneinheiten angegeben

werden. Wählt man für c schliesslich die Lichtgeschwindigkeit, so ergibt sich im Rahmen der speziellen Relativitätstheorie eine erstaunliche Erkenntnis: Räumliche Abstände sind keine absoluten Grössen mehr, vielmehr hängen sie ähnlich wie einzelne Koordinaten eines Raumpunktes von der Lage und dem Bewegungszustand eines Beobachters ab. Allerdings tritt an die Stelle des euklidischen Abstands (21) in Inertialsystemen mit kartesischer Raumzeit-Koordinatisierung eine neue Grösse

$$s(P_1, P_2)^2 = (ct_1 - ct_2)^2 - (x_1 - x_2)^2 - (y_1 - y_2)^2 - (z_1 - z_2)^2, \quad (23)$$

welche manchmal als (quadratischer) *Minkowski-Abstand* zweier Ereignisse $P_{1,2}$ bezeichnet wird. Es stellt sich heraus, dass die durch einen Satz des Pythagoras mit 'falschen Vorzeichen' berechenbare Grösse $s(P_1, P_2)^2$ eine Art Abstand zwischen beliebigen Ereignispaaren definiert, welcher unabhängig vom Bewegungszustand inertialer Beobachter ist, welche jeweils ein kartesisches Raumzeit-Koordinatensystem eingeführt haben, sodass eben Gleichung (23) uneingeschränkt gilt. Offensichtlich ist s^2 nicht positiv definit, es ist also nicht empfehlenswert, durch Wurzelziehen aus $s(P_1, P_2)^2$ einen im üblichen Sinne zu verstehenden Abstand zu berechnen. Die Existenz der Funktion s^2 wird aus dem Einsteinschen Postulat der Konstanz der Lichtgeschwindigkeit in allen Inertialsystemen hergeleitet.

Wir wollen die Existenz des Minkowski-Abstandes hier nicht in aller Glorie herleiten, zur Begründung aber des Pudels Kern kurz beleuchten. Dazu betrachten wir zwei zueinander unrotierte Inertialsysteme IS und IS' mit gemeinsamem Raumzeitursprung, wobei sich IS' aus Sicht von IS wie in Abbildung 3 dargestellt in x - bzw. x' -Richtung mit der Geschwindigkeit \vec{v} bewegt. Es ist also so, dass die Ursprünge der beiden Koordinatensysteme zu einem Startzeitpunkt $t = 0$ in IS beziehungsweise $t' = 0$ in IS' übereinstimmen und in diesem Moment auch alle drei räumlichen Koordinatenachsen von IS und IS' zur Deckung kommen. Vom Ursprung ($t = t' = 0, x = x' = y = y' = z = z' = 0$) soll nun eine von einem sehr kurzen Lichtblitz ausgelöste Lichtkugelwelle ausgehen. Ein Beobachter nimmt diesen Blitz etwas später wahr, wobei die Raumzeitkoordinaten dieses Ereignisses E in IS durch (ct, x, y, z) bzw. durch (ct', x', y', z') in IS' gegeben sind. Da nun die Licht-

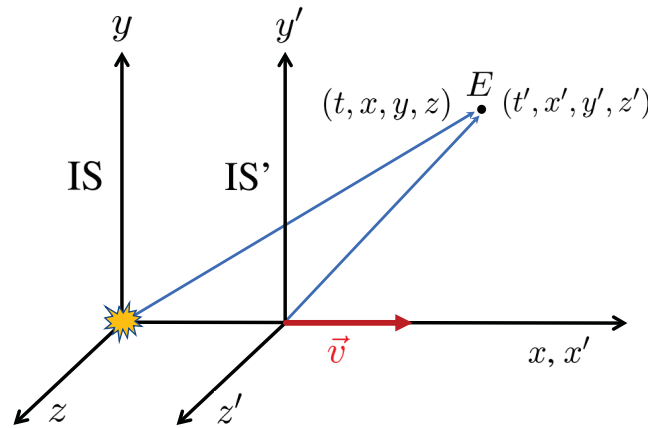


Abbildung 3: Ein kurzer Lichtblitz im gemeinsamen Raumzeitursprung von IS und IS' wird von einem Beobachter mit Koordinaten (t, x, y, z) in IS bzw. (t', x', y', z') in IS' beobachtet: Ereignis E .

geschwindigkeit c aus Sicht beider Inertialsysteme denselben Wert besitzt und der Lichtblitz vom Ursprung beider Inertialsysteme beim Startpunkt der Zeitmessung $t = t' = 0$ ausging, legt das Licht bis zum Ereignis E in IS die Strecke $r = ct = \sqrt{x^2 + y^2 + z^2}$ und in IS' die Strecke $r' = ct' = \sqrt{x'^2 + y'^2 + z'^2}$ zurück; in beiden Inertialsystemen gilt also mit $x^0 = ct$ und $x'^0 = ct'$, $x^1 = x$, $x^2 = y$, $x^3 = z$ und entsprechend $x'^1 = x'$, $x'^2 = y'$, $x'^3 = z'$

$$s(\text{Lichtblitz}, E)^2 = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = 0 = (x'^0)^2 - (x'^1)^2 - (x'^2)^2 - (x'^3)^2. \quad (24)$$

Man zeigt schliesslich leicht, dass aus der Forderung nach Homogenität und Isotropie des flachen physikalischen Raumes mit der Feststellung $s^2 = 0$ in Gleichung (24) die allgemeine Invarianz des Minkowski-Abstandes $s(P_1, P_2)^2$ bezüglich beliebigen Inertialsystemwechseln folgt.

Die Stellung eines Koordinatenindex - wie zum Beispiel bei der obigen Zuweisung $x^2 = y$ - wird, wie wir bald sehen werden, im Gegensatz zur galileischen, nicht-relativistischen Physik in relativistischer Notation eine nichttriviale Bedeutung erlangen, und tatsächlich ist dann $x^2 = -x_2$.

Dass wir in einer Welt leben, in welcher Abstands begriffe eine Rolle spielen, ist letztlich unerklärt; in der Topologie spielt der Abstands begriff schliesslich auch keine Rolle. Offensichtlich existieren in der Natur Mechanismen, welche den Dingen eine relative Grösse verleihen. Letztlich bleibt zu bemerken, dass der Raum - beziehungsweise die Raumzeit - wohl eine reichhaltige Struktur aufweist. Im Kleinsten entpuppt sich die Raumzeit wohl als 'Quantenschaum', weder flach noch sonst in einem durch klassische Begriffe beschreibbaren Zustand. Tatsächlich ist es trotz intensiver Bemühungen bislang nicht gelungen, eine befriedigende Quantentheorie der Raumzeit zu konstruieren, welche ja zugleich auch eine Quantentheorie der Gravitation darstellen sollte, wie es die eben dargelegten Ausführungen nahelegen. Letztlich sollte eine solche Theorie auch die nicht-gravitativen Aspekte der Physik mit einbeziehen.

In den folgenden Kapiteln sollen die obigen Betrachtungen zur klassischen, flachen Raumzeit für die mathematische Anwendung formalisiert werden.

Isometriegruppen des dreidimensionalen euklidischen Vektorraumes:

$$O(3) \ni SO(3)$$

Notation

Wir betrachten einen dreidimensionalen euklidischen Vektorraum V_E^3 , der strukturell isomorph zum Koordinatenraum \mathbb{R}_R^3 mit dem Standardskalarprodukt (\cdot, \cdot) ist. V_E^3 kann als Prototyp für den physikalischen, nichtrelativistischen dreidimensionalen flachen Raum dienen, wenn auch zu bemerken ist, dass in diesem wegen der Translationsinvarianz kein Ursprung oder Nullvektor ausgezeichnet ist; ein besseres Modell für den physikalischen Raum wäre daher durch den affinen euklidischen Punktraum E_3 gegeben, welcher in einer relativistischen Theorie schliesslich zu einer vierdimensionalen Raumzeit verallgemeinert werden muss.

Für zwei Vektoren $x, y \in \mathbb{R}_R^3$ ist in Anknüpfung an die obigen Bemerkungen mit der Einsteinschen Summenkonvention (Summation über doppelt auftretende, oft sowohl hoch- und tiefgestellte Indizes), dem Kronecker- δ -Symbol und der üblichen Matrizenmultiplikation das Standardskalarprodukt gegeben durch

$$(x, y) := x^1 y^1 + x^2 y^2 + x^3 y^3 = \sum_{i=1}^3 x^i y^i = x^i y_i = x^i \delta_{ij} y^j = x^T \mathbb{1}_3 y = x^T y = 'x \cdot y' = '\vec{x} \cdot \vec{y}' = 'xy', \quad (25)$$

mit

$$x = \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = (x^1, x^2, x^3)^T, \quad y = \begin{pmatrix} y^1 \\ y^2 \\ y^3 \end{pmatrix} = (y^1, y^2, y^3)^T \quad (26)$$

und

$$\delta_{ij} = \begin{cases} 1 & : i = j \\ 0 & : i \neq j \end{cases}, \quad \mathbb{1}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \text{diag}(1, 1, 1). \quad (27)$$

Es gibt in V_E^3 eine Orthonormalbasis (ONB) $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ (\rightarrow Orthogonalisierungsverfahren von Erhard Schmidt) mit $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij}$ für $i, j = 1, 2, 3$. Dabei bezeichnet $\langle \cdot, \cdot \rangle$ das Skalarprodukt in V_E^3 .

Zwei Vektoren $\mathbf{x}, \mathbf{y} \in V_E^3$ besitzen somit eine Darstellung

$$\mathbf{x} = x^i \mathbf{e}_i, \quad \mathbf{y} = y^j \mathbf{e}_j, \quad (28)$$

wobei nun die zwei Spaltenvektoren

$$x = (x^1, x^2, x^3)^T, \quad y = (y^1, y^2, y^3)^T \quad (29)$$

als *Koordinatenvektoren* von \mathbf{x} und \mathbf{y} bezüglich der Basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ bezeichnet werden. Offensichtlich ist unter Verwendung der Linearität des Skalarprodukts

$$\langle \mathbf{x}, \mathbf{y} \rangle = \langle x^i \mathbf{e}_i, y^j \mathbf{e}_j \rangle = x^i y^j \langle \mathbf{e}_i, \mathbf{e}_j \rangle = x^i y^j \delta_{ij} = x^T y = (x, y). \quad (30)$$

Aktive isometrische Transformationen des euklidischen Raumes V_E^3

Wir betrachten eine lineare, isometrische Abbildung auf V_E^3 , die sogar das Skalarprodukt zwischen allen Vektoren erhält

$$\mathbf{R} : V_E^3 \xrightarrow{\text{bijektiv}} V_E^3, \quad \mathbf{x} \mapsto \mathbf{x}' = \mathbf{R}(\mathbf{x}), \quad \langle \mathbf{R}(\mathbf{x}), \mathbf{R}(\mathbf{y}) \rangle = \langle \mathbf{x}', \mathbf{y}' \rangle = \langle \mathbf{x}, \mathbf{y} \rangle, \quad (31)$$

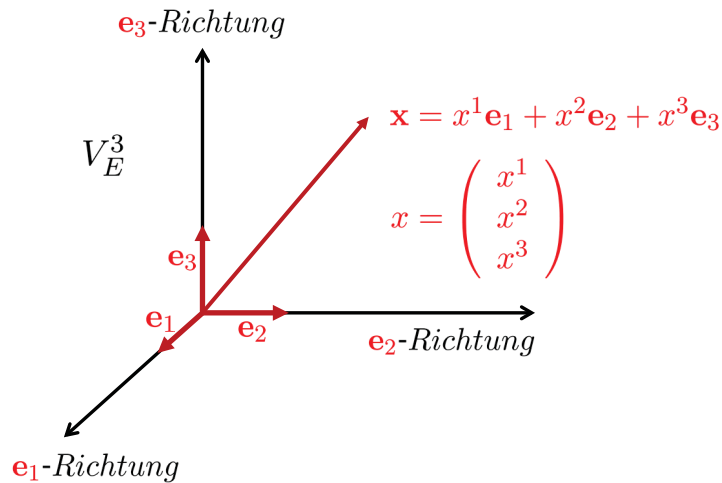


Abbildung 4: Der dreidimensionale euklidische Raum V_E^3 als Modell für dreidimensionale physikalische (Tangential-)Räume ist ein Vektorraum mit Skalarprodukt, $\langle \cdot, \cdot \rangle$, welches zwei 'physikalischen' Vektoren $\mathbf{x}, \mathbf{y} \in V_E^3$ einen reellen Wert $\langle \mathbf{x}, \mathbf{y} \rangle$ zuweist. Ein[e] Physiker[in] wählt in V_E^3 spontan und mit Vorteil eine *Orthonormal-Basis* (ONB) $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ mit $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij}$, durch welche jeder Vektor $\mathbf{x} \in V_E^3$ als Linearkombination $\mathbf{x} = x^1 \mathbf{e}_1 + x^2 \mathbf{e}_2 + x^3 \mathbf{e}_3$ dargestellt werden kann; dabei ist $x = (x^1, x^2, x^3)^T$ der von der Beobachter-Wahl der ONB abhängige *Koordinatenvektor*. Das Normquadrat $\|\mathbf{x}\|^2 = \langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ ist positiv definit, es verschwindet nur für den Nullvektor. In rein mathematisch motivierten Betrachtungen wird meist darauf verzichtet, die Koordinaten mit physikalischen Einheiten zu behaften.

bezüglich einer vorgängig in V_E^3 gewählten ONB repräsentiert durch eine Matrix R

$$R = \begin{pmatrix} R_1^1 & R_1^2 & R_1^3 \\ R_2^1 & R_2^2 & R_2^3 \\ R_3^1 & R_3^2 & R_3^3 \end{pmatrix}, \quad R^i_j = R^{ij} = R_{ij} = R_i^j, \quad i, j = 1, 2, 3, \quad (32)$$

wobei der linke Index i die Matricelemente in senkrechter Richtung von oben nach unten und der rechte Index j die Matricelemente von links nach rechts nummeriert. Die Höhe der Indizes spielt (hier!) keine Rolle. Die Koordinatenvektoren x, y und die entsprechenden Vektoren \mathbf{x}, \mathbf{y} transformieren dann gemäss

$$\mathbf{x}' = R^i_j x^j \mathbf{e}_i, \quad x'^i = R^i_j x^j, \quad \mathbf{y}' = R^k_l y^l \mathbf{e}_k, \quad y'^k = R^k_l y^l, \quad (33)$$

und das Skalarprodukt bleibt unter \mathbf{R} erhalten falls für beliebige $\mathbf{x}, \mathbf{y} \in V_E^3$

$$\langle \mathbf{x}', \mathbf{y}' \rangle = R^i_j x^j R^k_l y^l \langle \mathbf{e}_i, \mathbf{e}_k \rangle = R^i_j R^k_l x^j y^l \delta_{ik} = R^i_j R^i_l x^j y^l = \delta_{jl} x^j y^l = (x, y) = \langle \mathbf{x}, \mathbf{y} \rangle \quad (34)$$

erfüllt ist, d.h. es muss gelten

$$R^i_j R^i_l = \delta_{jl}, \quad \text{oder} \quad R^T R = \mathbb{1}_3 = R R^T, \quad \text{oder} \quad R^T = R^{-1}. \quad (35)$$

Eine *Symmetrie* ist in der Physik eine Transformation eines Objekts unserer Anschauung, welche Eigenschaften des Objekts erhält. Beim entsprechenden Objekt kann es sich sowohl um ein physikalisches System wie auch um eine mathematische Struktur (z.B. ein Naturgesetz) handeln. Symmetrien einer Objekteigenschaft bilden in natürlicher Weise eine (Transformations-)Gruppe.

Wir definieren daher die zur linearen Isometrie-Gruppe des dreidimensionalen euklidischen Raumes isomorphe Matrixgruppe

$$O(3) = \{R \in Mat(3, \mathbb{R}) \mid R^T R = \mathbb{1}_3\}, \quad (36)$$

welche als *orthogonale Gruppe in drei (reellen) Dimensionen* bezeichnet wird. Wegen

$$\det(R^T R) = \det(R^T) \det(R) = \det(R)^2 = \det(\mathbb{1}_3) = 1 \quad (37)$$

ist $\det(R) = \pm 1$ für $R \in O(3)$. Orthogonale Matrizen mit $\det(R) = +1$ beschreiben 'Rotationen' des euklidischen Raumes. Die entsprechende Rotationsgruppe

$$SO(3) = \{R \in Mat(3, \mathbb{R}) \mid R^T R = \mathbb{1}_3, \det(R) = 1\} \quad (38)$$

heisst *spezielle orthogonale Gruppe in drei (reellen) Dimensionen*. Die orthogonale Gruppe $O(3)$ enthält im Gegensatz zur $SO(3)$ auch 'Spiegelungen'. Die $SO(3)$ ist eng mit dem Begriff des Spins in der Quantenmechanik und der komplexen Gruppe der speziellen unitären Matrizen $SU(2)$ verknüpft.

Eine Verallgemeinerung der obigen Betrachtungen auf euklidische Vektorräume mit $\dim(V_E^n) = n \in \mathbb{N}$ ist relativ zwanglos möglich.

Passive isometrische Transformationen des euklidischen Raumes V_E^3

Unter einer *passiven Transformation* versteht man in der Physik einen Basiswechsel (Änderung des Koordinatensystems, 'Änderung des Blickwinkels'), bei welchem aber die beobachteten oder beschriebenen Objekte (im vorliegenden Fall die Vektoren) invariant bleiben.

Ersetzen wir die ONB $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ durch eine ONB' $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ gemäss $\mathbf{e}'_i = \tilde{R}_i^j \mathbf{e}_j$, so folgt für \tilde{R}

$$\delta_{ik} = \langle \mathbf{e}'_i, \mathbf{e}'_k \rangle = \tilde{R}_i^j \tilde{R}_k^j \Rightarrow \tilde{R} \tilde{R}^T = \mathbb{1}_3 = \tilde{R}^T \tilde{R} \Rightarrow \tilde{R} \in O(3). \quad (39)$$

Ein Vektor $\mathbf{x} \in V_E^3$ kann in den beiden alternativen Basen ONB und ONB' durch

$$\mathbf{x} = x^k \mathbf{e}_k = x'^i \mathbf{e}'_i = \hat{R}^i_k x^k \mathbf{e}'_i = \hat{R}^i_k x^k \tilde{R}_i^j \mathbf{e}_j = \delta_k^j x^k \mathbf{e}_j, \quad (40)$$

dargestellt werden, also kann aus $\hat{R}^i_k \tilde{R}_i^j = \delta_k^j$ oder $\hat{R}^T \tilde{R} = \mathbb{1}_3$ gefolgert werden, dass die Koordinatenvektoren mit der invers-transponierten (also kontragredienten) Basistransformationsmatrix $(\tilde{R}^T)^{-1}$ transformiert werden, wobei hier speziell gilt

$$\hat{R} = (\tilde{R}^{-1})^T = (\tilde{R}^T)^{-1} \stackrel{!}{=} \tilde{R}. \quad (41)$$

In der Relativitätstheorie mit einer nicht-euklidischen vierdimensionalen Raumzeit ist die Situation allerdings etwas komplizierter. Dort transformieren Basisvektoren *kovariant*, die Koordinatenvektoren *kontravariant*.

Minkowski-Raum (\mathbb{M}, η)

Eigenschaften

Der Minkowski-Raum (nach Hermann Minkowski, 1907) ist ein vierdimensionaler Vektorraum, mit dessen Hilfe sich die Relativitätstheorie elegant formulieren lässt. Verbreitete Bezeichnungen für den Minkowski-Raum (oder entsprechende Koordinatenvektorräume) sind \mathbb{M} , \mathbb{M}^4 oder $\mathbb{R}^{1,3}$.

Auf \mathbb{M} ist per Definition eine reelle, nicht ausgeartete Bilinearform $\eta(\cdot, \cdot)$ definiert, die aber im Gegensatz zu einem Skalarprodukt nicht positiv definit ist.

Definition 11 Sei \mathbb{M} ein vierdimensionaler reeller Vektorraum. Das Paar (\mathbb{M}, η) ist ein Minkowski-Raum, wenn die Bilinearform $\eta : \mathbb{M} \times \mathbb{M} \rightarrow \mathbb{R}$ folgende Eigenschaften aufweist:

$$(M1) \quad \eta(\lambda \mathbf{x} + \mathbf{y}, \mathbf{z}) = \lambda \eta(\mathbf{x}, \mathbf{z}) + \eta(\mathbf{y}, \mathbf{z}) \in \mathbb{R} \quad \forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{M} \text{ und } \lambda \in \mathbb{R} \quad (\text{Linearität})$$

$$(M2) \quad \eta(\mathbf{x}, \mathbf{y}) = \eta(\mathbf{y}, \mathbf{x}) \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{M} \quad (\text{Symmetrie})$$

$$(M3) \quad \eta(\mathbf{x}, \mathbf{y}) = 0 \quad \forall \mathbf{y} \in \mathbb{M} \quad \Rightarrow \quad \mathbf{x} = \mathbf{0} \quad (\eta \text{ nicht ausgeartet, perfekte Paarung})$$

$$(M4) \quad \eta \text{ hat die Signatur } (+, -, -, -)$$

(M4) sagt aus, dass η positiv definit auf einem höchstens eindimensionalen Untervektorraum von \mathbb{M} und negativ definit auf einem höchstens dreidimensionalen Untervektorraum von \mathbb{M} ist.

(\mathbb{M}, η) ist strukturell isomorph zum Koordinatenraum $\mathbb{R}_{\mathbb{R}}^4$ ($\mathbb{R}^{1,3}$) mit folgender Bilinearform $(\cdot, \cdot)_M$: Für zwei Vektoren $x, y \in \mathbb{R}_{\mathbb{R}}^4$ schreibt man unter Verwendung der Einsteinschen Summenkonvention

$$(x, y)_M := x^0 y^0 - x^1 y^1 - x^2 y^2 - x^3 y^3 = \sum_{\mu, \nu=0}^3 x^\mu g_{\mu\nu} y^\nu = g_{\mu\nu} x^\mu y^\nu = x^\mu y_\mu = x_\mu y^\mu = x^T g y, \quad (42)$$

mit

$$x = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} x_0 \\ -x_1 \\ -x_2 \\ -x_3 \end{pmatrix} = (x^0, x^1, x^2, x^3)^T, \quad y = \begin{pmatrix} y^0 \\ y^1 \\ y^2 \\ y^3 \end{pmatrix} = \begin{pmatrix} y_0 \\ -y_1 \\ -y_2 \\ -y_3 \end{pmatrix} = (y^0, y^1, y^2, y^3)^T \quad (43)$$

und dem metrischen Tensor (welcher *keine* Metrik und auch *kein* Skalarprodukt im ursprünglichen mathematischen Sinne definiert)

$$g = \begin{pmatrix} g_{00} & g_{01} & g_{02} & g_{03} \\ g_{10} & g_{11} & g_{12} & g_{13} \\ g_{20} & g_{21} & g_{22} & g_{23} \\ g_{30} & g_{31} & g_{32} & g_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \text{diag}(1, -1, -1, -1). \quad (44)$$

Es gibt in \mathbb{M} eine 'Orthonormalbasis' $\{\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ (\rightarrow Orthogonalisierungsverfahren von Erhard Schmidt) mit $\eta(\mathbf{e}_\mu, \mathbf{e}_\nu) = g_{\mu\nu}$ für $\mu, \nu = 0, 1, 2, 3$. Die $\mathbf{e}_{1,2,3}$ sind also *raumartige* Einheitsvektoren mit 'negativem Pseudo-Normquadrat' $\eta(\mathbf{e}_{1,2,3}, \mathbf{e}_{1,2,3}) = -1$, für den *zeitartigen* Einheitsvektor \mathbf{e}_0 gilt $\eta(\mathbf{e}_0, \mathbf{e}_0) = 1$.

Zwei Vektoren $\mathbf{x}, \mathbf{y} \in \mathbb{M}$ besitzen somit eine Darstellung

$$\mathbf{x} = x^\mu \mathbf{e}_\mu, \quad \mathbf{y} = y^\nu \mathbf{e}_\nu, \quad (45)$$

wobei nun die zwei Spaltenvektoren

$$x = (x^0, x^1, x^2, x^3)^T, \quad y = (y^0, y^1, y^2, y^3)^T \quad (46)$$

als (kontravariante, s.u.) *Koordinatenvektoren* von \mathbf{x} und \mathbf{y} bezüglich der Basis $\{\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ bezeichnet werden. Offensichtlich ist unter Verwendung der Linearität des Minkowski-'Skalarprodukts'

$$\eta(\mathbf{x}, \mathbf{y}) = \eta(x^\mu \mathbf{e}_\mu, y^\nu \mathbf{e}_\nu) = x^\mu y^\nu \eta(\mathbf{e}_\mu, \mathbf{e}_\nu) = x^\mu y^\nu g_{\mu\nu} = x^\mu y_\mu = x^T g y. \quad (47)$$

Elemente des Minkowski-Raumes \mathbb{M} oder Koordinatenvektoren x werden in der Physik oft als Vierervektoren bezeichnet. In Gleichung (43) wurden die sogenannten *kovarianten* Komponenten $y^* = g y$ beziehungsweise $y_\mu = g_{\mu\nu} y^\nu$ eines Koordinaten-Vierervektors y eingeführt, welche sich durch ein Vorzeichen im 'Raumanteil' des Vierervektors unterscheiden und mit einem tiefgestellten Index in der Komponentenschreibweise notiert werden. Die Herkunft der Beiworte *kovariant* und *kontravariant* wird unten klar werden.

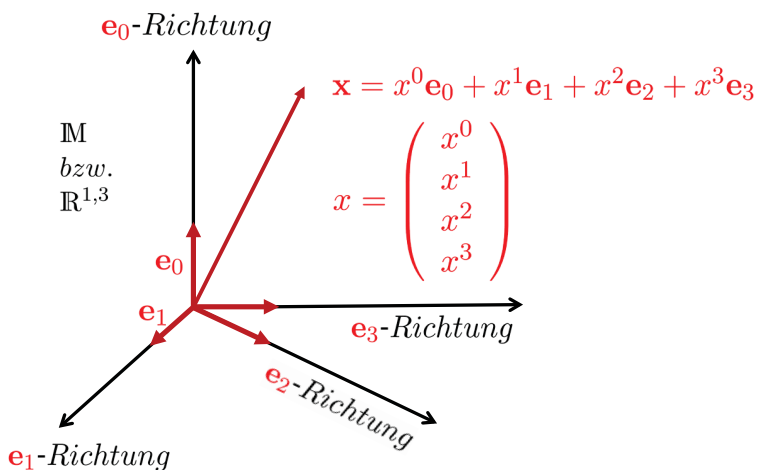


Abbildung 5: Im vierdimensionalen *pseudo*-euklidischen Minkowski-Raum \mathbb{M} wird das euklidische Skalarprodukt durch eine Bilinearform $\eta(\cdot, \cdot)$ (bzw. ein Pseudo-Skalarprodukt $\langle \cdot, \cdot \rangle_M$) ersetzt, welches zwei relativistischen Vektoren $\mathbf{x}, \mathbf{y} \in \mathbb{M}$ einen reellen Wert $\eta(\mathbf{x}, \mathbf{y})$ (bzw. $\langle \mathbf{x}, \mathbf{y} \rangle_M$) zuweist. In \mathbb{M} kann eine *Orthonormal-Basis* (ONB) $\{\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ im Sinne von $\langle \mathbf{e}_\mu, \mathbf{e}_\nu \rangle_M = g_{\mu\nu}$ gewählt werden, wobei der sogenannte metrische Tensor g durch $g_{00} = 1, g_{11} = g_{22} = g_{33} = -1$ und $g_{\mu\nu} = 0$ für $\mu \neq \nu$ gegeben ist. Jeder Vektor $\mathbf{x} \in \mathbb{M}$ kann wiederum als Linearkombination $\mathbf{x} = x^0 \mathbf{e}_0 + x^1 \mathbf{e}_1 + x^2 \mathbf{e}_2 + x^3 \mathbf{e}_3$ dargestellt werden, und $x = (x^0, x^1, x^2, x^3)^T$ ist wiederum der von der Beobachter-Wahl der ONB abhängige *Koordinatenvektor* ('Vierervektor'). \mathbf{e}_0 mit $\langle \mathbf{e}_0, \mathbf{e}_0 \rangle_M = 1$ zeigt in zeitliche Richtung, $\mathbf{e}_{1,2,3}$ sind raumartige Vektoren, deren 'Pseudo-Normquadrat' $\langle \mathbf{e}_1, \mathbf{e}_1 \rangle_M = \langle \mathbf{e}_2, \mathbf{e}_2 \rangle_M = \langle \mathbf{e}_3, \mathbf{e}_3 \rangle_M = -1$ in der in der Literatur verbreiteten Konvention negativ ist. Es ist also $\eta(\mathbf{x}, \mathbf{x}) = \langle \mathbf{x}, \mathbf{x} \rangle_M = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2$. Eine Definition von g mit umgekehrtem Vorzeichen wäre ebenso möglich.

Aktive 'isometrische' Transformationen des Minkowski-Raumes \mathbb{M}

Wir betrachten eine lineare Abbildung auf \mathbb{M} , welche die Minkowski-Bilinearform η auf allen Vektoren gemäss

$$\Lambda : \mathbb{M} \xrightarrow{\text{bijektiv}} \mathbb{M}, \quad \mathbf{x} \mapsto \mathbf{x}' = \Lambda(\mathbf{x}), \quad \eta(\Lambda(\mathbf{x}), \Lambda(\mathbf{y})) = \eta(\mathbf{x}', \mathbf{y}') = \eta(\mathbf{x}, \mathbf{y}) \quad (48)$$

erhält, bezüglich einer ONB repräsentiert durch eine Matrix Λ

$$\Lambda = \begin{pmatrix} \Lambda_0^0 & \Lambda_0^1 & \Lambda_0^2 & \Lambda_0^3 \\ \Lambda_1^0 & \Lambda_1^1 & \Lambda_1^2 & \Lambda_1^3 \\ \Lambda_2^0 & \Lambda_2^1 & \Lambda_2^2 & \Lambda_2^3 \\ \Lambda_3^0 & \Lambda_3^1 & \Lambda_3^2 & \Lambda_3^3 \end{pmatrix}. \quad (49)$$

Die Koordinatenvektoren x, y und die entsprechenden Vektoren \mathbf{x}, \mathbf{y} transformieren dann gemäss

$$\mathbf{x}' = \Lambda^\mu_\nu x^\nu \mathbf{e}_\mu, \quad x'^\mu = \Lambda^\mu_\nu x^\nu, \quad \mathbf{y}' = \Lambda^\rho_\sigma y^\sigma \mathbf{e}_\rho, \quad y'^\rho = \Lambda^\rho_\sigma y^\sigma, \quad (50)$$

als Matrizenprodukt notiert

$$x' = \Lambda x, \quad y' = \Lambda y, \quad (51)$$

und η bleibt erhalten falls für beliebige $\mathbf{x}, \mathbf{y} \in \mathbb{M}$

$$\eta(\mathbf{x}', \mathbf{y}') = \Lambda^\mu_\nu x^\nu \Lambda^\rho_\sigma y^\sigma \eta(\mathbf{e}_\mu, \mathbf{e}_\rho) = \Lambda^\mu_\nu \Lambda^\rho_\sigma x^\nu y^\sigma g_{\mu\rho} = \Lambda^\mu_\nu g_{\mu\rho} \Lambda^\rho_\sigma x^\nu y^\sigma = g_{\nu\sigma} x^\nu y^\sigma = (x, y)_M = \eta(\mathbf{x}, \mathbf{y}) \quad (52)$$

erfüllt ist, d.h. es muss gelten

$$\Lambda^\mu_\nu g_{\mu\rho} \Lambda^\rho_\sigma x^\nu y^\sigma = g_{\nu\sigma} x^\nu y^\sigma \quad \text{oder} \quad \Lambda^T g \Lambda = g. \quad (53)$$

Dasselbe Resultat folgt mit den Gleichungen (51) aus der kurzen Rechnung

$$(x', y')_M = x'^T g y' = (\Lambda x)^T g (\Lambda y) = x^T \underbrace{\Lambda^T g \Lambda}_g y = x^T g y = (x, y)_M \quad \forall x, y \in \mathbb{R}^{1,3}. \quad (54)$$

Wir definieren daher die in der Physik als *Lorentzgruppe* bezeichnete Matrixgruppe

$$\mathcal{L} = O(1, 3) = \{\Lambda \in Mat(4, \mathbb{R}) \mid \Lambda^T g \Lambda = g\} \quad (55)$$

welche in der Mathematik als pseudo-orthogonale Gruppe der Signatur $(1, 3)$ bezeichnet wird. Die Struktur dieser Gruppe muss aber noch einer eingehenderen Untersuchung unterzogen werden.

Aus $\Lambda^T g \Lambda = g$ folgt sofort $(\Lambda^T)^{-1} = g \Lambda g^{-1}$, wobei $g^{-1} = g^T = g$. Dies lässt sich mit $g_{\mu\nu} = g^{\mu\nu}$ und $g_{\mu\nu} g^{\nu\rho} = g_\mu^\rho = \delta_\mu^\rho$ auch in der Form

$$[(\Lambda^T)^{-1}]^\mu_\nu = g_{\mu\rho} \Lambda^\rho_\sigma g^{\sigma\nu} = \Lambda_\mu^\nu \quad (56)$$

schreiben, wobei das Hoch-(Herunter-)Ziehen des Index ν (μ) als rein formale Schreibweise für die rechts-seitige Matrixmultiplikation von Λ mit g^{-1} und die links-seitige Multiplikation mit g zu interpretieren ist.

Lorentzgruppe

Passive isometrische Transformationen im Minkowskiraum \mathbb{M}

Ein Vierervektor \mathbf{x} im Minkowski-Raum kann als formales Skalarprodukt des Koordinatenvektors x mit dem 'Basisvektorenvektor' $\mathbf{e} = (\mathbf{e}_0 \ \mathbf{e}_1 \ \mathbf{e}_2 \ \mathbf{e}_3)^T$

$$\mathbf{x} = x^0 \mathbf{e}_0 + x^1 \mathbf{e}_1 + x^2 \mathbf{e}_2 + x^3 \mathbf{e}_3 = (x^0 \ x^1 \ x^2 \ x^3) \begin{pmatrix} \mathbf{e}_0 \\ \mathbf{e}_1 \\ \mathbf{e}_2 \\ \mathbf{e}_3 \end{pmatrix} = x^T \mathbf{e} \quad (57)$$

geschrieben werden. In einer Orthonormalbasis $\{\mathbf{e}'_0, \mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$, in der der Koordinatenvektor von \mathbf{x} durch $x'^\mu = \Lambda^\mu_\nu x^\nu$ oder $x' = \Lambda x$ gegeben ist, lässt sich \mathbf{x} darstellen gemäss

$$\mathbf{x} = x'^T \mathbf{e}' = x'^T \Lambda^T (\Lambda^T)^{-1} \mathbf{e}, \quad (58)$$

also transformiert die Basis mit $(\Lambda^T)^{-1} = (\Lambda^{-1})^T = g \Lambda g^{-1}$, und das Transformationsverhalten der Basis wird als *kovariant*, jenes der Koordinatenvektoren als *kontravariant* bezeichnet. Natürlich hätte man in der obigen Definition der passiven Transformation die Transformationsmatrix Λ mit $(\Lambda^T)^{-1}$ vertauschen können. Da in praktischen Rechnungen oft mit Koordinatenvektoren hantiert wird, ist eine möglichst kurze Notation für die entsprechende Transformationsmatrix von Vorteil.

Die *kovarianten* Komponenten eines Koordinatenvektors transformieren gemäss

$$x'_\mu = g_{\mu\rho} \Lambda^\rho_\sigma x^\sigma = g_{\mu\rho} \Lambda^\rho_\sigma g^{\sigma\nu} x_\nu \quad \text{oder} \quad (gx') = g \Lambda g^{-1} (gx) = (\Lambda^{-1})^T (gx) \quad (59)$$

wie die Basis kovariant.

Im Gegensatz zum euklidischen Raum stimmt im Minkowski-Raum die Transformationsmatrix Λ der Koordinatenvektoren nicht mit der invers-transponierten, also der *kontragredienten* Matrix $(\Lambda^T)^{-1}$ überein.

Struktur der Lorentzgruppe

Wegen

$$\det(\Lambda^T g \Lambda) = \det(\Lambda^T) \det(g) \det(\Lambda) = \det(\Lambda)^2 \det(g) = \det g = -1 \quad (60)$$

ist $\det(\Lambda) = \pm 1$ für $\Lambda \in O(1, 3)$. Beschränkt man sich auf die Elemente in $O(1, 3)$ mit positiver Determinante, so erhält man die Untergruppe

$$\mathcal{L}_+ = SO(1, 3) = \{\Lambda \in Mat(4, \mathbb{R}) \mid \Lambda^T g \Lambda = g, \det \Lambda = 1\}, \quad (61)$$

welche *eigentliche Lorentzgruppe* genannt wird. Spalten wir eine Matrix $\Lambda \in O(1, 3)$ gemäss der Darstellung

$$\Lambda = \begin{pmatrix} \gamma & -\mathbf{a}^T \\ -\mathbf{b} & \mathbf{M} \end{pmatrix} \quad (62)$$

auf, wobei γ eine reelle Zahl, \mathbf{a} und \mathbf{b} Spaltenvektoren und \mathbf{M} eine 3×3 -Matrix sind, so folgt nach kurzer Rechnung mit Hilfe von $(\Lambda^{-1})^T = g \Lambda g^{-1}$

$$(\Lambda^{-1})^T = \begin{pmatrix} \gamma & \mathbf{a}^T \\ \mathbf{b} & \mathbf{M} \end{pmatrix}. \quad (63)$$

Zudem müssen die Relationen

$$\gamma^2 - \mathbf{b}^2 = 1, \quad \mathbf{b}^T \mathbf{M} = \gamma \mathbf{a}^T, \quad \mathbf{M}^T \mathbf{M} = \mathbf{a} \mathbf{a}^T + \mathbb{1}_3 \quad (64)$$

gelten. Daraus folgt, dass $\gamma = \Lambda_0^0$ entweder ≥ 1 oder ≤ -1 ist. Es zeigt sich, dass durch

$$\mathcal{L}_+^\uparrow = SO^+(1, 3) = \{\Lambda \in Mat(4, \mathbb{R}) \mid \Lambda^T g \Lambda = g, \det \Lambda = 1, \Lambda_0^0 \geq 1\} \quad (65)$$

wiederum eine Untergruppe der Gruppe $SO(1, 3) \subset O(1, 3)$ gegeben ist. Diese Gruppe gilt als echte Symmetriegruppe der Naturgesetze und wird *eigentliche orthochrone Lorentzgruppe* genannt. Die grösseren Gruppen $SO(1, 3)$ und $O(1, 3)$ enthalten Raum-, Zeit- und Raum-Zeit-Spiegelungen, welche tatsächlich keine Symmetriegruppen der bekannten Naturgesetze darstellen. Dennoch ist die $O(1, 3)$ für theoretische Betrachtungen wichtig.

Klassifikation der Vierervektoren

Vierektoren werden in folgende drei Klassen eingeteilt:

$$\eta(\mathbf{x}, \mathbf{x}) = x_\mu x^\mu > 0 : \mathbf{x} \text{ (oder } x) \text{ ist zeitartig.}$$

$$\eta(\mathbf{x}, \mathbf{x}) = x_\mu x^\mu < 0 : \mathbf{x} \text{ ist raumartig.}$$

$$\eta(\mathbf{x}, \mathbf{x}) = x_\mu x^\mu = 0 : \mathbf{x} \text{ ist lichtartig.}$$

Wurde in \mathbb{M} eine Basis gewählt und ist durch \mathbf{e}_0 die Zeitrichtung klar, so unterscheidet man bei *zeitartigen* Vektoren:

$$x^0 > 0 : \mathbf{x} \text{ ist zukunftsgerichtet.}$$

$$x^0 < 0 : \mathbf{x} \text{ ist vergangenheitsgerichtet.}$$

Für lichtartige Vektoren gilt schliesslich:

$$x^0 > 0 : \mathbf{x} \text{ ist zukunftsgerichtet.}$$

$$x^0 < 0 : \mathbf{x} \text{ ist vergangenheitsgerichtet.}$$

$$x^0 = 0 : \mathbf{x} \text{ ist der Nullvektor.}$$

Lichtartige Vektoren stehen auf sich selbst Minkowski-senkrecht. Die Tatsache, dass ihre Minkowski-Norm $x_\mu x^\mu$ verschwindet, bedeutet natürlich nicht, dass sie (verschwindende) Nullvektoren sind (falls $x^0 \neq 0$). Allerdings nennt man im Englischen solche Vektoren nebst 'lightlike vectors' auch 'null vectors', während der eigentliche Nullvektor mit 'zero vector' bezeichnet wird.

Alle eigentlichen orthochronen Lorentztransformationen erhalten die obigen Vektoreigenschaften, welche die Vektoren im Minkowskiraum in sechs verschiedene Lorentz-invariante Klassen (oder Mannigfaltigkeiten) einteilen.

Cauchy-Schwarzsche Ungleichung

Für zwei zeitartige zukunftsgerichtete Vektoren \mathbf{x} und \mathbf{y} gilt die Cauchy-Schwarzsche Ungleichung mit verkehrtem Vorzeichen

$$\eta(\mathbf{x}, \mathbf{y})^2 \geq \eta(\mathbf{x}, \mathbf{x})\eta(\mathbf{y}, \mathbf{y}) \quad (66)$$

Salopp schreiben wir künftig auch $xy = \eta(\mathbf{x}, \mathbf{y})$ und $x^2 = x_\mu x^\mu = \eta(\mathbf{x}, \mathbf{x})$, also gilt

$$(xy)^2 \geq x^2 y^2. \quad (67)$$

Das Gleichheitszeichen gilt genau dann, wenn \mathbf{x} und \mathbf{y} linear abhängig sind. Das Minkowski-Normquadrat ist für zeitartige Vektoren per Definition positiv, im Allgemeinen aber wie bereits erwähnt nicht definit und in diesem Sinne auch kein 'Längenquadrat'.

Minimale \mathcal{L}_+^\dagger -invariante Untermannigfaltigkeiten des Minkowski-Raumes

Figur 6 veranschaulicht die invarianten Mannigfaltigkeiten, welche entstehen, wenn alle Elemente der eigentlichen orthochronen Lorentzgruppe auf einem ausgewählten Punkt im Minkowski-Raum operieren. In vereinfachender Weise wird in der Notation nicht zwischen Vektoren $x \in \mathbb{M}$ und den zugehörigen kartesischen Koordinatenvektoren $x \in \mathbb{R}^{1,3}$ unterschieden. Oft werden Teilchenimpulse $p^\mu = (E/c = p^0 = \sqrt{\vec{p}^2 + m^2 c^2}, \vec{p})$ als Elemente eines Minkowskiraumes beschrieben, wobei die relativistische Energie-Impuls-Beziehung $c^2 p_0 p^0 = \vec{p}^2 c^2 + m^2 c^4$ beziehungsweise $p^2 = m^2$ für $c = 1$ die Bezeichnung "Massenschale" für die Mannigfaltigkeiten M_3^m und M_4^m motiviert.

- Zukunftslichtkegel: $M_1 = \{x \in \mathbb{M} \mid x^2 = 0, x^0 > 0\}$.
- Vergangenheitslichtkegel: $M_2 = \{x \in \mathbb{M} \mid x^2 = 0, x^0 < 0\}$.
- "Zukunftsmassenschale": $M_3^m = \{x \in \mathbb{M} \mid x^2 = m^2 > 0, x^0 > 0\}$.
- "Vergangenheitsmassenschale": $M_4^m = \{x \in \mathbb{M} \mid x^2 = m^2 > 0, x^0 < 0\}$.
- Raumartiges, einschaliges Hyperboloid: $M_5^m = \{x \in \mathbb{M} \mid x^2 = -m^2 < 0\}$.
- Nullpunkt: $M_6 = \{0\}$.

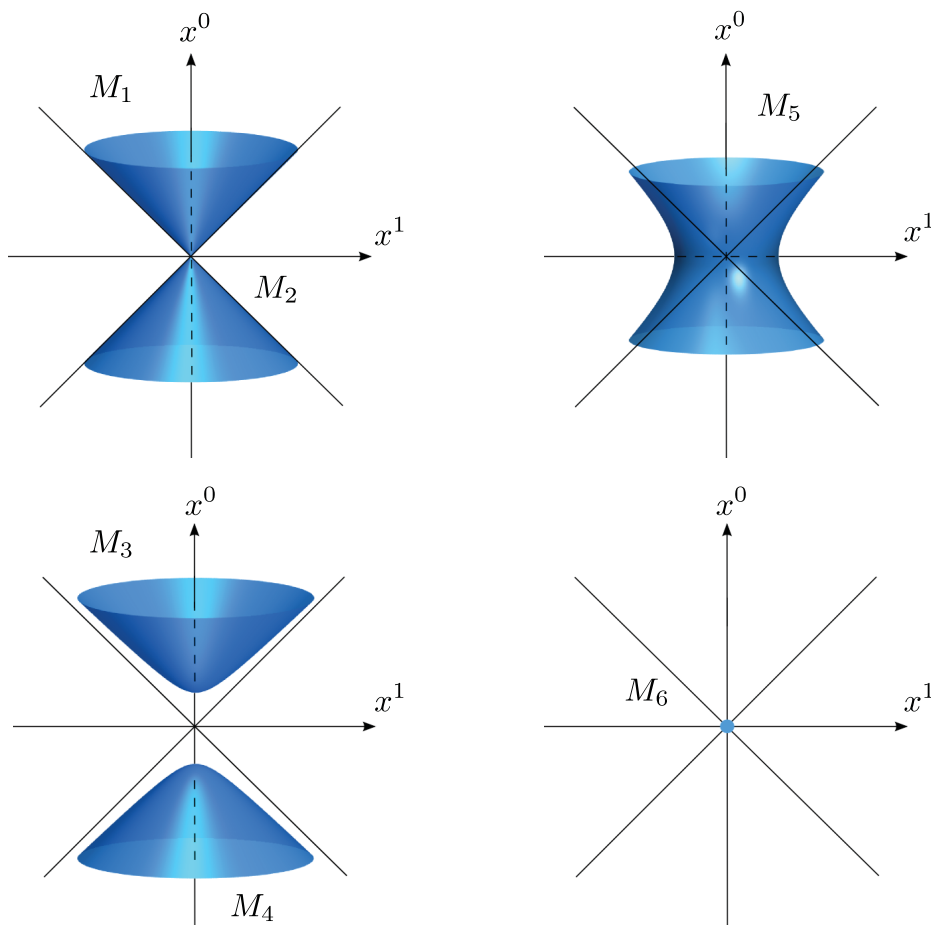


Abbildung 6: Lorentz-invariante Untermannigfaltigkeiten des Minkowski-Raumes.

Zusammenhangskomponenten und wichtige Untergruppen der Lorentzgruppe $\mathcal{L} = O(1, 3)$

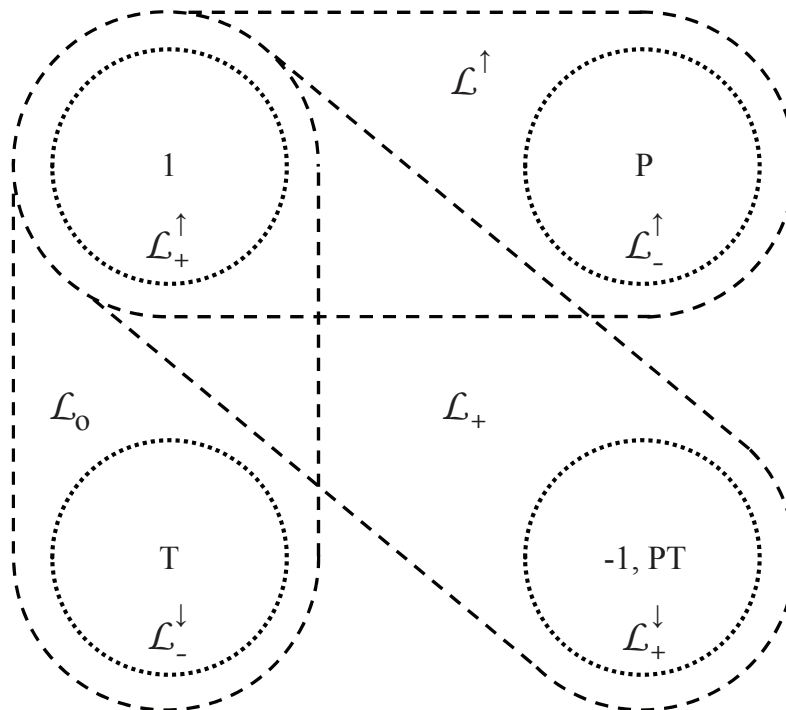


Abbildung 7: Die vier paarweise disjunkten, also nicht zusammenhängenden sowie nicht kompakten Komponenten der Lorentzgruppe $\mathcal{L} = O(1, 3)$ und entsprechende Untergruppen: Die eigentliche Lorentzgruppe $\mathcal{L}_+ = SO(1, 3)$, die orthochrone Lorentzgruppe \mathcal{L}^\uparrow , die orthochore Lorentzgruppe $\mathcal{L}_o = \mathcal{L}_+^\uparrow \cup T\mathcal{L}_+^\uparrow$ (s.u.) und die eigentliche orthochrone Lorentzgruppe $\mathcal{L}_+^\uparrow = SO^+(1, 3)$, welche die Identität enthält (Zusammenhangskomponente der Eins). Die Mengen \mathcal{L}_-^\downarrow , \mathcal{L}_-^\uparrow und \mathcal{L}_+^\downarrow sind natürlich *keine* Gruppen, da jeweils die Eins fehlt.

\mathcal{L}_+^\uparrow ist die Zusammenhangskomponente der Eins ($'1', \mathbb{1}_4$). \mathcal{L}_-^\uparrow enthält die Raumspiegelung P mit $\det P = -1$ und \mathcal{L}_-^\downarrow die Zeitumkehr T mit $\det T = -1$

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (68)$$

in \mathcal{L}_+^\downarrow ist Raumzeitpiegelung $PT = -\mathbb{1}_4$ mit $\det(PT) = 1$.

Die Transformationen $\{\mathbb{1}_4, P, T, PT\}$ bilden die Kleinsche Vierergruppe V . Diese diskrete Gruppe erhält ihre volle Bedeutung erst in der Quantentheorie der Elementarteilchenphysik im Zusammenhang mit dem *CPT*-Theorem.

Die *komplexe Lorentzgruppe* $O(4, \mathbb{C}) = \{\Lambda \in GL(4, \mathbb{C}) \mid \Lambda^T \Lambda = \mathbb{1}_4\}$ zerfällt lediglich in zwei Zusammenhangskomponenten, charakterisiert durch $\det \Lambda = \pm 1$. Die Zusammenhangskomponente der Eins ist die $SO(4, \mathbb{C})$. Man beachte, dass die Signatur des metrischen Tensors im Komplexen keine Rolle bei der abstrakten Definition der Gruppenstruktur spielt.

Darstellungen

Definitionen und Begriffe

In der Quantenmechanik werden die Zustände eines physikalischen Systems mit Hilfe von Hilbertraumvektoren dargestellt. Wird ein physikalisches System in irgendeiner Weise transformiert (z.B. rotiert), so überträgt sich diese Transformation auf die Beschreibung des Systems im Rahmen eines Hilbertraumformalismus und die entsprechenden Hilbertraum-Zustandsvektoren. Um diese Übertragung zu verstehen, ist die Kenntnis des mathematischen Begriffes der (linearen Gruppen-)Darstellung unerlässlich. Letztlich müssen in diesem Zusammenhang auch *Strahldarstellungen* untersucht werden.

Eine der häufig anzutreffenden und wichtigen Aufgaben der mathematischen Darstellungstheorie besteht in der Konstruktion aller (siehe unten: *irreduziblen*) Darstellungen von Gruppen. Für den Physiker und die Physikerin et cetera ist die Kenntnis der Darstellungstheorie der fundamentalen physikalischen Symmetriegruppen von grossem, praktischem Vorteil.

Definition 12 Eine (reelle oder komplexe) Darstellung ϑ einer Gruppe G ordnet jedem Element $s \in G$ eine bijektive lineare Abbildung $\vartheta(s) : V_{\mathbb{K}} \rightarrow V_{\mathbb{K}}$ auf einem \mathbb{K} -Vektorraum $V_{\mathbb{K}}$ mit $\mathbb{K} = \mathbb{R}$ oder \mathbb{C} zu, so dass gilt:

$$\vartheta(s) \circ \vartheta(t) = \vartheta(s \cdot t) \quad \forall s, t \in G. \quad (69)$$

\circ bezeichnet die Komposition der Abbildungen in $\text{Aut}(V_{\mathbb{K}})$, welche also der *Automorphismengruppe* von $V_{\mathbb{K}}$ angehören. Diese Gruppe unter Komposition wird oft auch mit $GL(V_{\mathbb{K}})$ bezeichnet.

Für $n = \dim(V_{\mathbb{K}}) < \infty$ kann mit jedem $\vartheta(s) \in GL(V_{\mathbb{K}})$ eine $n \times n$ Abbildungsmatrix $D(s) \in GL(n, \mathbb{K})$ bezüglich einer Basis in $V_{\mathbb{K}}$ assoziiert werden. n ist die Dimension der Abbildung. Entsprechend ist Gleichung (69) äquivalent zu folgender Bedingung für Matrizenprodukte:

$$D(s)D(t) = D(s \cdot t) \quad \forall s, t \in G. \quad (70)$$

Es ergeben sich aus der obigen Definition folgende Schlussfolgerungen:

- $D(e \cdot e) = D(e)D(e) = D(e) \Rightarrow D(e) = id|_{\mathbb{K}^n} = \mathbb{1}_n$, $\vartheta(e) = id|_{V_{\mathbb{K}}}$, e : neutrales Element in G .
- $D(s \cdot s^{-1}) = D(s)D(s^{-1}) = D(e) \Rightarrow D(s^{-1}) = D(s)^{-1}$.

Weiter definiert man folgende Begriffe:

- Eine Darstellung ϑ heisst *treu*, wenn der Darstellungshomomorphismus $\vartheta : G \rightarrow GL(V_{\mathbb{K}})$, $s \mapsto \vartheta(s)$ injektiv ist, d.h. verschiedene Gruppenelemente in G werden durch verschiedene lineare Abbildungen (oder Abbildungsmatrizen) dargestellt: $s \neq t \Rightarrow \vartheta(s) \neq \vartheta(t)$.
- Eine Darstellung $\vartheta : G \rightarrow id|_{V_{\mathbb{K}}}$ heisst *trivial*.
- Zwei Darstellungen $\vartheta_1 : G \rightarrow GL(V_{\mathbb{K}})$ und $\vartheta_2 : G \rightarrow GL(W_{\mathbb{K}})$ heissen äquivalent, falls ein fester linearer \mathbb{K} -Vektorraum-Isomorphismus (für $V_{\mathbb{K}} = W_{\mathbb{K}}$: Automorphismus) $\Sigma : V_{\mathbb{K}} \xrightarrow{\text{bijektiv}} W_{\mathbb{K}}$ existiert mit

$$\vartheta_2(a) = \Sigma \circ \vartheta_1(a) \circ \Sigma^{-1} \quad \forall a \in G. \quad (71)$$

Für Abbildungsmatrizen bedeutet dies, dass eine feste invertierbare Matrix S existiert, so dass $D_2(a) = SD_1(a)S^{-1} \forall a \in G$ gilt, wobei die $D_{1,2}$ die den Darstellungen $\vartheta_{1,2}$ entsprechenden Darstellungsmatrizen sind bezüglich fest gewählter Basen in $V_{\mathbb{K}}, W_{\mathbb{K}}$. Man schreibt oft $\vartheta_1 \simeq \vartheta_2$.

- Ein Untervektorraum $U \subseteq V_{\mathbb{K}}$ heisst G -invariant (oder ϑ -invariant), falls

$$\vartheta(a)(U) \subseteq U \quad \forall a \in G. \quad (72)$$

- Eine Darstellung heisst *irreduzibel*, falls es nur die trivial G -invarianten Unterräume $\{0\}$ und $V_{\mathbb{K}}$ gibt. Andernfalls ist die Darstellung reduzibel.
- Ist $V_{\mathbb{K}}$ eine direkte Summe von irreduziblen G -invarianten Untervektorräumen, so heisst ϑ *vollreduzibel*.
- Ist $\vartheta : G \rightarrow GL(n, \mathbb{C})$ eine Matrix-Darstellung der Gruppe G durch invertierbare komplexwertige $n \times n$ -Matrizen $G \ni s \mapsto D(s) \in GL(n, \mathbb{C})$, so definiert die Zuordnung

$$s \mapsto D^*(s) \quad \forall s \in G \quad (73)$$

die *komplex konjugierte* Darstellung ϑ^* . ϑ und ϑ^* sind im Allgemeinen nicht äquivalent.

- Ist $\vartheta : G \rightarrow GL(n, \mathbb{K})$ eine Matrix-Darstellung einer Gruppe G , so definiert die Zuordnung

$$s \mapsto D(s^{-1})^T \quad \forall s \in G \quad (74)$$

die *kontragrediente* Darstellung ϑ^K . ϑ und ϑ^K sind im Allgemeinen nicht äquivalent, im Falle vieler physikalisch relevanter Gruppen aber doch.

- Eine Darstellung $\vartheta : G \rightarrow GL(V_{\mathbb{C}})$ heisst unitär, falls auf $V_{\mathbb{C}}$ ein G -invariantes Skalarprodukt $\langle \cdot, \cdot \rangle$ existiert, so dass

$$\langle x, y \rangle = \langle \vartheta(a)(x), \vartheta(a)(y) \rangle \quad \forall a \in G, \quad x, y \in V_{\mathbb{C}}. \quad (75)$$

Unitäre Darstellungen spielen in der Quantenmechanik eine wichtige Rolle, da dort der Begriff der Unitarität eng mit der Wahrscheinlichkeitserhaltung verknüpft ist. Im Formalismus der Eichtheorien wie der Quantenchromodynamik (QCD) und Quantenelektrodynamik (QED) wird der Begriff der unitären Darstellung zur pseudo-unitären Darstellung auf sogenannten Krein-Räumen (Pseudo-Hilberträumen) verallgemeinert. Diese Verallgemeinerung hat mit der relativistischen Formulierung dieser Theorien und der damit verbundenen Einführung 'unphysikalischer' Objekte wie der Faddeev-Popov-Geistfelder und longitudinaler sowie zeitartiger masseloser Spin-1-Freiheitsgrade (von Gluonen, Photonen etc.) zu tun.

- Ist ϑ_1 eine Darstellung einer Gruppe G auf einem m -dimensionalen Vektorraum $V_{\mathbb{K}}$ mit einer Basis $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m$ und ϑ_2 eine Darstellung derselben Gruppe G auf einem n -dimensionalen Vektorraum $W_{\mathbb{K}}$ mit einer Basis $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n$, so lässt sich das *Tensorprodukt* (oder *Kronecker-Produkt*) $\vartheta = \vartheta_1 \otimes \vartheta_2$ definieren. Dazu betrachtet man zuerst die formal gebildeten $m \cdot n$ Produkte $\mathbf{e}_i \otimes \mathbf{f}_k, i = 1, \dots, m, k = 1, \dots, n$, welche eine Basis des $n \cdot m$ -dimensionalen *Tensorproduktraumes* $V_{\mathbb{K}} \otimes W_{\mathbb{K}}$ bilden. Ganz allgemein definiert man das Tensorprodukt zweier Vektoren gemäss der linearen Beziehung

$$v \otimes w = \sum_{i=1}^m \sum_{k=1}^n v_i w_k \mathbf{e}_i \otimes \mathbf{f}_k \in V_{\mathbb{K}} \otimes W_{\mathbb{K}}, \quad v = \sum_{i=1}^m v_i \mathbf{e}_i \in V_{\mathbb{K}}, \quad w = \sum_{k=1}^n w_k \mathbf{f}_k \in W_{\mathbb{K}}. \quad (76)$$

Mit $s \in G$ wirkt dann $\vartheta(s) = \vartheta_1(s) \otimes \vartheta_2(s)$ linear auf $V_{\mathbb{K}} \otimes W_{\mathbb{K}}$ gemäss

$$\mathbf{e}_i \otimes \mathbf{f}_k \xrightarrow{\vartheta(s)} \sum_{\mu=1}^m \sum_{\nu=1}^n D_{\mu i}^1(s) D_{\nu k}^2(s) \mathbf{e}_{\mu} \otimes \mathbf{f}_{\nu}, \quad \vartheta(s)(v \otimes w) = \vartheta_1(s)(v) \otimes \vartheta_2(s)(w), \quad (77)$$

wobei die $D^{1,2}$ die mit den Darstellungen $\vartheta_{1,2}$ assoziierten Matrizen bezüglich der Basen $\{\mathbf{e}_i\}_{i=1}^m$ und $\{\mathbf{f}_k\}_{k=1}^n$ sind. ϑ kann in expliziter Matrixform D dargestellt werden, wenn man eine Reihenfolge der Basisvektoren in $V_{\mathbb{K}} \otimes W_{\mathbb{K}}$ festgelegt hat. Wählt man in naheliegender Weise die Reihenfolge $\mathbf{e}_1 \otimes \mathbf{f}_1$,

$\mathbf{e}_1 \otimes \mathbf{f}_2, \dots, \mathbf{e}_1 \otimes \mathbf{f}_n, \mathbf{e}_2 \otimes \mathbf{f}_1, \mathbf{e}_2 \otimes \mathbf{f}_2, \dots, \mathbf{e}_2 \otimes \mathbf{f}_n, \dots, \mathbf{e}_m \otimes \mathbf{f}_1, \dots, \mathbf{e}_m \otimes \mathbf{f}_n$, so erhält man das $n \cdot m \times n \cdot m$ -Matrizen-Tensorprodukt (oder *Kronecker-Produkt*)

$$D(s) = D^1(s) \otimes D^2(s) = \begin{pmatrix} D_{11}^1(s)D^2(s) & D_{12}^1(s)D^2(s) & \dots & D_{1m}^1(s)D^2(s) \\ D_{21}^1(s)D^2(s) & D_{22}^1(s)D^2(s) & \dots & D_{2m}^1(s)D^2(s) \\ \vdots & \vdots & \ddots & \vdots \\ D_{m1}^1(s)D^2(s) & D_{m2}^1(s)D^2(s) & \dots & D_{mm}^1(s)D^2(s) \end{pmatrix}. \quad (78)$$

Durch Ausnützung der Linearität in der Definition von Tensorprodukten lassen sich die obigen Begriffe auch basisunabhängig und im unendlichdimensionalen Fall formulieren. Die Wahl einer basisbezogenen Darstellung hat aber einige intuitionsfördernde Vorteile.

Projektive Darstellungen und Strahldarstellungen

In der Physik spielen ausgehend von den obigen Betrachtungen auch projektive Darstellungen und - als Spezialfall derselben die Strahldarstellungen - eine wichtige Rolle:

Definition 13 Eine projektive Darstellung ρ einer Gruppe G ist ein Gruppenhomomorphismus

$$\rho : G \rightarrow \mathrm{PGL}(V_{\mathbb{K}}), \quad \mathrm{PGL}(V_{\mathbb{K}}) = \mathrm{GL}(V_{\mathbb{K}}) / \sim, \quad (79)$$

wobei die Äquivalenzrelation \sim auf $\mathrm{GL}(V_{\mathbb{K}})$ definiert ist durch

$$M_1 \sim M_2 \Leftrightarrow M_1 = \lambda \cdot M_2, \quad \lambda \in \mathbb{K}^\times. \quad (80)$$

Einem Gruppenelement $a \in G$ wird also eine Äquivalenzklasse $\rho(a)$ linearer Automorphismen des Vektorraumes $V_{\mathbb{K}}$ (oder entsprechender Abbildungsmatrizen) zugeordnet, innerhalb der sich die Elemente durch einen nicht-verschwindenden skalaren Multiplikationsfaktor $\lambda \neq 0$ unterscheiden dürfen. In der Physik ist in vielen Fällen $\mathbb{K} = \mathbb{C}$. Man nennt die durch einen 'Operator' M definierte Menge $\{\lambda M\}$ einen *Operatorstrahl*.

Manchmal wird die Menge der zulässigen λ in Gleichung (80) auf eine multiplikative Untergruppe $\mathcal{A} \subset \mathbb{C}$ der komplexen Zahlen eingeschränkt, also z.B. auf $\{1, -1\}$ oder auf die $U(1)$, der Menge der komplexen Zahlen vom Betrag Eins. Die so entstehende mehrwertige Darstellung nennt man dann *Strahldarstellung*.

Alternativ kann eine projektive Darstellung ρ auch durch eine Darstellung ϑ definiert werden, zu welcher eine skalare Funktion $\omega : G \times G \rightarrow \mathbb{K}^\times$ existiert, so dass

$$\vartheta(g_1 g_2) = \omega(g_1, g_2) \vartheta(g_1) \vartheta(g_2) \quad \forall g_1, g_2 \in G. \quad (81)$$

Man folgert sofort wegen

$$\vartheta(g_1 g_2 g_3) = \omega(g_1 g_2, g_3) \vartheta(g_1 g_2) \vartheta(g_3) = \omega(g_1, g_2 g_3) \vartheta(g_1) \vartheta(g_2 g_3), \quad (82)$$

dass ω die Konsistenz- beziehungsweise *Kohärenzbedingung*

$$\omega(g_1, g_2) \omega(g_1 g_2, g_3) = \omega(g_2, g_3) \omega(g_1, g_2 g_3) \quad \forall g_1, g_2, g_3 \in G \quad (83)$$

erfüllen muss. Im Falle einer Strahldarstellung fordert man, dass ω eine Phasenfunktion ist: $|\omega(g_1, g_2)| = 1$.

Ein einfaches Beispiel für eine Strahldarstellung erhält man, wenn man den Elementen der Kleinschen Vierergruppe - nennen wir sie hier in Anlehnung an die Raumzeitspiegelungen $1, P, T$ und PT - die verallgemeinerten Pauli-Matrizen gemäss $1 \rightarrow \sigma_0 = \mathbb{1}_2, P \rightarrow \sigma_1, T \rightarrow \sigma_2$ und $PT \rightarrow \sigma_3$ zuordnet. Diese Zuordnung respektiert Verknüpfungsstruktur der Kleinschen Vierergruppe $V \simeq C_2 \times C_2$ modulo einer Phase; so ist beispielsweise $\rho(P)\rho(T) = \sigma_1\sigma_2 = i\sigma_3 = i\rho(PT)$. Diese *echte* Strahldarstellung kann *nicht* durch eine Umzeichnung der Zuordnung mit passend gewählten Phasen in eine gewöhnliche Darstellung überführt werden.

Beispiele für Darstellungen endlicher und kontinuierlicher Gruppen

Eine reduzible reelle zweidimensionale Darstellung $D_{\dot{\mathbb{R}}}$ der multiplikativen Gruppe der reellen Zahlen ohne die Null $(\mathbb{R} \setminus \{0\}, \cdot) = (\dot{\mathbb{R}}, \cdot)$ wird durch die Zuordnung

$$\dot{\mathbb{R}} \ni x \mapsto D_{\dot{\mathbb{R}}}(x) = \begin{pmatrix} 1 & \ln|x| \\ 0 & 1 \end{pmatrix} \in GL(2, \mathbb{R}) \quad (84)$$

definiert. Man überprüft sofort die Homomorphiseigenschaft

$$xy \mapsto D_{\dot{\mathbb{R}}}(x)D_{\dot{\mathbb{R}}}(y) = \begin{pmatrix} 1 & \ln|x| \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \ln|y| \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \ln|xy| \\ 0 & 1 \end{pmatrix} = D_{\dot{\mathbb{R}}}(xy) \quad (85)$$

für $x, y \in \dot{\mathbb{R}}$. Die Darstellung (84) ist offensichtlich reduzibel, denn es existiert der unter allen Darstellungsmatrizen invariante $\mathbb{R}_{\dot{\mathbb{R}}}^2$ -Unterraum $\text{span}[(1, 0)^T]$. Die Darstellung ist allerdings nicht vollreduzibel, da der Vektor $(0, 1)^T$ keinen $D_{\dot{\mathbb{R}}}$ -invarianten Unterraum aufspannt.

Die symmetrische Gruppe S_3 ist die Gruppe aller Permutationen einer dreielementigen Menge. Nehmen wir als Menge die Zahlen $\{1, 2, 3\}$, so können die sechs Gruppenelemente der S_3 folgendermassen durch ihre permutierende Wirkung auf dieser Menge notiert werden:

$$\begin{aligned} n &: \downarrow \begin{matrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{matrix}, & a &: \downarrow \begin{matrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{matrix}, & b &: \downarrow \begin{matrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{matrix}, \\ c &: \downarrow \begin{matrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{matrix}, & d &: \downarrow \begin{matrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{matrix}, & e &: \downarrow \begin{matrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{matrix}. \end{aligned} \quad (86)$$

Die drei Elemente $\{n, a, b\}$ bilden die alternierende (Unter-)Gruppe $A_3 \subset S_3$. Offensichtlich ist die S_3 nicht abelsch. So gilt zum Beispiel $ac \neq ca$, was auf einfache Weise eingesehen werden kann, wenn man die Wirkung von ac und ca auf die 1 betrachtet. Es ist $(ac)(1) = 2 \neq (ca)(1) = 3$,

$$2 \xleftarrow{a} 1 \xleftarrow{c} 1, \quad 3 \xleftarrow{c} 2 \xleftarrow{a} 1. \quad (87)$$

Tatsächlich gilt $ac = e$ und $ca = d$.

Eine dreidimensionale reelle Darstellung der S_3 konstruiert man, indem die Permutationen der S_3 auf die drei kanonischen Basisvektoren des $\mathbb{R}_{\mathbb{R}}^3$ gemäss

$$n \mapsto \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad a \mapsto \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad b \mapsto \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad \dots \quad (88)$$

angewendet werden. Die Gruppenverknüpfung überträgt sich in natürlicher Weise auf die Matrixmultiplikation, z.B. ist

$$D(a)D(c) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = D(ac) = D(e). \quad (89)$$

Die Darstellung D ist aber reduzibel, da ein unter allen Darstellungsmatrizen $D(g)$, $g \in S_3$ invarianter eindimensionaler Unterraum existiert, welcher durch den Vektor

$$v = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad (90)$$

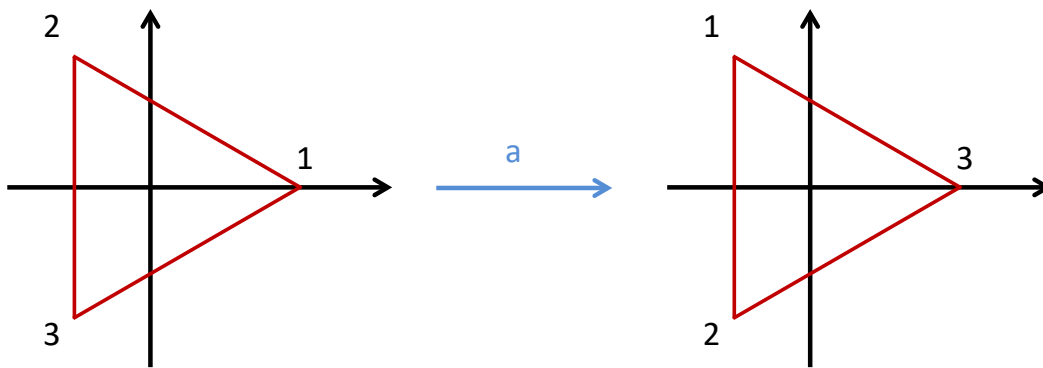


Abbildung 8: Die Permutation der Dreiecksecken $123 \rightarrow 231$ unter der Permutation a entspricht einer Drehung um $120^\circ = 2\pi/3$.

aufgespannt wird. Eine weitere zweidimensionale (reelle oder komplexe) Darstellung der S_3 erhält man, indem man die Permutationen auf die Ecken eines in der zweidimensionalen euklidischen Ebene liegenden Dreiecks anwendet. Beispielsweise wirkt a dann gemäss Abbildung 8 als Rotation in der Ebene und wird durch die Matrix

$$D(a) = \begin{pmatrix} \cos 120^\circ & -\sin 120^\circ \\ \sin 120^\circ & \cos 120^\circ \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \in SO(2) \quad (91)$$

dargestellt, c spiegelt das Dreieck durch transponieren der Ecken 2 und 3 und führt so auf die Spiegelungsmatrix

$$D(c) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \in O^-(2) = O(2) \setminus SO(2). \quad (92)$$

Die so konstruierte Darstellung ist als *reelle* und als *komplexe* Darstellung irreduzibel. Allerdings gilt dies für die Untergruppe $A_3 = \{n, a, a^2\}$ nur im reellen Fall, denn im Komplexen sind die beiden Vektoren

$$v_{1,2} = \begin{pmatrix} 1 \\ \pm i \end{pmatrix} \quad (93)$$

Eigenvektoren aller Darstellungsmatrizen in A_3 und spannen so zwei invariante eindimensionale Darstellungsräume auf. Es gilt nämlich

$$D(a)v_{1,2} = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 \\ \pm i \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \mp \frac{\sqrt{3}}{2}i \\ \frac{\sqrt{3}}{2} \mp \frac{1}{2}i \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \mp \frac{\sqrt{3}}{2}i \\ \frac{\sqrt{3}}{2} \mp \frac{1}{2}i \end{pmatrix} = \lambda_{1,2}v_{1,2} \quad (94)$$

und

$$D(b)v_{1,2} = D(a)^2v_{1,2} = \lambda_{1,2}^2v_{1,2}. \quad (95)$$

Fundamental-(Strahl-)Darstellung der Gruppen $SU(2)$ und $SO(3)$

Konstruktion der Darstellungen

Bilden wir jede Matrix aus der $SO(3)$ trivial auf die Identität in der multiplikativen Gruppe der komplexen invertierbaren 1×1 -Matrizen $GL(1, \mathbb{C}) = \dot{\mathbb{C}}$ ab, so haben wir damit alle eindimensionalen Darstellungen der $SO(3)$ konstruiert. Um die niedrigst-dimensionale nicht-triviale (Strahl-)Darstellung der $SO(3)$ in zwei komplexen Dimensionen zu konstruieren, verwenden wir die Pauli-Matrizen

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (96)$$

und bilden einen beliebigen (ONB-Koordinaten-)Vektor \mathbf{x} (x) mit den Komponenten $x_j = (x_1, x_2, x_3)$ durch die lineare Vorschrift $x \mapsto \underline{x}$ (in diesem Abschnitt gelte noch euklidisch $x_j = x^j$ und $x_j y_j = x_1 y_1 + x_2 y_2 + x_3 y_3$)

$$\underline{x} = \sigma_j x_j = \vec{\sigma} \cdot \vec{x} = \begin{pmatrix} x_3 & x_1 - ix_2 \\ x_1 + ix_2 & -x_3 \end{pmatrix} \in Herm(2, \mathbb{C})_{\text{tr}=0} \quad (97)$$

bijektiv auf die Menge der spurlosen hermiteschen komplexen 2×2 -Matrizen ab. Die Umkehrung dieser Beziehung ergibt sich in einfacher Weise aus der Spur-Identität ($\text{tr} = \text{trace}$, Spur)

$$\frac{1}{2} \text{tr}(\sigma_j \sigma_k) = \delta_{jk}. \quad (98)$$

Dann ist

$$\frac{1}{2} \text{tr}(\underline{x} \sigma_k) = \frac{1}{2} \text{tr}(x_j \sigma_j \sigma_k) = x_j \delta_{jk} = x_k = x^k. \quad (99)$$

Das Standard-Normquadrat $\|\mathbf{x}\|^2 = \langle \mathbf{x}, \mathbf{x} \rangle = (x, x) = x^2 = x_j x^j$ im \mathbb{R}^3 kann mit Hilfe der Determinante umgeschrieben werden (hier mit $i^2 = -1$) gemäss

$$\det \underline{x} = -(x_3)^2 - (x_1 + ix_2)(x_1 - ix_2) = -(x_1)^2 - (x_2)^2 - (x_3)^2 = -(x, x) = -x^2, \quad (100)$$

ausserdem gilt die Polarisationsformel⁵

$$(x + y, x + y) = x^2 + y^2 + 2(x, y) \quad (101)$$

und somit ist das Skalarprodukt

$$\langle \mathbf{x}, \mathbf{y} \rangle = (x, y) = \frac{1}{2} (\det(\underline{x}) + \det(\underline{y}) - \det(\underline{x} + \underline{y})) \quad (102)$$

und es gilt auch

$$\langle \mathbf{x}, \mathbf{y} \rangle = \frac{1}{2} \text{tr}(\underline{x} \underline{y}). \quad (103)$$

Die *spezielle unitäre Gruppe* $SU(2)$ ist definiert durch

$$SU(2) = \{U \in GL(2, \mathbb{C}) \mid U^{-1} = U^+, \det U = +1\}. \quad (104)$$

Der Trick besteht nun darin, $U \in SU(2)$ auf $Herm(2, \mathbb{C})_{\text{tr}=0}$ durch

$$\underline{x} \mapsto \underline{x}' = U \underline{x} U^+ = U \underline{x} U^{-1}, \quad (105)$$

⁵Im Komplexen führt das durch eine im ersten Argument semilineare hermitesche Sesquilinearform $(,) : V_{\mathbb{C}} \times V_{\mathbb{C}} \rightarrow \mathbb{C}$ auf einem Vektorraum $V_{\mathbb{C}}$ definierte Normquadrat $\|v\|^2 = (v, v)$ durch die komplexe Polarisationsformel $(v, w) = \frac{1}{4} \sum_{n=0}^3 (-i)^n \|v + i^n w\|^2$ wieder auf die ursprüngliche Form zurück.

operieren zu lassen, wobei $+$ hier die hermitesche Konjugation bezeichne. Offensichtlich ist \underline{x}' wiederum hermitisch und spurlos, und die Norm und das Skalarprodukt sind erhalten, z.B.

$$-x'^2 = \det \underline{x}' = \det (U \underline{x} U^+) = \det U \det \underline{x} \det U^+ = \det \underline{x} = -x^2. \quad (106)$$

Offensichtlich kann \underline{x}' wiederum in der Form

$$\underline{x}' = \sigma_j x'_j, \quad \text{mit} \quad x'^2 = x^2 \quad (107)$$

geschrieben werden, und es ist explizit

$$x'_k = \frac{1}{2} \text{tr} (\underline{x}' \sigma_k) = \frac{1}{2} \text{tr} (U x_j \sigma_j U^{-1} \sigma_k) = \frac{1}{2} \text{tr} (U \sigma_j U^+ \sigma_k) x_j = R_{kj} x_j. \quad (108)$$

Offensichtlich sind x und x' durch eine Isometrie $R \in O(3)$ verknüpft, und tatsächlich ist sogar $R \in SO(3)$, was zu zeigen hier als Übung vorgeschlagen sei (die Gruppenmannigfaltigkeit der $SO(3)$ sowie der $SU(2)$ sind zusammenhängend). Allerdings vermitteln U und $-U$ dieselbe Rotation, denn

$$U \underline{x} U^{-1} = (-U) \underline{x} (-U)^{-1}. \quad (109)$$

Tatsächlich besteht der Gruppenisomorphismus ($\mathbb{Z}_2 = \{\mathbb{1}_2, -\mathbb{1}_2\}$)

$$SU(2)/\mathbb{Z}_2 \cong SO(3). \quad (110)$$

Durch den Trick gemäss Gleichung (105) wurde also eine reelle Darstellung der $SU(2)$ durch die $SO(3)$ konstruiert, welche eben die Darstellungseigenschaft

$$R(U_2)R(U_1) = R(U_2U_1). \quad (111)$$

wegen $U_2U_1 \underline{x} U_1^+ U_2^+ = (U_2U_1) \underline{x} (U_2U_1)^+$ erfüllt. Bis auf ein Vorzeichen kann diese Beziehung umgekehrt werden, und in einem saloppen Stil kann dann $U(R_2R_1) = \pm U(R_2)U(R_1)$ geschrieben werden, was die Tatsache beschreibt, dass wir zugleich eine zweiwertige *Strahl-Darstellung* der Rotationsgruppe $SO(3)$ durch $SU(2)$ -Matrizen (oder eine Darstellung der $SU(2)$ durch sich selbst) gefunden haben.

Man mag sich fragen, ob eine solche Darstellung zu ihrer komplex konjugierten Darstellung äquivalent ist, bei welcher einem U (oder R) jeweils die komplex konjugierte Matrix U^* (oder $\pm U^*(R)$) zugeordnet wird. Tatsächlich existiert eine Matrix $\epsilon = i\sigma_2$, der antisymmetrische Tensor in zwei Dimensionen, so dass für alle $U \in SU(2)$

$$U^* = \epsilon U \epsilon^{-1}. \quad (112)$$

Eine spezielle unitäre Matrix $U \in SU(2)$ kann nämlich in der Form

$$U = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}, \quad \det U = aa^* + bb^* = 1, \quad (113)$$

geschrieben werden, und eine kurze Rechnung zeigt, dass für alle $U \in SU(2)$ offensichtlich Gleichung (112) gilt. Damit ist die Strahldarstellung der Rotationsgruppe $SO(3)$ durch die $SU(2)$ äquivalent zu ihrer komplex konjugierten Darstellung.

Eine $SU(2)$ -Matrix U kann übrigens auch in der eleganten Form

$$U = \sigma_0 q_0 + i\sigma_j q_j = \sigma_0 q_0 + i\vec{\sigma} \cdot \vec{q} \begin{pmatrix} q_0 + iq_3 & q_2 + iq_1 \\ -q_2 + iq_1 & q_0 - iq_3 \end{pmatrix} \quad \text{mit} \quad \sum_{j=0}^3 (q_j)^2 = (q_0)^2 + \vec{q}^2 = 1 \quad (114)$$

geschrieben werden.

Die Umkehrung von Gleichung (108) ist gegeben durch

$$U = \pm \frac{\mathbb{1}_2 + R_{kj} \sigma_k \sigma_j}{2\sqrt{1 + \text{tr} R}}. \quad (115)$$

Diese Formel entartet allerdings für exakte 180° -Drehungen zu einem Ausdruck $0/0$.

Eigenschaften der Pauli-Matrizen σ_1, σ_2 und σ_3

- Definition:

$$\sigma_j = \begin{pmatrix} \delta_{j3} & \delta_{j1} - i\delta_{j2} \\ \delta_{j1} + i\delta_{j2} & -\delta_{j3} \end{pmatrix}, \quad i^2 = -1. \quad (116)$$

- Involution:

$$\sigma_j^2 = -i\sigma_1\sigma_2\sigma_3 = \mathbb{1}_2 =: \sigma_0, \quad \sigma_j^{-1} = \sigma_j. \quad (117)$$

- Hermitezität:

$$\sigma_j^\dagger = \sigma_j. \quad (118)$$

- Transposition:

$$\sigma_j^T = (-1)^{j+1} \sigma_j. \quad (119)$$

- Invarianten:

$$\det \sigma_j = -1, \quad \text{tr} \sigma_j = 0. \quad (120)$$

- Eigenwerte:

$$EW(\sigma_j) = \{\pm 1\}. \quad (121)$$

- Produkte:

$$\sigma_l \sigma_m = \delta_{lm} \sigma_0 + i \epsilon_{lmn} \sigma_n, \quad \epsilon_{lmn} = \frac{1}{2}(l-m)(m-n)(n-l). \quad (122)$$

- Kommutatoren:

$$[\sigma_l, \sigma_m] = 2i \epsilon_{lmn} \sigma_n. \quad (123)$$

- Antikommutatoren:

$$\{\sigma_l, \sigma_m\} = 2\delta_{lm} \sigma_0. \quad (124)$$

- Ähnlichkeitstransformationen:

$$\sigma_j \sigma_k \sigma_j^{-1} = \sigma_j \sigma_k \sigma_j = -\sigma_k, \quad j \neq k. \quad (125)$$

- Kreuzprodukt:

$$(\vec{a} \cdot \vec{\sigma})(\vec{b} \cdot \vec{\sigma}) = (\vec{a} \cdot \vec{b}) \sigma_0 + i(\vec{a} \times \vec{b}) \cdot \vec{\sigma}. \quad (126)$$

- Potenzen:

$$(\hat{n} \cdot \vec{\sigma})^{2n} = \sigma_0, \quad (\hat{n} \cdot \vec{\sigma})^{2n+1} = (\hat{n} \cdot \vec{\sigma}), \quad \hat{n} = \frac{\vec{n}}{|\vec{n}|} \in \mathbb{R}^3. \quad (127)$$

- Exponentialabbildung:

$$e^{i\alpha(\hat{n} \cdot \vec{\sigma})} = \sigma_0 \cos \alpha + i(\hat{n} \cdot \vec{\sigma}) \sin \alpha. \quad (128)$$

- Spurformeln:

$$\text{tr}(\sigma_l \sigma_m \sigma_n) = 2i \epsilon_{lmn}, \quad \text{tr}(\sigma_k \sigma_l \sigma_m \sigma_n) = 2(\delta_{kl} \delta_{mn} - \delta_{km} \delta_{ln} + \delta_{kn} \delta_{lm}). \quad (129)$$

- Die vier generalisierten Pauli-Matrizen $\sigma_{0,1,2,3}$ bilden eine Basis des reellen Vektorraums der hermiteschen komplexen 2×2 -Matrizen und eine Basis des komplexen Vektorraums $Mat(2, \mathbb{C})$ der komplexen 2×2 -Matrizen.

Fundamental-(Strahl-)Darstellungen der eigentlichen orthochronen Lorentzgruppe $SO^+(1, 3)$

Explizite Konstruktion der zweidimensionalen Fundamental(strahl)darstellungen

Um die niedrigst-dimensionalen nicht-trivialen Darstellungen der eigentlichen orthochronen Lorentzgruppe zu konstruieren, führen wir die relativistisch verallgemeinerten Pauli-Matrizen

$$\sigma_\mu = \bar{\sigma}^\mu = (\mathbb{1}_2, \vec{\sigma}) = (\sigma_0, \sigma_1, \sigma_2, \sigma_3), \quad \bar{\sigma}_\mu = \sigma^\mu = (\mathbb{1}_2, -\vec{\sigma}) \quad (130)$$

ein, wobei die drei Komponenten in $\vec{\sigma}$ durch die Pauli-Matrizen gegeben sind. Das für die Einheitsmatrix in zwei Dimensionen $\mathbb{1}_2$ verwendete Symbol σ_0 fügt sich für das Auge angenehm in die Notation ein.

Beliebigen Vierervektoren \mathbf{x} mit kontravarianten Komponenten $x^\mu = (x^0, x^1, x^2, x^3)$ werden durch die lineare Bijektion $x \mapsto \underline{x}$

$$Herm(2, \mathbb{C}) \ni \underline{x} = \sigma_\mu x^\mu = \begin{pmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{pmatrix} \quad (131)$$

hermitesche 2×2 -Matrizen zugeordnet. Die Umkehrung dieser Beziehung ergibt sich in einfacher Weise aus der Verallgemeinerung der Spur-Identität $\frac{1}{2} \text{tr}(\sigma_i \sigma_j) = \delta_{ij}$ (tr = trace, Spur), nämlich

$$\frac{1}{2} \text{tr}(\sigma_\mu \bar{\sigma}_\nu) = g_{\mu\nu}. \quad (132)$$

Dann ist

$$\frac{1}{2} \text{tr}(\underline{x} \bar{\sigma}^\mu) = \frac{1}{2} \text{tr}(x^\nu \sigma_\nu \bar{\sigma}^\mu) = x^\nu g_\nu^\mu = x^\mu. \quad (133)$$

Das Minkowski-Skalarprodukt überträgt sich von \mathbb{M} auf $Herm(2, \mathbb{C})$ gemäss

$$\det \underline{x} = (x^0 + x^3)(x^0 - x^3) - (x^1 + ix^2)(x^1 - ix^2) = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = x_\mu x^\mu = x^2, \quad (134)$$

also

$$\eta(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) = (x + y, x + y)_M = (x + y)_\mu (x + y)^\mu = x^2 + y^2 + 2x_\mu y^\mu \quad (135)$$

und somit

$$\eta(\mathbf{x}, \mathbf{y}) = x_\mu y^\mu = \frac{1}{2} (\det(\underline{x} + \underline{y}) - \det(\underline{x}) - \det(\underline{y})). \quad (136)$$

Ebenso lässt sich mit $\bar{y} = \bar{\sigma}_\mu y^\mu$ elegant schreiben

$$\eta(\mathbf{x}, \mathbf{y}) = \frac{1}{2} \text{tr}(\underline{x} \bar{y}). \quad (137)$$

Die *spezielle lineare Gruppe* $SL(2, \mathbb{C})$ in zwei komplexen Dimensionen ist definiert durch

$$SL(2, \mathbb{C}) = \{A \in GL(2, \mathbb{C}) \mid \det A = +1\}. \quad (138)$$

Der Trick besteht darin, $A \in SL(2, \mathbb{C})$ auf $\underline{x} \in Herm(2, \mathbb{C})$ durch

$$\underline{x} \mapsto \underline{x}' = A \underline{x} A^+ \quad (139)$$

operieren zu lassen, wobei $^+$ hier die hermitesche Konjugation bezeichne. Offensichtlich ist \underline{x}' wiederum hermitisch, und das Minkowski-Skalarprodukt ist im folgenden Sinne erhalten

$$\det \underline{x}' = \det(A \underline{x} A^+) = \det A \det \underline{x} \det A^+ = \det \underline{x}. \quad (140)$$

\underline{x}' kann daher wiederum in der Form

$$\underline{x}' = \sigma_\mu x'^\mu \quad \text{mit} \quad x'_\mu x'^\mu = x_\mu x^\mu \quad (141)$$

geschrieben werden, und es ist explizit

$$x'^\mu = \frac{1}{2} \text{tr}(\underline{x}' \bar{\sigma}^\mu) = \frac{1}{2} \text{tr}(A x^\nu \sigma_\nu A^+ \bar{\sigma}^\mu) = \frac{1}{2} \text{tr}(A \sigma_\nu A^+ \bar{\sigma}^\mu) x^\nu = \Lambda^\mu_\nu x^\nu. \quad (142)$$

Offensichtlich sind \mathbf{x} und \mathbf{x}' beziehungsweise x und x' durch eine Lorentztransformation $\Lambda \in O(1, 3)$ verknüpft, und tatsächlich ist sogar $\Lambda \in SO^+(1, 3)$, was zu zeigen hier als Übung vorgeschlagen sei. Tatsächlich ist $SL(2, \mathbb{C})$ zusammenhängend und die Abbildung $A \mapsto \Lambda(A)$ stetig, die Abbildung $\lambda : A \rightarrow \Lambda(A)$ ist also ein Homomorphismus der $SL(2, \mathbb{C})$ in die eigentliche orthochrone Lorentzgruppe $\mathcal{L}_+^\uparrow = SO^+(1, 3)$. $SL(2, \mathbb{C})$ ist die (doppelte) universelle Überlagerungsgruppe der $SO^+(1, 3)$.

Bildet man übrigens die kovarianten Komponenten x_μ gemäss $\bar{x} = \bar{\sigma}^\mu x_\mu = \bar{\sigma}_\mu x^\mu$ in $Herm(2, \mathbb{C})$ ab, so lautet das Gleichung (139) entsprechende Transformationsgesetz $\bar{x}' = (A^+)^{-1} \bar{x} A^{-1}$. Dann ist ja auch

$$x'^\mu y'_\mu = \frac{1}{2} \text{tr}(A \underline{x} A^+ (A^+)^{-1} \bar{y} A^{-1}) = \frac{1}{2} \text{tr}(A^{-1} A \underline{x} \bar{y}) = \frac{1}{2} \text{tr} \underline{x} \bar{y} = x^\mu y_\mu. \quad (143)$$

Wir wollen uns davon überzeugen, dass dieser Homomorphismus λ Zwei-zu-Eins ist. Offensichtlich vermitteln ja sowohl A wie auch $-A$ dieselbe Lorentztransformation. Um dies im Detail zu verstehen, berechnen wir den Kern von λ , also die Menge aller $A \in SL(2, \mathbb{C})$ welche für jede hermitesche Matrix \underline{x} die Gleichung

$$\underline{x} = A \underline{x} A^+ \quad (144)$$

erfüllen. Wählt man speziell

$$\underline{x} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

so führt dies zur Bedingung $A = (A^+)^{-1}$, und Gleichung (144) reduziert sich auf $\underline{x} A - A \underline{x} = [\underline{x}, A] = 0$ für jedes hermitesche \underline{x} . Dies impliziert $A = \alpha \mathbb{1}_2$, $\alpha \in \mathbb{C}$. Aus der Bedingung $\det A = +1$ folgt schliesslich $A = \pm \mathbb{1}_2$.

Abgesehen vom soeben gefundenen wichtigen Gruppenisomorphismus

$$\mathcal{L}_+^\uparrow \cong SL(2, \mathbb{C}) / \{\pm 1\}, \quad \underline{\Lambda}(\pm A) \underline{x} = A \underline{x} A^+ \quad (145)$$

haben Matrizen in der speziellen linearen Gruppe $SL(2, \mathbb{C})$ eine weitere interessante Eigenschaft. Definiert man den vollständig antisymmetrischen Tensor ϵ in zwei Dimensionen

$$\epsilon = i\sigma_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \epsilon = -\epsilon^{-1} = -\epsilon^T, \quad (146)$$

so kann man zugleich eine *symplektische*, also schiefe Bilinearform $\langle u, v \rangle = -\langle v, u \rangle$ für je zwei Elemente im zwei-dimensionalen komplexen Vektorraum \mathbb{C}_c^2 (dem sogenannten Spinorraum)

$$u = \begin{pmatrix} u^1 \\ u^2 \end{pmatrix}, \quad v = \begin{pmatrix} v^1 \\ v^2 \end{pmatrix} \quad (147)$$

gemäss

$$\langle u, v \rangle = u^1 v^2 - u^2 v^1 = u^T \epsilon v \quad (148)$$

definieren. Diese symplektische Form ist wiederum invariant unter der Gruppe $SL(2, \mathbb{C})$

$$\langle u, v \rangle = u^T \epsilon v = \langle Au, Av \rangle = u^T A^T \epsilon Av. \quad (149)$$

Setzt man nämlich

$$A = \begin{pmatrix} a_1^1 & a_1^2 \\ a_2^1 & a_2^2 \end{pmatrix}, \quad \text{mit} \quad \det A = a_1^1 a_2^2 - a_1^2 a_2^1, \quad (150)$$

so zeigt eine kurze Rechnung

$$\begin{aligned} A^T \epsilon A &= \begin{pmatrix} a_1^1 & a_1^2 \\ a_2^1 & a_2^2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} a_1^1 & a_1^2 \\ a_2^1 & a_2^2 \end{pmatrix} = \\ &= \begin{pmatrix} a_1^1 & a_2^1 \\ a_1^2 & a_2^2 \end{pmatrix} \begin{pmatrix} a_2^1 & a_2^2 \\ -a_1^1 & -a_1^2 \end{pmatrix} = \begin{pmatrix} 0 & \det A \\ -\det A & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \end{aligned} \quad (151)$$

und tatsächlich besteht ein weiterer Gruppenisomorphismus

$$\mathcal{L}_+^\uparrow \cong Sp(2, \mathbb{C}) / \{\pm 1\}, \quad SL(2, \mathbb{C}) \cong Sp(2, \mathbb{C}) \quad (152)$$

mit der *komplexen symplektischen Gruppe* $Sp(2, \mathbb{C})$

$$Sp(2, \mathbb{C}) = \{A \in GL(2, \mathbb{C}) \mid A^T \epsilon A = \epsilon\}. \quad (153)$$

Abschliessend sei hier auch noch ohne Herleitung der Isomorphismus $SO(3, \mathbb{C}) \cong \mathcal{L}_+^\uparrow$ erwähnt.

Zusammenfassend kann also gesagt werden, dass durch den Darstellungs-Trick in Gleichung (139) eine reelle Darstellung der $SL(2, \mathbb{C})$ durch die $SO^+(1, 3)$ konstruiert wurde, welche eben die Darstellungseigenschaft

$$\Lambda(A_1)\Lambda(A_2) = \Lambda(A_1 A_2). \quad (154)$$

erfüllt. Bis auf ein Vorzeichen kann diese Beziehung umgekehrt werden, und in ungezwungenem Stil kann dann $A(\Lambda_1 \Lambda_2) = \pm A(\Lambda_1)A(\Lambda_2)$ geschrieben werden, was die Tatsache beschreibt, dass wir zugleich eine zweiwertige *Strahldarstellung* der Lorentzgruppe \mathcal{L}_+^\uparrow gefunden haben.

Ohne Herleitung sei hier auch die Umkehrung von Gleichung (142) notiert, welche aus topologischen Gründen für gewisse Λ versagt:

$$A = \pm \frac{\Lambda_\nu^\mu \sigma_\mu \bar{\sigma}^\nu}{\sqrt{\det(\Lambda_\nu^\mu \sigma_\mu \bar{\sigma}^\nu)}}. \quad (155)$$

Wie im Falle der $SU(2)$ mag man sich fragen, ob die obige zweiwertige Darstellung der eigentlichen orthochronen Lorentzgruppe beziehungsweise die Darstellung der $SL(2, \mathbb{C})$ durch sich selbst zu ihrer komplex konjugierten Darstellung äquivalent ist. Tatsächlich existiert keine durch eine Matrix B vermittelte Basistransformation, so dass für alle $A \in SL(2, \mathbb{C})$

$$A^* = BAB^{-1} \quad (156)$$

gilt. Betrachtet man allerdings lediglich die Untergruppe

$$SU(2) = \{U \in GL(2, \mathbb{C}) \mid U^+ = U^{-1}, \det U = 1\} \subset SL(2, \mathbb{C}), \quad (157)$$

so ist die Situation anders, da für eine spezielle unitäre Matrix $U \in SU(2)$

$$U = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix} = (U^{-1})^+, \quad \det U = aa^* + bb^* = 1 \quad (158)$$

offensichtlich mit dem antisymmetrischen Tensor in zwei Dimensionen $\epsilon = i\sigma_2$ gilt

$$\epsilon U \epsilon^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} b & -a \\ a^* & b^* \end{pmatrix} = \begin{pmatrix} a^* & b^* \\ -b & a \end{pmatrix} = U^*. \quad (159)$$

Damit ist die Strahldarstellung der Rotationsgruppe $SO(3)$ durch die $SU(2)$, welche durch Einschränkung der $SL(2, \mathbb{C})$ auf die $SU(2)$ in Gleichung (139) gewonnen werden kann, äquivalent zu ihrer komplex konjugierten Darstellung. Die $SU(2)$ -Matrizen in der $SL(2, \mathbb{C})$ erzeugen übrigens Rotationen in der Raumzeit; hermitesche Matrizen in der $SL(2, \mathbb{C})$ generieren die Boosts.

Konstruktion aller endlichdimensionalen irreduziblen komplexen Strahldarstellungen der Gruppen $SU(2)$ und $SL(2, \mathbb{C})$

“Die Physik ist für die Physiker eigentlich viel zu schwer.”

David Hilbert (1862-1943).

Verpflanzungstrick

Die Konstruktion aller endlichdimensionalen irreduziblen Darstellungen der $SU(2)$ und $SL(2, \mathbb{C})$ als universelle Überlagerungsgruppen der $SO(3)$ und $SO^+(1, 3)$ soll nun mit dem mathematischen Trick der Verpflanzung von Funktionen vorgeführt werden. Um der Kürze Willen wollen wir hier nicht zeigen, dass die in der Folge verwendete Methode tatsächlich alle gewünschten Darstellungen in erschöpfender Weise erzeugt.

Man betrachte also einen linearen komplexen Funktionen-Vektorraum F^m von vorerst nicht weiter spezifizierten Funktionen $f(x) : \mathbb{C}^m \rightarrow \mathbb{C}$ und definiere die *Verpflanzung* $\vartheta_U : F^m \rightarrow F^m$ einer dieser Funktionen durch ein Element U in einer komplexen Matrix-Gruppe $G \subseteq GL(m, \mathbb{C})$ durch

$$\vartheta_U : f(x) \mapsto g(x) = (\vartheta_U(f))(x) = f(U^{-1}x). \quad (160)$$

Damit ist auf F^m tatsächlich eine Darstellung der Gruppe G definiert, wovon man sich unmittelbar überzeugen kann. Beispielsweise gilt

$$\vartheta_{U_2 U_1} : f(x) \mapsto f((U_2 U_1)^{-1}x) = f(U_1^{-1}(U_2^{-1}x)), \quad (161)$$

die zentrale Darstellungseigenschaft für die Abbildungsverknüpfung $\vartheta_{U_2 U_1} = \vartheta_{U_2} \circ \vartheta_{U_1}$ ist also erfüllt. Natürlich lassen sich alle eben eingeführten Begriffe auch für Gruppen in $GL(m, \mathbb{R})$ in völliger Analogie auf den rein reellen Fall einschränken.

Um die gewünschten endlichdimensionalen Darstellungen der interessierenden Untergruppen der $GL(2, \mathbb{C})$ zu konstruieren, betrachten wir nun für $m = 2$ konkrete Funktionenräume F_{n+1}^2 der komplexen algebraischen Dimension $(n + 1) \in \mathbb{N}$, welche durch die Basis-Monome

$$[x_1^n, x_1^{n-1}x_2, x_1^{n-2}x_2^2, \dots, x_1x_2^{n-1}, x_2^n] = F_{n+1}^2 \quad (162)$$

aufgespannt werden. Um ein konkretes Beispiel zu studieren, wählen wir an dieser Stelle $n = 2$ und lassen eine spezielle unitäre Matrix $U \in SU(2)$ beziehungsweise die inverse Matrix $U^{-1} \in SU(2)$

$$U = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}, \quad U^{-1} = \begin{pmatrix} \bar{a} & -b \\ \bar{b} & a \end{pmatrix} = U^+, \quad a\bar{a} + b\bar{b} = 1 = \det U \quad (163)$$

auf die Argumente der Basis-Monome wirken: Mit

$$U^{-1}x = \begin{pmatrix} \bar{a} & -b \\ \bar{b} & a \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \bar{a}x_1 - bx_2 \\ \bar{b}x_1 + ax_2 \end{pmatrix} \quad (164)$$

ergibt sich die Zuordnung

$$\begin{aligned} x_1^2 &\mapsto \vartheta_U(x_1^2) = \bar{a}^2 x_1^2 - 2\bar{a}b x_1 x_2 + b^2 x_2^2, \\ x_1 x_2 &\mapsto \vartheta_U(x_1 x_2) = \bar{a}\bar{b} x_1^2 + (a\bar{a} - b\bar{b})x_1 x_2 - abx_2^2, \\ x_2^2 &\mapsto \vartheta_U(x_2^2) = \bar{b}^2 x_1^2 + 2a\bar{b}x_1 x_2 + a^2 x_2^2, \end{aligned} \quad (165)$$

welche sich mit Hilfe einer Darstellungsmatrix

$$D(U) = \begin{pmatrix} \bar{a}^2 & \bar{a}\bar{b} & \bar{b}^2 \\ -2\bar{a}\bar{b} & a\bar{a} - b\bar{b} & 2a\bar{b} \\ b^2 & -ab & a^2 \end{pmatrix} \quad (166)$$

auf dem durch die Basis-Funktionen x_1^2, x_1x_2 und x_2^2 aufgespannten Raum darstellen lässt. $D(U)$ wird durch den Basiswechsel $x_1^2, x_1x_2, x_2^2 \rightarrow x_1^2, \sqrt{2}x_1x_2, x_2^2$ sogar unitär, wenn wir die Basisfunktionen $x_1^2, \sqrt{2}x_1x_2, x_2^2$ zur Orthonormalbasis in F_3^2 erheben

$$D'(U) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \bar{a}^2 & \bar{a}\bar{b} & \bar{b}^2 \\ -2\bar{a}\bar{b} & a\bar{a} - b\bar{b} & 2a\bar{b} \\ b^2 & -ab & a^2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \bar{a}^2 & \sqrt{2}\bar{a}\bar{b} & \bar{b}^2 \\ -\sqrt{2}\bar{a}\bar{b} & a\bar{a} - b\bar{b} & \sqrt{2}a\bar{b} \\ b^2 & -\sqrt{2}ab & a^2 \end{pmatrix}. \quad (167)$$

Leicht prüft man beispielsweise nach, dass die Norm der linken Spalte der Matrix $D(U')$

$$a^2\bar{a}^2 + 2a\bar{a}b\bar{b} + \bar{b}^2b^2 = (a\bar{a} + b\bar{b})^2 = 1 \quad (168)$$

beträgt.

In der Physik ist es üblich, die für die Dimension $(n + 1)$ der Darstellungen charakteristische Zahl n mit der Drehimpulsquantenzahl $l = n/2$ zu verknüpfen. Das oben konstruierte Beispiel (167) entspricht also $l = 1$, und wir wollen künftig die auf dem $n + 1 = (2l + 1)$ -dimensionalen Raum F_{n+1}^2 nach obiger Vorschrift (160) bis auf Äquivalenz konstruierten Darstellungen mit ϑ_l bezeichnen. Die Tatsache, dass ϑ_1 eine dreidimensionale Darstellung ist, ist eng verknüpft mit der Tatsache, dass der Zustand eines ruhenden elementaren Spin-1-Systems als Superposition dreier Zustände $|1, -1\rangle, |1, 0\rangle$ und $|1, 1\rangle$ mit magnetischen Quantenzahlen $m = -1, 0, 1$ dargestellt werden kann. Eine $SU(2)$ -Matrix U (oder $-U$) beschreibt eine Rotation im dreidimensionalen euklidischen Raum; die entsprechende unitäre Darstellungsmatrix $D'(U)$ dreht den Zustandsvektor des Spin-1-Systems und überträgt so die reale räumliche Rotation in den abstrakten, unter Umständen auch höherdimensionalen Hilbert-Zustandsraum des Systems.

Eine eingehende Analyse zeigt, dass mit den (unitarisierbaren) Darstellungen $\vartheta_l, l = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots$ alle Darstellungen der $SU(2)$ wie bereits erwähnt in den entsprechenden endlichen Dimensionen in erschöpfender Weise aufgelistet sind. In jeder Dimension existiert bis auf Äquivalenz also genau eine irreduzible Darstellung der $SU(2)$, insbesondere gilt $\vartheta_l = \vartheta_l^*$, es sind also alle Darstellungen äquivalent zu ihrer komplex konjugierten Version. Die Fundamentaldarstellung der $SU(2)$ ist die Darstellung der Gruppe durch sich selbst, die triviale Darstellung in einer Dimension erhält man für $l = 0$. Für $l \in \mathbb{N}_0$ lassen sich bei geeigneter Basiswahl im Darstellungsraum sämtliche Darstellungsmatrizen durch rein reelle Matrizen darstellen, entsprechend nennt man die ganzzahligen Darstellungen *reell*. Im Gegensatz zu den halbzahligen Darstellungen sind diese Darstellungen nicht treu, da sie $\pm U \in SU(2)$ auf dieselbe Darstellungsmatrix abbilden.

In genau derselben Weise wie im Falle der $SU(2)$ konstruiert man (Strahl)-Darstellungen der $(SO^+(1, 3))$ $SL(2, \mathbb{C})$, welche wir in weiser Voraussicht mit $\vartheta_{j,0}$ bezeichnen wollen, wobei die Dimension der Darstellung wiederum durch $\dim \vartheta_{j,0} = n + 1 = 2j + 1$ gegeben ist. Wie wir bei der zweidimensionalen Fundamentaldarstellung $\vartheta_{\frac{1}{2},0}$ der $SL(2, \mathbb{C})$ bereits gesehen haben, ist diese im Gegensatz zur Fundamentaldarstellung $\vartheta_{\frac{1}{2}} = \vartheta_{\frac{1}{2}}^*$ der $SU(2)$ nicht äquivalent zur komplex konjugierten Darstellung $\vartheta_{0,\frac{1}{2}} = \vartheta_{\frac{1}{2},0}^*$; es stellt sich sogar heraus, dass die mit $\vartheta_{0,j} = \vartheta_{j,0}^*$ bezeichneten komplex konjugierten Darstellungen einen weiteren Satz irreduzibler endlichdimensionaler Darstellungen der $SL(2, \mathbb{C})$ liefern, die bis auf die triviale Darstellung allesamt nicht-unitär sind. Damit aber nicht genug. Die vollständige Liste aller endlichdimensionalen irreduziblen Darstellungen der $SL(2, \mathbb{C})$ ergibt sich schliesslich und endlich, wenn man auch die direkten und irreduziblen Produktdarstellungen

$$\vartheta_{j,j'} = \vartheta_{j,0} \otimes \vartheta_{0,j'} = \vartheta_{j,0} \otimes \vartheta_{j',0}^* \quad (169)$$

hinzu nimmt, wobei aber bei gleichen Indexpaaren $j = j'$ eine Darstellung $\vartheta_{j,j}$ jeweils äquivalent zu $\vartheta_{j,j}^*$ ist; man nennt eine solche Darstellung auch insofern reell, weil sie in einer geeigneten Basis durch rein reelle Matrizen dargestellt werden kann.

Tatsächlich spielen in der Quantenmechanik auch *reelle irreduzible Darstellungen* der eigentlichen orthochronen Lorentzgruppe eine wichtige Rolle. Diese reellen Darstellungen stehen im Zusammenhang mit der Konstruktion reeller Wellengleichungen und deren Wellenlösungen, welche der Beschreibung von ladungsneutralen Einteilchenzuständen in der relativistischen Wellenmechanik dienen. So ist das Higgsfeld klassisch gesehen ein neutrales skalares Feld, das elektrisch neutrale elektromagnetische Feld wird durch ein reelles Vektorfeld oder einen reellen antisymmetrischen Feldstärketensor beschrieben, und das Gravitationsfeld lässt sich mit Hilfe eines symmetrischen reellen metrischen Tensors beschreiben. Diese klassischen Zustände werden dann aber in der vollen Quantenfeldtheorie nach einer 'zweiten Quantisierung' wiederum komplex superponiert, da die Quantentheorie eine Theorie mit manifest komplexer Struktur ist.

Die reell-irreduziblen (Strahl-)Darstellungen der $SO^+(1, 3)$ lassen sich in zwei Typen einteilen:

- Typ 1: $\vartheta_j^{\mathbb{R}}$ entsteht durch Einschränkung einer komplexen, zur $\vartheta_{j,j}$ äquivalenten $SL(2, \mathbb{C})$ -Darstellung auf den reellen Unterraum $\mathbb{R}_{\mathbb{R}}^{(2j+1)^2} \subset \mathbb{C}_{\mathbb{C}}^{(2j+1)^2}$, wie bereits oben erwähnt. Suggestiv kann man die Einschränkung der komplexen Darstellungen (j, j) auf den reellen Fall mit $(j, j)_{\mathbb{R}}$ notieren.
- Typ 2: $\vartheta_{j,k}^{\mathbb{R}} \simeq \vartheta_{k,j}^{\mathbb{R}}$ mit $j \neq k$ entsteht durch Einschränkung der direkten Summe zweier zueinander komplex konjugierter $SL(2, \mathbb{C})$ -Darstellungen $\vartheta_{j,k} \oplus \vartheta_{k,j}$ auf den reellen Unterraum $\mathbb{R}_{\mathbb{R}}^{2(2j+1)(2k+1)} \subset \mathbb{C}_{\mathbb{C}}^{2(2j+1)(2k+1)}$. Aus 'komplexer Sicht' wären solche Darstellungen also reduzibel, aber sie sind es nicht als reelle Darstellungen. Kurz können diese Darstellungen, auf deren eingehendere Beschreibung hier nicht eingegangen werden soll, auch durch $(j, k)_{\mathbb{R}} \simeq (k, j)_{\mathbb{R}}$ notiert werden.

Nochmals sei bemerkt: Nur die triviale Darstellung $\vartheta_{0,0}$ ist unitär, ansonsten ist es unmöglich, eine irreduzible und endlich-dimensionale stetige Darstellung der nicht-kompakten Lorentzgruppe zu unitarisieren. Hier ist zu erwähnen, dass für Darstellungen von Lie-Gruppen die Stetigkeit gefordert wird. Tatsächlich kann man für endlichdimensionale Darstellungen dann sogar zeigen, dass aus der Stetigkeit sogar die *Glatttheit* des Darstellungshomomorphismus folgt.

In unendlich vielen Dimensionen existieren (irreduzible) unitäre Darstellungen der Lorentzgruppe. Das ist auch gut so, denn die entsprechenden Darstellungsräume spielen in der Physik oft die Rolle von physikalischen Zustandsräumen oder 'Wellenfunktionen-Räumen', und die Unitarität der Darstellungen steht dann in direkter Beziehung zur Wahrscheinlichkeitserhaltung in der Quantenmechanik.

Einschub: Lösungen und Eigenschaften der Dirac-Gleichung

“Es gibt keinen Gott und Dirac ist sein Prophet.”

Wolfgang Pauli (1900 - 1958)
an der *fünften Solvay-Konferenz* im Jahre 1927.

Ebene Wellenlösungen

In der Folge wollen wir die Lösungen der freien Dirac-Gleichung herleiten, wobei wir zuerst die in der Literatur verbreitete Standard-Darstellung der Dirac-Matrizen

$$\gamma^0 = \begin{pmatrix} \mathbf{1}_2 & 0 \\ 0 & -\mathbf{1}_2 \end{pmatrix} = \begin{pmatrix} \sigma_0 & 0 \\ 0 & -\sigma_0 \end{pmatrix}, \quad \gamma^k = \begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix} \quad (170)$$

verwenden. Die freie Dirac-Gleichung lautet ja mit der reduzierten Compton-Wellenlänge $\lambda = \hbar(mc)^{-1}$

$$i\gamma_\mu \partial^\mu \Psi(x) - \lambda^{-1} \Psi(x) = i\gamma^0 \frac{\partial}{c \partial t} \Psi(ct, \vec{x}) + i\vec{\gamma} \cdot \vec{\nabla} \Psi(ct, \vec{x}) - \frac{mc}{\hbar} \Psi(ct, \vec{x}) = 0 \quad (171)$$

beziehungsweise mit $\hbar = c = 1$

$$(i\gamma_\mu \partial^\mu - m)\Psi(x) = i\gamma_\mu \partial^\mu \Psi(x) - m\Psi(x) = i\gamma^0 \dot{\Psi}(x) - \vec{\gamma} \cdot \frac{\vec{\nabla}}{i} \Psi(x) - m\Psi(x) = 0. \quad (172)$$

Operiert man auf der obigen Gleichung von links mit $(-i\gamma_\nu \partial^\nu - m)$, so folgt mit der Vertauschbarkeit partieller Ableitungen glatter Wellenfunktionen die Klein-Gordon-Gleichung

$$\begin{aligned} (-i\gamma_\nu \partial^\nu - m)(i\gamma_\mu \partial^\mu - m)\Psi(x) &= (\gamma_\nu \gamma_\mu \partial^\nu \partial^\mu + m^2)\Psi(x) = \frac{1}{2} \{\gamma_\nu, \gamma_\mu\} \partial^\nu \partial^\mu \Psi(x) + m^2 \Psi(x) \\ &= (g_{\mu\nu} \partial^\mu \partial^\nu + m^2)\Psi(x) = (\square + m^2)\Psi(x) = 0 \end{aligned} \quad (173)$$

für alle vier Komponenten des Dirac-Spinors.

Nun führt der ebene Wellenansatz ($p w \sim$ plane wave)

$$\Psi_{pw}(p; x) = u(p)e^{-ipx} = u(p)e^{-ip^0 x^0 + i\vec{p} \cdot \vec{x}} = u(p)e^{-i\omega t + i\vec{p} \cdot \vec{x}} \quad (174)$$

mit einem Viererimpulsvektor $p^\mu = (p^0, p^1, p^2, p^3) = (p^0, \vec{p})$ und einem zu bestimmenden Viererspinor $u(p)$ wegen $\partial^\mu e^{-ipx} = -ip^\mu e^{-ipx}$ beziehungsweise $i\gamma_\mu \partial^\mu e^{-ipx} = \gamma_\mu p^\mu e^{-ipx}$ auf die elegante Bedingung

$$(\not{p} - m)u(p) = 0. \quad (175)$$

Dabei haben wir die praktische Feynman-Slash-Notation \not{p} für die in der theoretischen Behandlung von Dirac-Teilchen häufig vorkommende 4×4 -Matrix $p_\mu \gamma^\mu$ eingeführt:

$$\not{p} \equiv p^\mu \gamma_\mu. \quad (176)$$

Diese Notation wird in der Literatur generell zur Darstellung einer mit einem Vierervektor a gebildeten Linearkombination von Dirac-Matrizen

$$\not{a} = a_\mu \gamma^\mu = a^\mu \gamma_\mu = a^0 \gamma^0 - a^1 \gamma^1 - a^2 \gamma^2 - a^3 \gamma^3 = a^0 \gamma^0 - \vec{a} \cdot \vec{\gamma} \quad (177)$$

verwendet. Analog schreibt man auch $\not{\partial} = \gamma^\mu \partial_\mu = \gamma_\mu \partial^\mu$.

Da $\Psi_{pw}(p; x)$ die Klein-Gordon-Gleichung erfüllt, gilt für die Komponenten des Viererimpulsvektors die Massenschalenbedingung

$$p_\mu p^\mu = p_0^2 - \vec{p}^2 = m^2 \quad \text{beziehungsweise} \quad p^0 = \pm \sqrt{\vec{p}^2 + m^2}. \quad (178)$$

Es gilt also immer $|p^0| \geq m \geq 0$. Zerlegt man nun den zeitunabhängigen Dirac-Spinor u in einen oberen und einen unteren Teil gemäss

$$u(p) = \begin{pmatrix} \xi(p) \\ \eta(p) \end{pmatrix} \quad (179)$$

mit zweikomponentigen Subspinoren ξ und η , so lässt sich die Dirac-Gleichung schreiben wie folgt

$$(p^0 \gamma^0 - \vec{p} \cdot \vec{\gamma} - m)u(p) = 0 \quad (180)$$

beziehungsweise

$$\begin{pmatrix} p^0 - m & -\vec{p} \cdot \vec{\sigma} \\ \vec{p} \cdot \vec{\sigma} & -p^0 - m \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} = 0 \quad (181)$$

beziehungsweise

$$(p^0 - m)\xi - (\vec{p} \cdot \vec{\sigma})\eta = 0, \quad (182a)$$

$$(\vec{p} \cdot \vec{\sigma})\xi - (p^0 + m)\eta = 0. \quad (182b)$$

Aus Gleichung (182b) folgt sofort für $p_0 \geq m > 0$

$$\eta = \frac{\vec{p} \cdot \vec{\sigma}}{p^0 + m} \xi, \quad (183)$$

und damit ist automatisch auch Gleichung (182a) erfüllt, da ja p^μ auf der Massenschale $p^2 = p_0^2 - \vec{p}^2 = m^2$ liegt:

$$\left(p^0 - m - \frac{(\vec{p} \cdot \vec{\sigma})^2}{p^0 + m} \right) \xi = \left(p^0 - m - \frac{\vec{p}^2}{p^0 + m} \right) \xi = \frac{p_0^2 - m^2 - \vec{p}^2}{p^0 + m} \xi = 0. \quad (184)$$

Wir betrachten also hier zuerst den negativ-frequenten (*massiven*) 'Normalfall' $p^0 (\geq m) \geq 0$; der zweikomponentige Spinor ξ ist dann völlig beliebig wählbar, und zu einem Viererimpuls p existieren jeweils *zwei* linear unabhängige und superponierbare ebene Wellen-Lösungen der Dirac-Gleichung, die wir an dieser Stelle folgendermassen normieren wollen:

$$u_s(\vec{p}) = u_s(p^0, \vec{p}) = \sqrt{p^0 + m} \begin{pmatrix} \chi_s \\ \frac{\vec{p} \cdot \vec{\sigma}}{p_0 + m} \chi_s \end{pmatrix} \quad \text{mit} \quad \chi_{\frac{1}{2}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{und} \quad \chi_{-\frac{1}{2}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (185)$$

Natürlich lassen sich diese Zustände nach der Konstruktion passender Drehimpulsoperatoren zur Beschreibung von Fermionen mit einer Spinausrichtung in positiver oder negativer 3-Richtung verwenden. Es fällt auf, dass sich der u -Spinor bei Teilchen mit kleinem Impuls aus einer oberen 'grossen Komponente' und einer unteren 'kleinen Komponente' zusammensetzt.

Wir könnten nun mit der Betrachtung des Falles $p^0 \leq -m \leq 0$ fortfahren. Im Wissen, dass auch Antiteilchen positive Energie haben⁶, wollen wir aber p^0 in der Folge als Teilchenenergie auffassen und für die positiv-frequenten Lösungen der freien Dirac-Gleichung den Ansatz

$$\Psi_{pw}^a(p; x) = v(p)e^{ipx} = \begin{pmatrix} \bar{\xi}(p) \\ \bar{\eta}(p) \end{pmatrix} e^{ipx}, \quad p^0 = \sqrt{\vec{p}^2 + m^2} > 0 \quad (186)$$

⁶Die Existenz betragsmässig beliebig grosser negativer Teilchenenergien würde zu einer explosiven Reorganisation des Universums führen. Die elektrisch geladenen Elektronen könnten durch Emission von Photonen in beliebig tiefe Energiezustände übergehen.

in die Dirac-Gleichung einsetzen. Es folgt

$$(p^0 + m)\bar{\xi} - (\vec{p} \cdot \vec{\sigma})\bar{\eta} = 0, \quad (187a)$$

$$(\vec{p} \cdot \vec{\sigma})\bar{\xi} - (p^0 - m)\bar{\eta} = 0, \quad (187b)$$

also ist analog zu Gleichung (183)

$$\bar{\xi} = \frac{\vec{\sigma} \cdot \vec{p}}{p^0 + m} \bar{\eta} \quad (188)$$

mit einem frei wählbaren Zweierspinor $\bar{\eta}$. Rechentechnisch praktisch sind die folgenden linear unabhängigen positiv-frequenten ebenen Wellenlösungen zu gegebenem Antiteilchenviererimpuls p und Spin $s = \pm 1/2$:

$$v_s(p) = v_s(\vec{p}) = -\sqrt{p^0 + m} \begin{pmatrix} \frac{\vec{p} \cdot \vec{\sigma}}{p^0 + m} \epsilon \chi_s \\ \epsilon \chi_s \end{pmatrix} \quad \text{mit} \quad \epsilon = i\sigma_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \quad (189)$$

Eine allgemeine Lösung der freien Dirac-Gleichung ergibt sich als Superposition ebener Wellen zu positiver und negativer Frequenz

$$\begin{aligned} \Psi(x) &= \int \frac{d^3p}{(2\pi)^3 2p^0} \sum_{s=\pm\frac{1}{2}} [a_s(\vec{p}) u_s(\vec{p}) e^{-ipx} + b_s^*(\vec{p}) v_s(\vec{p}) e^{+ipx}] \\ &= \int \frac{d^4p}{(2\pi)^3} \Theta(p_0) \delta(p^2 - m^2) \sum_{s=\pm\frac{1}{2}} [a_s(\vec{p}) u_s(\vec{p}) e^{-ipx} + b_s^*(\vec{p}) v_s(\vec{p}) e^{+ipx}]. \end{aligned} \quad (190)$$

Dabei bezeichnet das *-Symbol wie üblich die komplexe Konjugation und $a_s(\vec{p})$ und $b_s(\vec{p})$ sind hinreichend anständige komplexe Funktionen, die für grosse Impulse betragsmässig schnell genug abfallen und nach der Prozedur der Feldquantisierung in den Stand der Feldoperatoren (d.h. der operatorwertigen Distributionen) erhoben werden.

Nun wird die Dirac-Gleichung also zur Beschreibung der Dynamik von Spin- $\frac{1}{2}$ -Fermionen verwendet, wobei bei deren Lösung offensichtlich zwei Klassen von ebenen Wellenlösungen auftreten: negativ-frequente Lösungen $\sim e^{-ipx}$ und positiv-frequente Lösungen e^{+ipx} . Während die negativ-frequenten Lösungen das aus der nicht-relativistischen Wellenmechanik eines Teilchens der Energie p^0 und Impuls \vec{p} vertraute Vorzeichen aufweisen, bringen uns die neuartigen positiv-frequenten Wellen in eine gewisse Erklärungsnot. Beschreiben diese Lösungen Zustände negativer Energie? Auch Antiteilchen besitzen positive Energie. Dirac wusste 1928 aber, als er seine Gleichung fand, noch nichts von den Anti-Elektronen, den Positronen; und schon gar nichts wusste er von den Myonen und Anti-Myonen, den Tauonen und Anti-Tauonen und den (Anti-)Quarks, welche sich alle mit Hilfe von Diracs Theorie exzellent beschreiben lassen - während die (Anti-)Neutrinos vielleicht eine gesonderte Behandlung erfordern. Als Gegenstück des Elektrons kam für Dirac bestenfalls das positiv geladene, aber viel schwerere Proton in Frage.

Kann man sich nun mit einem vereinfachten Bild erklären, dass ein Antiteilchen mit Spin in 3-Richtung gemäss obiger Behauptung durch eine positiv-frequente Wellenfunktion und einen Spinor

$$\epsilon \chi_{\frac{1}{2}} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \quad (191)$$

beschrieben wird, welcher mit einem Teilchen mit negativ-frequenter Wellenfunktion mit zur 3-Richtung entgegengesetztem Spin assoziiert ist? Vielleicht so: Da die positiv-frequenten ebenen Antiteilchen-Lösungen $\sim e^{+ipx}$ naiv betrachtet Zustände negativer Energie beschreiben, stellte sich Dirac vor, dass im physikalischen Vakuum praktisch alle energetisch tief liegenden Zustände besetzt seien (was natürlich die Frage aufwarf, weshalb das Vakuum ungeladen ist). Ein *unbesetzter* Zustand in diesem 'Dirac-See' negativer Energie würde dann

einem beobachtbaren Antiteilchen entsprechen. Damit kehren sich die beobachtbaren Teilcheneigenschaften um: Ein unbesetzter Zustand negativer Energie mit Spin s in z -Richtung entspricht einem Teilchen positiver Energie mit Spin $-s$.

Die Betrachtung der sogenannten Ladungskonjugation ist für das Verständnis der Lösungen positiver Frequenz ebenfalls dienlich. Man muss sich aber klar machen, dass eine befriedigendere Betrachtung der vorliegenden Problematik erst im Rahmen der Quantenfeldtheorie sinnvoll ist. Und um die Sache noch etwas zu komplizieren, sei hier zudem erwähnt dass in der realen Welt wechselwirkender Teilchen die Teilchen mit ihren Antiteilchen nicht durch eine einfache Ladungskonjugation, sondern durch eine sogenannte CPT-Transformation miteinander verknüpft sind.

Klassifikation relativistischer Felder

Felder mit Spin $0, \frac{1}{2}$ und 1

'Felder' oder 'Wellenfunktionen' als raumzeitabhängige Funktionen dienen in der Physik dazu, das räumliche und zeitliche Verhalten physikalischer Grössen oder 'Observablen' verschiedenster Art zu beschreiben. Zu diesen Grössen gehören in erster Linie quantenmechanische Wahrscheinlichkeits(strom)dichten, Ladungsdichten und Ladungsstromdichten verschiedenster Natur. Raumzeitfunktionen wie die elektromagnetischen Felder können bei geeigneter Interpretation als mit dem Photon assoziierte Felder oder Wellenfunktionen aufgefasst werden. Die Wellenfunktionen entsprechen in vielen Fällen keiner unmittelbar beobachtbaren Grösse, sondern dienen der Konstruktion oder Berechnung einer solchen. Oft werden Wellenfunktionen als Elemente eines Funktionen-Hilbertraums zur Beschreibung einzelner Teilchen oder entsprechender Antiteilchen konstruiert. Werden solche Teilchen-Wellenfunktionen zudem einer sogenannten *zweiten Quantisierung* unterworfen, dann werden sie zu *operatorwertigen Distributionen*. Aus der Quantenmechanik als 'naive' Wellenmechanik entsteht so die Quantenfeldtheorie als eine mathematische Theorie operatorwertiger Distributionen. Während in der Wellenmechanik die zeitliche Entwicklung eines quantenmechanischen Systems wie z.B. des Wasserstoffatoms unter Zuhilfenahme von klassischen Konzepten wie dem Coulombpotential beschrieben wird, so hat die Quantenfeldtheorie den weitreichenderen Anspruch, die quantenmechanische Ursache der Existenz und Form des in der klassischen Physik auftretenden Coulombpotentials als Folge tiefergreifenderer physikalischer Prinzipien zu erklären, wobei das Coulombfeld in der Quantenfeldtheorie selbst einem Quantisierungsprozess unterworfen wird.

Unter dem *manifest Lorentz-invarianten Transformationsverhalten* von Feldern unter der eigentlichen orthochronen Lorentzgruppe $SO^+(1, 3)$ versteht man ein Transformationsverhalten gemäss der Vorschrift

$$\Psi'_\lambda(x') = D_\lambda^\rho(\Lambda)\Psi_\rho(x), \quad x'^\mu = \Lambda^\mu_\nu x^\nu, \quad (192)$$

wobei $D(\Lambda)$ die $n \times n$ -Darstellungsmatrix einer nicht notwendigerweise irreduziblen (Strahl-)Darstellung der Lorentzgruppe $SO^+(1, 3) \supset \Lambda$ in beliebiger Dimension n mit der Darstellungseigenschaft

$$D(\Lambda_2\Lambda_1)/\{\pm\mathbb{1}\} = D(\Lambda_2)D(\Lambda_1)/\{\pm\mathbb{1}\} \quad \text{oder} \quad D_\lambda^\rho(\Lambda_2\Lambda_1) = \pm D_\lambda^\alpha(\Lambda_2)D_\alpha^\rho(\Lambda_1) \quad (193)$$

ist. Auf die Spitzfindigkeit, dass n in gewissem Sinne unendlich gross sein könnte, wollen wir hier nicht eingehen. Natürlich sind die durch 'Spinor-Indizes' indizierten Feldkomponenten *nicht* Lorentz-invariant, aber das Transformationsverhalten drückt eine manifeste Invarianz in dem Sinne aus, dass jeder Beobachter in jedem Inertialsystem weiss, wie er seine Felder durch eine invariante Rezept (192) in ein anderes Inertialsystem umrechnen könnte, wenn er dies wollte.

Folgende ausgewählte Felder, welche unter den irreduziblen (Strahl-)Darstellungen $\vartheta_{j,j'}$ der Lorentzgruppe niedrigster Dimension wie unten dargestellt bei einem passiven Wechsel von einem Inertialsystem IS mit kartesischen Ereigniskordinaten x in ein Inertialsystem IS' mit Ereigniskordinaten $x' = \Lambda x$ transformieren, spielen in Quanten(feld)theorien der $3 + 1$ -dimensionalen Raumzeit eine wichtige Rolle:

- $(j, j') = (0, 0)$: Reelles oder komplexes skalares Feld $\varphi(x)$.

$$\varphi'(x') = D_{0,0}(\Lambda)\varphi(x) = \varphi(x) = \varphi(\Lambda^{-1}x'). \quad (194)$$

- $(j, j') = (\frac{1}{2}, 0)$: Komplexes, zweikomponentiges rechts-chirales Spinorfeld $\psi_R(x)$.

$$\psi'_{R,\alpha}(x') = A_\alpha^\beta(\Lambda)\psi_{R,\beta}(x) = A_\alpha^\beta(\Lambda)\psi_{R,\beta}(\Lambda^{-1}x'), \quad A(\Lambda) = D_{\frac{1}{2},0}(\Lambda). \quad (195)$$

- $(j, j') = (0, \frac{1}{2})$: Komplexes, zweikomponentiges links-chirales Spinorfeld $\psi_L(x)$.

$$\psi'_{L,\bar{\alpha}}(x') = A_{\bar{\alpha}}^{*\bar{\beta}}(\Lambda)\psi_{L,\bar{\beta}}(x) = A_{\bar{\alpha}}^{*\bar{\beta}}(\Lambda)\psi_{L,\bar{\beta}}(\Lambda^{-1}x'), \quad A^*(\Lambda) = D_{0,\frac{1}{2}}(\Lambda). \quad (196)$$

Die überstrichenen Indizes verdeutlichen das Transformationsverhalten des links-chiralen Spinorfeldes unter der komplex konjugierten fundamentalen Strahldarstellung der eigentlichen orthochronen Lorentzgruppe. Links- und rechts-chirale Spinorfelder spielen eine wichtige Rolle in der Neutrinophysik jenseits des Standardmodells der Teilchenphysik und in allgemeinerem Sinne bei der Beschreibung von Spin- $\frac{1}{2}$ -Fermionen.

- $(j, j') = (\frac{1}{2}, \frac{1}{2})$: Reelles oder komplexes Vektorfeld $V_{\alpha\bar{\beta}}(x)/V^\mu(x)$.
Das Transformationsgesetz unter dem direkten Produkt der Darstellungen $\vartheta_{\frac{1}{2},0}$ und der $\vartheta_{0,\frac{1}{2}} = \vartheta_{\frac{1}{2},0}^*$

$$V'_{\alpha\bar{\beta}}(x') = A_\alpha^\gamma A_{\bar{\beta}}^{*\delta} V_{\gamma\bar{\delta}}(\Lambda^{-1}x') \quad (197)$$

kann in eine interessante Form gebracht werden, wenn man die verallgemeinerten Pauli-Matrizen als komplexe Basis des $Mat(2, \mathbb{C})$ als Vektorraum über \mathbb{C} verwendet. Man schreibt bei Verwendung der Einsteinschen Summenkonvention

$$V_{\gamma\bar{\delta}}(x) = \sum_{\mu=0}^3 V^\mu(x) (\sigma_\mu)_{\gamma\bar{\delta}} = V^\mu(x) (\sigma_\mu)_{\gamma\bar{\delta}} \sim \begin{pmatrix} V^0 + V^3 & V^1 - iV^2 \\ V^1 + iV^2 & V^0 - V^3 \end{pmatrix} (x) \quad (198)$$

und erhält somit

$$\begin{aligned} V'^{\mu}(\sigma_\mu)_{\alpha\bar{\beta}}(x') &= V'_{\alpha\bar{\beta}}(x') = A_\alpha^\gamma A_{\bar{\beta}}^{*\delta} V_{\gamma\bar{\delta}}(\Lambda^{-1}x') \\ &\stackrel{!}{=} \underbrace{(AV(x)A^+)_{\alpha\bar{\beta}}} \stackrel{!}{=} (\Lambda^\mu_\nu V^\nu(x)) (\sigma_\mu)_{\alpha\bar{\beta}}. \\ &= \begin{pmatrix} V'^0 + V'^3 & V'^1 - iV'^2 \\ V'^1 + iV'^2 & V'^0 - V'^3 \end{pmatrix} (x') \end{aligned} \quad (199)$$

V^μ ist also ein *Vektorfeld* und transformiert wie die kartesischen Raumzeitkoordinaten. V ist nicht notwendigerweise reell; während das aus der Elektrodynamik bekannte, aus dem elektrostatischen Potential V und dem drei-komponentigen magnetischen Vektorpotential \vec{A} zusammengesetzte reelle Viererpotential $V^\mu \sim A^\mu = (\Phi/c, \vec{A})$ das elektromagnetische Feld beschreibt, werden die geladenen W^\pm -Bosonen im Standardmodell der Elementarteilchenphysik durch ein komplexes Viererpotential $W^{*\mu} \neq W^\mu$ beschrieben. Die elektrisch neutralen und massiven Z -Bosonen werden wiederum durch ein reelles Viererpotential Z^μ repräsentiert.

- $(j, j') = (1, 0)$ oder $(0, 1)$: Komplexe Riemann-Silberstein-Vektoren $F(x) \in \mathbb{C}^3$. Bildet man die direkte Summe der Darstellungen $(1, 0) \oplus (0, 1)$, so lässt sich diese zur Konstruktion einer reellen sechsdimensionalen Darstellung der Lorentzgruppe verwenden, welche zur Beschreibung des Lorentz-Transformationsverhaltens des elektromagnetischen Feldes $(\vec{E}(x), \vec{B}(x))$ verwendet werden kann.
- $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$: Dirac-Spinoren $\psi(x)$, welche der Beschreibung geladener Spin- $\frac{1}{2}$ -Teilchen (z. B. Leptonen, Quarks) dienen. Die Darstellung $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$ kann auf den reellen vierdimensionalen Raum eingeschränkt werden und führt so auf eine reelle, vierdimensionale Darstellung der Lorentzgruppe, die *nicht* im reellen Sinne äquivalent ist zur vierdimensionalen reellen Darstellung der Lorentzgruppe durch sich selbst. Die mit dieser Darstellung verknüpften Felder $\psi_M(x)$ heissen *Majorana-Felder*.

Einschub: Teilchen, Wechselwirkungen und Wirkungsquerschnitte

“Wenn man für ein Experiment einen Statistiker benötigt, dann hätte man ein besseres Experiment durchführen sollen.”

Ernest Rutherford (1871-1937).

Der Begriff des *Wirkungsquerschnitts* ist von zentraler Bedeutung für die experimentelle und theoretische Atom-, Astro-, Kern- und Teilchenphysik und darüber hinaus für viele weitere naturwissenschaftlich-technische Anwendungs- und Forschungs-Bereiche. Als Formelzeichen für den Wirkungsquerschnitt wird üblicherweise der griechische Buchstabe σ verwendet.

Der Wirkungsquerschnitt ist ein bestimmendes Mass für die Wahrscheinlichkeit, dass zwischen einfallenden Teilchen (oder *Projektilen* in einem Teilchenstrahl) und Zielteilchen (welche im experimentellen Rahmen einen Streukörper oder das *Target* bilden) ein bestimmter Prozess wie beispielsweise eine Streuung (d.h., eine elastische oder inelastische Ablenkung des Projektils an einem Zielteilchen), eine Absorption oder eine weitere spezifische Reaktion stattfindet, bei welcher durchaus auch neue Teilchen erzeugt werden können. Entsprechend spricht man präziser von Streuquerschnitten σ_s , Absorptionsquerschnitten σ_a und vielen weiteren Reaktionstypen entsprechenden Reaktionsquerschnitten σ_r . Die Unterscheidung kollidierender Teilchen in Projektilen und Targets ist natürlich künstlicher Natur. Sie widerspiegelt aber die häufig experimentell vorliegende Situation, bei welcher aus einer Beschleunigeranlage stammende Projektilen auf eine im Labor ruhende Probe geschossen werden.

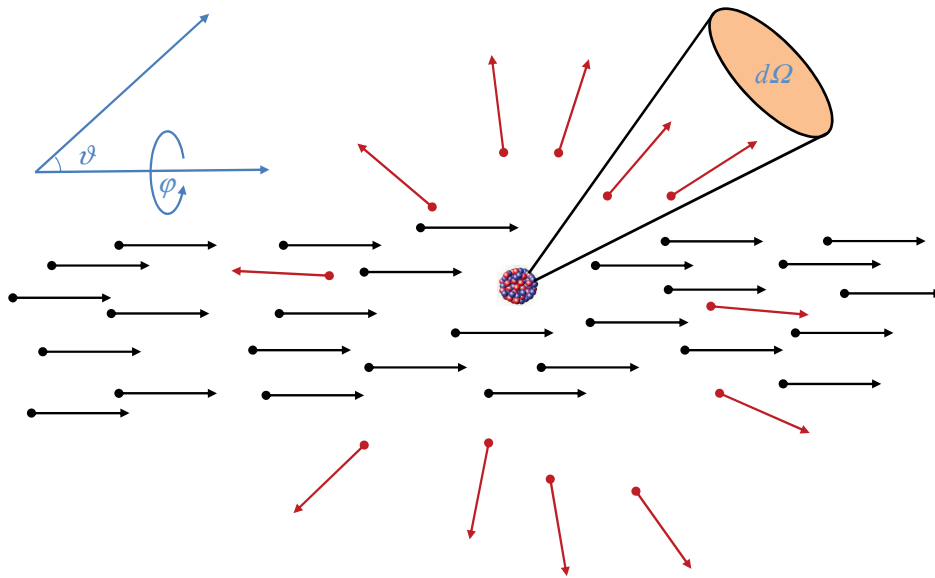


Abbildung 9: Klassisches Sinnbild für die Streuung von Projektilen an einem anfänglich ruhenden Atomkern. Weit vom als Streuzentrum agierenden Atomkern befindliche Detektoren stellen fest, wieviele Teilchen des von links einfallenden Teilchenstroms in einen (bei theoretischen Betrachtungen oft infinitesimalen) Raumwinkel $d\Omega$ gestreut werden. Bei axialer Symmetrie bezüglich der Einfallsrichtung der Projektilen charakterisiert die Angabe der kinetischen Energie der Projektilen E_{kin} und eines Streuwinkels ϑ einen elastischen Streuvorgang bei Abwesenheit innerer Drehimpulse vollständig. Bei komplizierteren experimentellen Anordnungen muss detaillierter nach Streuwinkeln, Energien, Polarisationen und Reaktionskanälen differenziert werden.

Die experimentelle Untersuchung von Wirkungsquerschnitten liefert gewissermassen die Grundlage der Theorien über die Wechselwirkungen, welche unsere Welt im Innersten, im Äussersten und dazwischen regieren.

Wie durch den Ausdruck selbst angedeutet handelt es sich bei einem Wirkungsquerschnitt um eine als effektive Wirkfläche darstellbare Grösse. Man stellt sich vor, dass das Targetteilchen eine Zielscheibe der Grösse des Wirkungsquerschnitts vor den Projektilen aufspannt, welche die dem Wirkungsquerschnitt entsprechende Reaktion auslösen, wenn sie die Zielscheibe treffen. Auch wenn diese Vorstellung der diffusen Quantenrealität des Streuvorganges nicht gerecht wird, so stellt sie dennoch eine effiziente Strategie dar, mit welcher sich zahlreiche experimentelle Daten und theoretische Aussagen anschaulich beschreiben lassen. Natürlich ergibt sich ein Wirkungsquerschnitt in physikalisch realisierten Situationen jeweils aus den *Wechselwirkungen*, an welchen *sowohl* das Projektil *wie auch* das Target *wechselseitig* beteiligt sind; in idealisierten Betrachtungen wird das Target oft als *unverrückbare*, unendlich schwere Quelle eines *statischen* Kraftfeldes angesetzt.

Betrachten wir eine typische experimentelle Anordnung zur Analyse eines Streuvorganges, bei der ein Targetteilchen einer möglichst *homogenen*, *kollinearen* und *monoenergetischen* Teilchenstromdichte I von Projektilen ausgesetzt ist, so sind natürlich die Teilchen von Interesse, welche das Targetteilchen nicht unbeeinflusst passieren. Die Teilchenstromdichte $\vec{I} = \rho_i \vec{v}_i$ beschreibt die Anzahl der Teilchen, die pro Flächen- und Zeit-Einheit aus grosser Entfernung mit einheitlichem *initialem* Geschwindigkeitsvektor $\vec{v}_i \neq 0$ auf das Target zuströmen, welches sich also im Rahmen der vorliegenden Diskussion aus Gründen der Einfachheit in Ruhe befindet möge. Dabei ist $\rho_i \neq 0$ die *initiale* Projektil-Teilchendichte in grosser Entfernung vom Targetteilchen in zur asymptotischen Geschwindigkeit \vec{v}_i entgegengesetzten Richtung.

Die Zahl aller pro Zeiteinheit abgelenkten Teilchen lässt sich durch eine *totale Streurrate* R_s angeben. Das Verhältnis

$$\sigma_s = \frac{R_s}{I} \quad (200)$$

besitzt dann die physikalische Dimension einer *Fläche* und wird als *totaler Streuquerschnitt* σ_s bezeichnet.

Um im Rahmen einer *klassischen*⁷ Beschreibung von Streuvorgängen präzise zu argumentieren, muss auf gewisse Idealisierungen zurückgegriffen werden. Die obigen Forderungen lassen sich formal durch

$$\lim_{t \rightarrow -\infty} \vec{v}(\vec{x} + \vec{v}_i t) = \vec{v}_i \quad (201)$$

für das stationäre Geschwindigkeitsfeld $\vec{v}(\vec{x})$ der Projektilen und

$$\lim_{t \rightarrow -\infty} \rho(\vec{x} + \vec{v}_i t) = \rho_i \quad (202)$$

für die lokale stationäre Projektildichte $\rho(\vec{x})$ ausdrücken. Da für die meisten Streuexperimente im Labor die Effekte der Raumzeitkrümmung vernachlässigbar sind, kann \vec{x} getrost als kartesische Koordinate oder Ortsvektor im flachen dreidimensionalen Raum aufgefasst werden. In einer erweiterten Diskussion könnten auch noch Eigendrehimpulszustände der Reaktionspartner mit einbezogen werden.

Wie andere (totale) Wirkungsquerschnitte auch hängt der totale Streu(wirkungs)querschnitt im Allgemeinen von der kinetischen Energie der Reaktionspartner ab. Bleibt die gesamte kinetische Energie der aneinander streuenden Teilchen beim Streuvorgang erhalten, so spricht man von einer *elastischen* Streuung; andernfalls ist die Streuung *inelastisch*, und ein Teil der kinetischen Energie geht in die Anregung der Streupartner über. So kann ein Elektron bei der Streuung an einem Atom dieses anregen; ebenso kann ein Neutron einen Atomkern im Grundzustand in einen angeregten Zustand versetzen, sofern die kinetische Energie des Neutrons ausreichend

⁷Die klassische Physik umfasst die Teilgebiete der Physik, welche ohne Konzepte der modernen Physik, vor allem der Quantisierung, auskommen.

gross ist. In einem erweiterten Sinne kann auch ein Absorptionsvorgang, bei welchem das Projektil verschwindet, oder noch allgemeiner jegliche Reaktion als Streuvorgang gedeutet werden; die ursprünglich vorhandenen Teilchen werden in neue Zustände 'hineingestreut'.

Bei der klassischen Streuung von *Massepunkten* an einer starren und unendlich harten Kugel mit Radius R ist, wie man sich leicht vorstellt, der totale Streuquerschnitt durch den Querschnitt πR^2 der Kugel gegeben. Streut man aber ein elektrisch geladenes Teilchen - beispielsweise ein Elektron - an einem anderen geladenen Teilchen - beispielsweise an einem Proton - so erweist sich der totale Streuquerschnitt als unendlich gross, denn schliesslich werden *alle* Teilchen durch ein elektrisches Feld unendlicher Reichweite gestreut. Selbst wenn sich die Wechselwirkung zwischen den aneinander streuenden Teilchen mit wachsendem Abstand r viel schneller abschwächt als im Falle der elektromagnetischen Wechselwirkung, wie dies beispielsweise bei der manchmal durch das exponentiell abfallende Yukawa-Potential $\sim e^{-\kappa r}/r$ phänomenologisch beschriebenen Kernkraft der Fall ist, so muss man im Rahmen einer *klassischen* Betrachtung schliessen, dass die Streurate und der totale Streuquerschnitt unendlich gross sind, da unendlich viele Teilchen zwar nur marginal, aber letztlich immer am Zielkern gestreut werden.

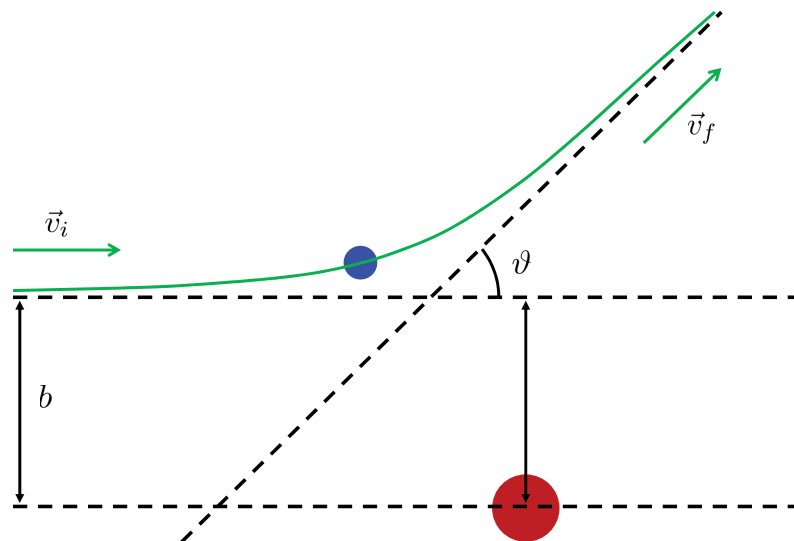


Abbildung 10: Ein von links einfallendes Projektil (blau) streut mit Impaktparameter b an einem Targetteilchen (rot) und wird um einen asymptotischen Winkel ϑ abgelenkt. Tatsächlich muss ein solcher klassischer Vorgang entgegen der Darstellung nicht planar ablaufen. Relevant sind letztlich die vorgängig experimentell präparierten Ausgangszustände der streuenden Teilchen und deren Endzustände, welche nach hinreichend langer Zeit beobachtet und durch diverse Parameter wie Streuwinkel, Geschwindigkeiten oder Spinvariablen et cetera beschrieben werden.

Im Rahmen einer *klassischen* Betrachtung stellt man sich ja im Sinne der Forderungen (201) und (202) eine unendlich ausgedehnte Verteilung von Projektilen vor, welche in grosser Ferne mit fester Projektilsdichte ρ_i und Geschwindigkeit \vec{v}_i auf das Targetteilchen zuströmt, wobei einem *einzelnen* Projektil jeweils ein sogenannter *Impaktparameter* (oder *Stossparameter*) zugeordnet werden kann. Als Impaktparameter bezeichnet man dabei den minimalen Abstand der jeweiligen Stosspartner, der sich ergeben würde, wenn zwischen den beiden Teilchen keine Kraft wirken würde und sie sich daher auf geradlinigen Bahnen passieren würden (siehe Abbildung 10).

Bei Streuprozessen besteht aber ein bemerkenswerter Unterschied zwischen dem klassischen und dem quanten-

mechanischen Fall. Die quantenmechanische Analyse der Streuung an einem Yukawa-Potential ergibt, dass der totale Streuquerschnitt tatsächlich *endlich* ist, trotz der unendlichen Reichweite des Potentials und der durch das Potential vermittelten Kräfte. Interferenzeffekte sorgen dafür, dass in der Quantenmechanik der totale Yukawa-Streuquerschnitt endlich und nicht mit einer *Infrarotdivergenz* wie im Falle der Coulomb-Wechselwirkung behaftet ist.

Deckt nun ein Detektor fern des Reaktionszentrums einen (für die theoretische Diskussion infinitesimalen) Raumwinkel $d\Omega$ ab, so kann die Anzahl der Teilchen gezählt werden, welche pro Zeiteinheit in diesen Raumwinkel hinein gestreut werden.

Damit gelangt man zum Begriff des *differentiellen Wirkungsquerschnitts* $d\sigma/d\Omega$. Bei der naheliegenden Wahl von Kugelkoordinaten, deren Ursprung durch die initiale Lage des Targetteilchens gegeben ist und deren Polachse in Richtung der initialen Projektilgeschwindigkeit \vec{v}_i durch den Ursprung verläuft, ist die Streurichtung der Projektile durch den Polarwinkel ϑ und einen Azimutwinkel φ , wie in Abbildung 9 dargestellt, gegeben. Die Anzahl der pro Zeiteinheit in einen infinitesimalen Raumwinkel $d\Omega$ in Richtung ϑ und φ gestreuten Teilchen ist dann

$$R(\vartheta, \varphi)d\Omega = I \frac{d\sigma}{d\Omega} d\Omega. \quad (203)$$

Dabei muss man sich immer im Klaren darüber sein, dass die Streuwinkel ϑ und φ die Flugrichtung gestreuter Teilchen nach *unendlich langer Zeit* charakterisieren.

Die sogenannte *Rutherford'sche Streuformel*

$$\frac{d\sigma_s^R}{d\Omega} = \frac{1}{16} \left(\frac{qQ}{4\pi\epsilon_0 E} \right)^2 \frac{1}{\sin^4(\vartheta/2)} = \frac{d^2}{16} \sin^{-4}(\vartheta/2) \quad (204)$$

beschreibt den klassischen nicht-relativistischen differentiellen Wirkungsquerschnitt der Streuung eines punktförmigen, strukturlosen Teilchens der elektrischen Ladung q und kinetischen Energie E an einer unendlich schweren, ruhenden und unbeweglichen Punktladung Q gleichen Vorzeichens im Ursprung des Kugelkoordinatensystems, wobei die auf das Projektil ausgeübte Kraft $\vec{F}_C = -\vec{\nabla}V_C$ durch das Coulomb-Potential $V_C(r) = qQ/(4\pi\epsilon_0 r)$ im SI-System bestimmt ist. Dabei ist ϵ_0 die Dielektrizitätskonstante im Vakuum und r der Abstand des Projektils vom Target im Koordinatenursprung. Es wird vorausgesetzt, dass das beschleunigte Projektil nicht strahlt. d ist der kleinste Abstand zwischen den repulsiven Streupartnern, welcher bei einem zentralen Stoss mit Impactparameter $b = 0$ erreicht wird (der '*closest approach*').

Durch Streuexperimente von α -Teilchen - also Heliumatomkernen, welche sich jeweils aus zwei Protonen und zwei Neutronen zusammensetzen - an dünnen Goldfolien konnte der bedeutende neuseeländische Physiker Ernest Rutherford zusammen mit den Kollegen Hans Geiger und Ernest Marsden schliesslich im Jahre 1911 zeigen, dass Goldatome einen kleinen schweren und elektrisch geladenen Kern besitzen müssen, welcher von viel leichteren Elektronen umgeben ist. α -Teilchen, welche von einer Radium-226-Quelle mit einer Geschwindigkeit von etwa 15000km/s emittiert werden, wurden von Rutherford et al. als gerichteter Teilchenstrahl auf Goldatome geschossen. α -Teilchen werden durch die viel leichteren Elektronen in einer Goldfolie kaum von ihrer Bahn abgelenkt, jedoch kommt es vor, dass einige der α -Projektile durch die kleinen, aber schwereren und elektrisch geladenen Goldatomkerne in guter Näherung gemäss der Rutherford'schen Streuformel (204) aus ihrer ursprünglichen Bahn herausgestreut werden. Solche α -Teilchen konnte Rutherford auf einem Zinksulfid-Fluoreszenzschirm nachweisen und dadurch für den Goldatomkernradius eine obere Grenze von $3.4 \cdot 10^{-14}\text{m}$ angeben, wobei zur damaligen Zeit bereits bekannt war, dass sich der Atomradius in der Grössenordnung von 10^{-10}m bewegt. Die Erkenntnisse Rutherfords unterstützten das atomare 'Planeten-Modell' des japanischen Physikers Nagaoka Hantarō aus dem Jahre 1904 und ebneten schliesslich auch den Weg zu Niels Bohrs Atommodell (1913).

Tatsächlich waren die von Geiger, Marsden und Rutherford verwendeten Goldfolien lediglich einige 0.0001cm dick, wodurch eine *Mehrfachstreuung* der α -Teilchen beim Durchgang durch die Goldfolie an zwei oder mehreren Goldatomkernen sehr unwahrscheinlich war.

Nun ist ja der totale Rutherford-Streuquerschnitt

$$\int_{4\pi} d\Omega \frac{d\sigma_s^R}{d\Omega} = \int_0^{2\pi} d\varphi \int_0^\pi d\vartheta \sin\vartheta \frac{d\sigma_s^R}{d\Omega} = \frac{\pi}{8} \left(\frac{qQ}{4\pi\epsilon_0 E} \right)^2 \int_0^\pi d\vartheta \frac{\sin\vartheta}{\sin^4(\vartheta/2)} = \infty \quad (205)$$

wegen des divergenten Verhaltens des differentiellen Rutherford-Streuquerschnitts für $\vartheta \rightarrow 0$ unendlich gross; beobachtet man allerdings nur die Teilchen, welche um einen Winkel $\vartheta > \vartheta_{min}$ aus der Einfallsrichtung gestreut werden, so erhält man nach kurzer Rechnung das vom Minimalwinkel ϑ_{min} abhängige endliche Resultat

$$\sigma_s^R(\vartheta > \vartheta_{min}) = 2\pi \int_{\vartheta_{min}}^\pi d\vartheta \sin\vartheta \frac{d\sigma_s^R}{d\Omega} = \frac{\pi}{4} d^2 \cot^2(\vartheta_{min}/2). \quad (206)$$

Berechnet man damit den integralen Rutherfordschen Streuquerschnitt für α -Teilchen mit einem minimalen Streuwinkel $\vartheta_{min} = 1^\circ = \pi/180$, so ergibt sich unter Verwendung der entsprechenden Ladungen des α -Teilchens (zwei Elementarladungen, $q = 2 \cdot 1.602176634 \cdot 10^{-19} \text{C}$), des Goldatomkerns ($Q = 79 \cdot 1.602176634 \cdot 10^{-19} \text{C}$) sowie der α -Energie ($E \simeq 4.7843 \text{MeV} = 7.665 \cdot 10^{-13} \text{J}$) der Wert $\sigma_s^R(\vartheta > 1^\circ) \simeq 2.33 \cdot 10^{-23} \text{m}^2$. Die minimale Annäherung d beträgt dabei $4.755 \cdot 10^{-14} \text{m} = 47.55 \text{fm}^8$.

Eine in der Atom-, Kern- und Teilchen-Physik oft zur Angabe von Wirkungsquerschnitten verwendete Flächeneinheit ist das *Barn* mit dem Einheitenzeichen b; es liegt in der Grössenordnung typischer Wirkungsquerschnitte von Atomkernen und ist durch $1\text{b} = 10^{-24} \text{cm}^2 = (10 \text{fm})^2 = 10^{-28} \text{m}^2$ definiert.

Damit ergibt sich also ein integrierter Streuquerschnitt $\sigma_s^R(\vartheta > 1^\circ)$ von 233000b; ein Goldatomkern präsentiert sich den α -Teilchen als Zielscheibe dieser Fläche, und jedes Teilchen, welches diese Zielscheibe passiert, wird um mindestens ein Grad aus seiner initialen Bewegungsrichtung abgelenkt. Die tatsächliche Querschnittsfläche eines Goldatomkerns mit einem ungefähren Radius $r_{Au} \simeq 7 \cdot 10^{-15} \text{m} = 7 \text{fm}$ ist mit $\pi \cdot r_{Au}^2 \simeq 1.5\text{b}$ bedeutend kleiner als $\sigma_s^R(\vartheta > 1^\circ)$, doch erstreckt sich ja das Coulomb-Feld des Goldatomkerns über einen beliebigen Raumbereich.

Eine interessante Beobachtung ergibt sich aus der Berechnung des integrierten Streuquerschnitts der *rückwärts-gestreuten* α -Teilchen

$$\sigma_s^R(\vartheta > 90^\circ = \pi/2) = \frac{\pi}{4} d^2 \tan^{-2} \frac{\pi}{4} = \frac{\pi}{4} d^2 \simeq 17.76\text{b}. \quad (207)$$

Diese Trefferfläche ist von der Grössenordnung d^2 und damit bedeutend kleiner als $\sigma_s^R(\vartheta > 1^\circ)$. Immerhin können α -Teilchen von einer Goldfolie abprallen, und diese Beobachtung macht klar, dass feste Materie wie Gold schwere kleine Kerne enthält und nicht als homogene 'puddingartige' Masse aufzufassen ist, welche α -Teilchen lediglich abbremsen würde. Rutherford verglich den erstaunlichen Vorgang der Rückwärtsstreuung mit dem Zurückprallen einer auf ein Stück Seidenpapier abgefeuerten 15-Zoll-Granate.

Als praktisches Beispiel sollen schliesslich kurz die wichtigsten Reaktionen *thermischer* Neutronen (n) mit einem Uran-235-Kern (^{235}U) und die quantitative Beschreibung durch entsprechende Wirkungsquerschnitte betrachtet werden.

⁸ 1fm = 10^{-15}m = 1 Femtometer = 1 Fermi. Aus dem Dänischen 'femten' = 'fünfzehn' bzw. nach Enrico Fermi (1901-1954), Nobelpreis für Physik 1938.

Die bei einer Kernspaltung von ^{235}U oder anderen spaltbaren Nukliden erzeugten Neutronen werden mit einer Geschwindigkeit der Grössenordnung 10000km/s von den beiden meist verschieden grossen Kernbruchstücken abgedampft. Thermalisiert man diese Neutronen anschliessend durch wiederholte Streuvorgänge an Atomkernen eines *Moderators* bei Raumtemperatur (wie beispielsweise Grafit oder Wasser bei $T = 293\text{K}$), welcher bei diesem Vorgang die Neutronen selbst in nur geringem Masse absorbiert, so gewinnt man schliesslich Neutronen, deren typische Geschwindigkeit durch die sogenannte *Westcott-Geschwindigkeit* $v_W = 2.2\text{km/s}$ charakterisiert wird. Neutronen dieser Geschwindigkeit besitzen eine kinetische Energie $E_{kin,n}$ von $4 \cdot 10^{-21}\text{J} = 0.025\text{eV}$, was dem Produkt der Boltzmann-Konstante $k = 1.380649 \cdot 10^{-23}\text{J} \cdot \text{K}^{-1}$ mit der Raumtemperatur T entspricht. Während sich die Wirkungsquerschnitte schneller Neutronen mit ^{235}U als ruhendes Target im Bereich weniger Barn befinden, sind thermische Neutronen bedeutend reaktionsfreudiger. Der Spaltquerschnitt (oder Fissionsquerschnitt) σ_f für Neutronen mit Westcott-Geschwindigkeit beträgt nämlich etwa 586b . In Abbildung 11 ist dieser Wirkungsquerschnitt als grösste Fläche mittlerer Graustufe angedeutet.

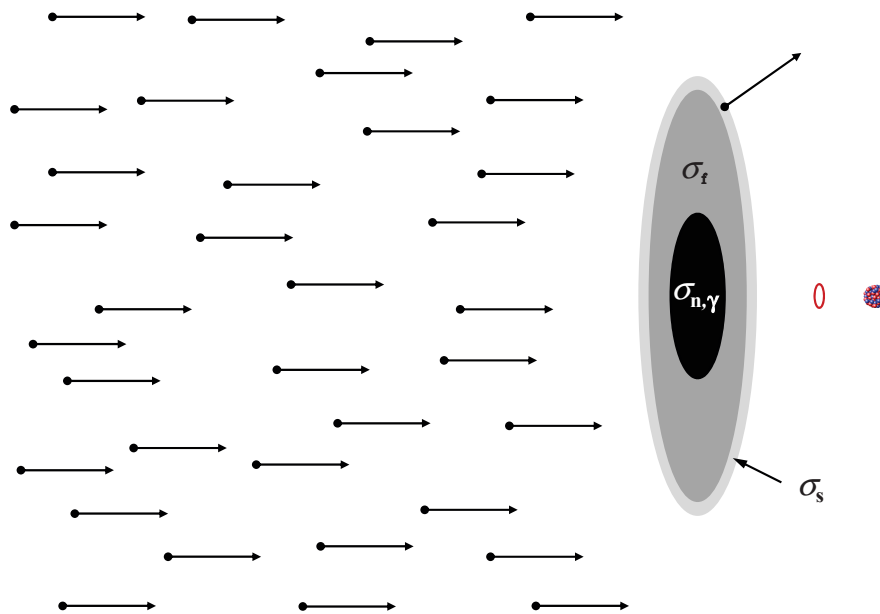


Abbildung 11: Wechselwirkungen von Neutronen mit einem ^{235}U -Kern. Die Zahl der pro Flächen- und Zeiteinheit den Kern passierenden Neutronen wird durch die *Neutronenstromdichte* I_n beschrieben. Die *totale Streurrate* $\dot{N}_s = I_n \cdot \sigma_s$ gibt die Zahl der pro Zeiteinheit unter einem beliebigen *nicht-verschwindenden* Winkel gestreuten Neutronen an.

Es besteht allerdings auch die Möglichkeit, dass ein Neutron von einem Uran-235-Kern ohne Spaltung eingefangen wird. Der so entstehende Uran-236-Zwischenkern ist angeregt und gibt seine Anregungsenergie durch Emission von γ -Quanten ab, weshalb der in Abbildung 11 schwarz dargestellte sogenannte thermische *Einfangquerschnitt* $\sigma_{n,\gamma}$ von etwa 98b mit den entsprechenden Symbolen indiziert ist. Der schmale hellgraue Rand in Abbildung 11 symbolisiert schliesslich den (totalen) Neutronenstreuquerschnitt σ_s von rund 15b für thermische Neutronen. Weitere Querschnitte sind im thermischen Bereich kaum erwähnenswert.

In Abbildung 12 ist die Energieabhängigkeit der in vielerlei Hinsicht wichtigsten Wirkungsquerschnitte für Neutronen an ^{235}U dargestellt. Erwähnenswert ist, dass sich der Wert des elastischen Streuquerschnitts über einen weiten Energiebereich von etwa 10^{-4}eV bis 0.1MeV vorwiegend zwischen 10b und 20b bewegt, im thermischen Bereich also bedeutend kleiner ist als der Einfang- und der Spaltquerschnitt. Dieses Verhältnis ändert sich aber bei Energien im MeV-Bereich. Die Tatsache, dass Neutronen in einer überkritischen Ansammlung

von Uran-235 oder Plutonium-239 von den Atomkernen vorwiegend gestreut werden und nicht ganz ungehindert aus der Anordnung entkommen können hat einen massgeblichen Einfluss auf die Grösse der kritischen Masse.

Bei geringen thermischen Neutronenenergien verhalten sich die in Abbildung 12 dargestellten Reaktionsquerschnitte näherungsweise gemäss einem sogenannten $1/v$ -Gesetz; je langsamer ein Neutron auf einen Uran-235-Kern trifft, desto länger hat es Zeit, eine Spaltung oder einen Einfang auszulösen, und entsprechend verhalten sich die Reaktionsquerschnitte σ_f und $\sigma_{n,\gamma}$. Bei höheren Energien im Bereich von etwa 1eV bis einige 1000eV lassen sich zahlreiche Einfangresonanzen beobachten, bei welchen in einem jeweils sehr engen Resonanz-Energiebereich die entsprechenden Querschnitte hohe Werte erreichen. Das komplexe Verhalten der Querschnitte in diesem Bereich spiegelt die komplexe Struktur des grossen, aus 92 Protonen und 143 Neutronen zusammengesetzten Kerns wider, welcher als oszillationsfähiges quantenmechanisches System resonant auf die äussere Anregung durch die einfallende Neutronenwellenfunktion reagieren kann.

Wiederum sind die zu beobachtenden Wirkungsquerschnitte grösser als die Querschnittfläche der Urankerns als kompakte Nukleonenkugel; allerdings sind die angegebenen Querschnitte wegen der kurzen Reichweite der Kernkräfte und der Ladungsfreiheit des Neutrons endlich.

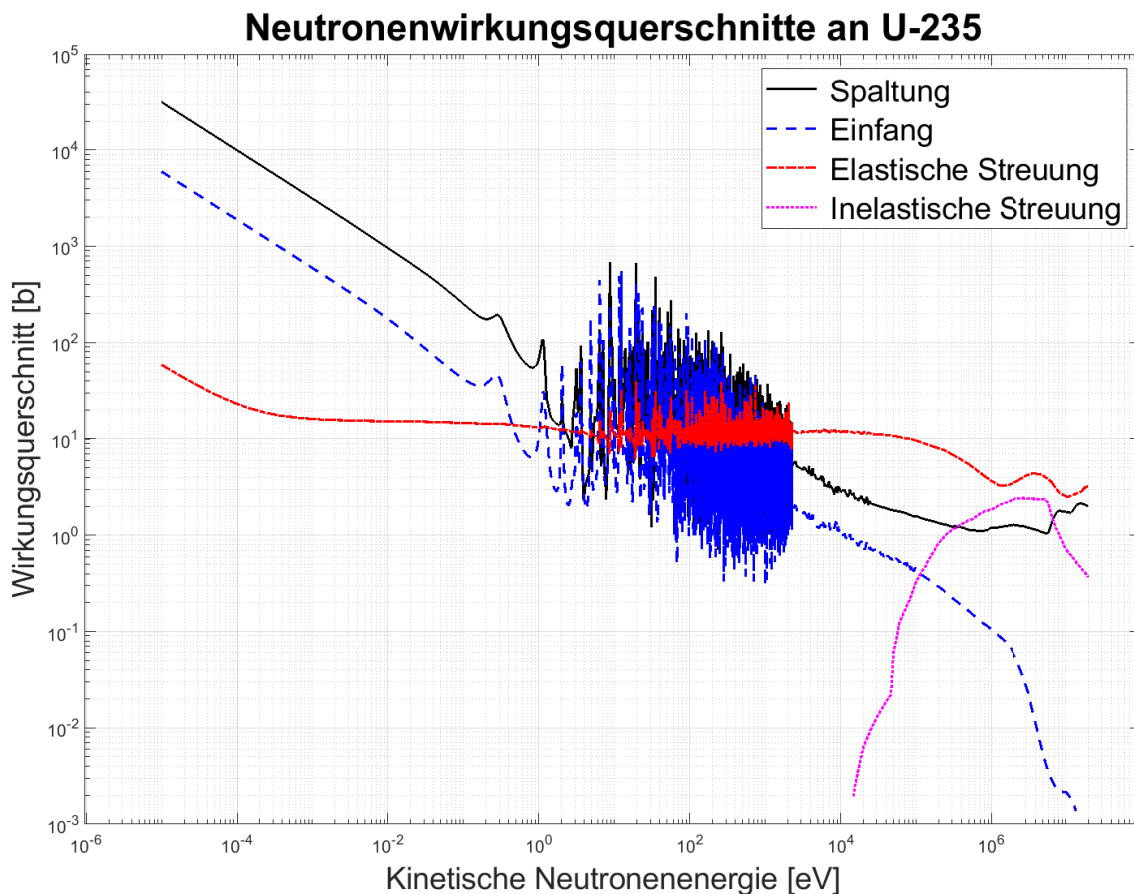


Abbildung 12: Energieabhängigkeit wichtiger Neutronenwirkungsquerschnitte.

Die bisherigen Betrachtungen machten durchwegs vom klassischen Teilchenbild Gebrauch. In der quantenmechanischen Behandlung von Streuprozessen werden einzelne Projektile schliesslich durch Wellenfunktionen repräsentiert, welche in grosser Entfernung entgegen der Einfallsrichtung der Projektile einen einheitlichen lokalen Impulserwartungswert aufweisen.

Der Thomson-Streuquerschnitt

Zur Illustration des Streuquerschnittbegriffs wollen wir durch eine vereinfachte klassische Betrachtung den Streuvorgang einer ebenen elektromagnetischen Welle beziehungsweise mit dieser Welle assoziierter niederenergetischer Photonen an einer Punktladung - einem 'Elektron' - untersuchen.

Um den entsprechenden klassischen, nach dem Entdecker der Elektrons Joseph John Thomson benannten *Thomson-Streuquerschnitt* (oder *Thomson-Wirkungsquerschnitt*) zu berechnen, betrachten wir eine ebene elektromagnetische Welle, wobei das kartesische Koordinatensystem mit Bedacht so gewählt sei, dass sich die Welle in x -Richtung bewegt, das elektrische Feld $\vec{E} = (E_x, E_y, E_z)$ in y -Richtung oszilliert und das magnetische Feld $\vec{B} = (B_x, B_y, B_z)$ in z -Richtung.



Abbildung 13: Ein geladenes Teilchen (Elektron) im Feld einer ebenen elektromagnetischen Welle beginnt durch die Kraftwirkung des \vec{E} -Feldes zu oszillieren. Die Kraftwirkung des magnetischen Feldes kann dabei in erster Näherung vernachlässigt werden. Als beschleunigte Ladung beginnt das geladene Teilchen wiederum elektromagnetische Energie abzustrahlen.

Aus den Maxwell-Gleichungen folgt das raumzeitliche Verhalten des elektromagnetischen Feldes im oben beschriebenen Szenario, wobei in der Folge irrelevante Phasenterme vernachlässigt seien

$$E_{x,z} \equiv 0, \quad E_y(x, y, z, t) = \hat{E} \sin(kx - \omega t), \quad B_{x,y} \equiv 0, \quad B_z(x, y, z, t) = \frac{\hat{E}}{c} \sin(kx - \omega t) \quad (208)$$

mit $\omega = kc$, wobei c die Lichtgeschwindigkeit ist. $2\pi/k$ ist natürlich die Wellenlänge und ω die Kreisfrequenz der Welle. Man überprüft sofort, dass beispielsweise das Induktionsgesetz $\vec{\nabla} \times \vec{E} = -\dot{\vec{B}}$ erfüllt ist.

Die mit $\sim \sin^2(kx - \omega t)$ variierende Energiedichte w des Feldes ist mit $1/\mu_0 = \epsilon_0 c^2$

$$w(x, y, z, t) = \frac{1}{2} \epsilon_0 \vec{E}^2(x, y, z, t) + \frac{1}{2} \mu_0^{-1} \vec{B}^2(x, y, z, t) = \epsilon_0 \vec{E}^2(x, y, z, t)^2, \quad (209)$$

und der räumliche Mittelwert der elektromagnetischen Energiedichte ergibt sich wegen

$$\frac{1}{2\pi} \int_0^{2\pi} dx \sin^2(x) = \frac{1}{2} \quad (210)$$

sogleich zu

$$\bar{w} = \frac{1}{2} \epsilon_0 \hat{E}^2. \quad (211)$$

Die entsprechende mittlere Energiestromdichte, welche in x -Richtung durch den Raum strömt, ist $j = \bar{w}c$, und da ein Photon die Energie $\hbar\omega$ besitzt, strömen also pro Flächen- und Zeiteinheit

$$I_\gamma = \frac{j}{\hbar\omega} = \frac{c\epsilon_0\hat{E}^2}{2\hbar\omega} \quad (212)$$

Photonen in x -Richtung durch den Raum.

Durch das elektrische Feld wird das Elektron in y -Richtung in Schwingung versetzt, und im nicht-relativistischen Fall kann man zur Beschreibung der Auslenkung $y(t)$ des Elektrons der Ruhemasse m_e und mit der Elementarladung e_0 die Bewegungsgleichung

$$m_e\ddot{y}(t) = F_{el} = -e_0\hat{E}\sin(\omega t + \varphi_0), \quad (213)$$

verwenden, wobei die für die vorliegende Betrachtung irrelevante Phase φ_0 von der Position des Elektrons und dem zeitlichen Beginn der Feldeinwirkung auf das Elektron abhängt. Abgesehen von einem homogenen Anteil in der Lösung der Differentialgleichung (213) folgt als partikuläre Lösung

$$y(t) = \frac{e_0\hat{E}}{m\omega^2}\sin(\omega t + \varphi_0), \quad (214)$$

das Elektron oszilliert also mit der Kreisfrequenz ω und einer Amplitude $e_0\hat{E}^2/(m\omega^2)$. In der Elektrodynamik-Vorlesung leitet man her, dass eine mit a beschleunigte (nicht-relativistische) Ladung q elektromagnetische Strahlung mit einer Leistung

$$P_{el} = \frac{q^2a^2}{6\pi\epsilon_0c^3} \quad (215)$$

abgibt. Die mittlere quadratische Beschleunigung $\langle a^2 \rangle$ des betrachteten Elektrons beträgt $\frac{1}{2}e_0^2\hat{E}^2/m^2$, die zeitlich gemittelte abgestrahlte Leistung also

$$\bar{P}_{el} = \frac{e_0^4\hat{E}^2}{12\pi\epsilon_0m^2c^3}, \quad (216)$$

entsprechend emittiert das Elektron

$$\dot{N}_\gamma = \frac{e_0^4\hat{E}^2}{12\pi\epsilon_0\hbar\omega m^2c^3} \quad (217)$$

Photonen der Frequenz $\nu = \omega/(2\pi)$ pro Zeiteinheit. Es folgt der Thomson-Wirkungsquerschnitt

$$\sigma_{Th} = \frac{\dot{N}_\gamma}{I_\gamma} = \frac{8\pi}{3} \frac{e_0^4}{16\pi^2\epsilon_0^2m^2c^4} = \frac{8\pi}{3} \left(\frac{e_0^2}{4\pi\epsilon_0mc^2} \right)^2 = \frac{8\pi}{3} r_e^2 \simeq 0.665 \cdot 10^{-28} \text{m}^2 = 0.665 \text{b}, \quad (218)$$

wobei r_e der *klassische Elektronenradius* ist. Der Ausdruck für den Thomson-Wirkungsquerschnitt enthält die Plancksche Konstante nicht; er ist klassischer Natur. Das Elektron erscheint dem Photon also gewissermassen als eine Zielscheibe mit einer linearen Ausdehnung in der Grössenordnung von r_e .

Die Thomson-Streuung entspricht dem niederenergetischen Grenzfall der Compton-Streuung, deren differenzieller Streuquerschnitt für unpolarisierte Photonen in der Literatur durch die Klein-Nishina-Formel beschrieben wird. Diese ergibt sich aus quantenelektrodynamischen Rechnungen in tiefster nicht-trivialer Ordnung. Bei höheren Energien überträgt ein Photon im Allgemeinen einen merklichen Impuls auf das Elektron; auch im klassischen Fall lässt sich die auf das Elektron wirkende Kraft einer elektromagnetischen Welle präziser untersuchen, wenn berücksichtigt wird, dass das magnetische Feld der Welle auf das oszillierende Elektron eine (Lorentz-)Kraft ausübt.

Matrix-Lie-Gruppen und Generatoren

Definition 14 Eine bezüglich der $GL(n, \mathbb{K})$ -Topologie abgeschlossene Untergruppe G der $GL(n, \mathbb{K})$ wird *Matrix-Lie-Gruppe* genannt. Dabei wird die $GL(n, \mathbb{K})$ -Topologie durch die \mathbb{K}^{n^2} -Standard-Topologie auf dem Raum der $n \times n$ -Matrizen $Mat(n, \mathbb{K})$ induziert.

Dass $G \subseteq GL(n, \mathbb{K})$ eine abgeschlossene Teilmenge ist, bedeutet genau folgendes: Ist $\{A_i \in Mat(n, \mathbb{K})\}_{i \in \mathbb{N}}$ eine in $Mat(n, \mathbb{K})$ konvergente Folge von Matrizen mit $A_i \in G$ für alle $i \in \mathbb{N}$ und ist $A = \lim_{i \rightarrow \infty} A_i$, dann gilt entweder $A \in G$ oder A ist nicht invertierbar. Der zweite Fall ist erlaubt, da nicht verlangt wird, dass G eine abgeschlossene Teilmenge von $Mat(n, \mathbb{K})$ ist, sondern lediglich von $GL(n, \mathbb{K})$. Die $GL(n, \mathbb{K})$ induziert mit ihrer von $Mat(n, \mathbb{K})$ geerbten Topologie ja eine Topologie auf jeder ihrer Teilmengen. Tatsächlich ist $GL(n, \mathbb{K})$ selbst nicht abgeschlossen (sondern offen) in der Topologie von $Mat(n, \mathbb{K})$.

Definition (14) lässt auch diskrete, *null-dimensionale* Matrix-Lie-Gruppen zu, deren Eigenschaften als 'kontinuierliche Gruppen' natürlich nichtig sind. Ein Beispiel für eine nicht-triviale diskrete Matrixgruppe stellt die $SL(2, \mathbb{Z})$ dar, also die Gruppe der 2×2 -Matrizen mit Determinante 1 und ganzzahligen Elementen.

$$Z = \begin{pmatrix} 5 & 8 \\ 8 & 13 \end{pmatrix}, \quad Z^{-1} = \begin{pmatrix} 13 & -8 \\ -8 & 5 \end{pmatrix} \quad (219)$$

sind solche Matrizen. Normalerweise wird gefordert, dass Lie-Gruppen mit diskreter Topologie abzählbar viele Elemente besitzen, und tatsächlich stellt sich heraus, dass die diskreten *Matrix-Lie-Gruppen* automatisch abzählbar sind. Führt man also auf der abzählbaren Menge der rationalen Zahlen die diskrete Topologie ein, so wird $(\mathbb{Q}, +)$ zu einer null-dimensionalen Lie-Gruppe, allerdings nicht zu einer Matrix-Lie-Gruppe im obigen Sinne.

Der mathematische Begriff der Lie-Gruppe ist noch allgemeiner als jener der Matrix-Lie-Gruppe. Eine Lie-Gruppe ist eine endlich-dimensionale glatte Mannigfaltigkeit, so dass die Gruppenverknüpfung und die Inversenbildung kompatibel mit dieser glatten Struktur sind. Allgemein fordert man von Mannigfaltigkeiten, dass ihre Topologie eine abzählbare Basis besitzt. Als Dimension der Lie-Gruppe bezeichnet man üblicherweise die Dimension der Gruppenmannigfaltigkeit. Die Gruppe der rationalen Zahlen $(\mathbb{Q}, +)$ mit der von \mathbb{R} induzierten Topologie ist eine topologische Gruppe, aber keine *ein-dimensionale* Lie-Gruppe, da sie sich nicht glatt kartografieren lässt. Nicht alle Lie-Gruppen lassen sich übrigens als Matrix-Lie-Gruppen darstellen. Auf diese spezielleren Aspekte der Theorie der Lie-Gruppen soll hier nicht eingegangen werden.

Generatoren

Die Definition einer Matrix-Lie-Gruppe $L \subseteq GL(n, \mathbb{K})$ schließt die Tatsache mit ein, dass sie eine bezüglich der durch die Standardtopologie von $Mat(n, \mathbb{K})$ in L induzierten Topologie offene Umgebung $U_1 \subset L$ der Eins besitzt. Lokal ist eine solche Gruppe also homöomorph zum $\mathbb{R}_{\mathbb{R}}^{\dim_{\text{top}} L}$ mit Standardtopologie. Die Gruppenverknüpfung und die Inversion von Matrizen besitzt in Matrix-Lie-Gruppen sogar automatisch eine 'analytische Struktur'. Es sei nun

$$\gamma : [-1, 1] \rightarrow U_1, \quad t \mapsto \gamma(t) \quad (220)$$

ein stetig differenzierbarer Weg in U_1 mit $\gamma(0) = 1_n$. Um der Anschauung Willen kann der Parameter t als Zeitvariable aufgefasst werden, $\gamma(t)$ bezeichnet dann die Position eines sich bewegenden Punktes innerhalb der Lie-Gruppenmannigfaltigkeit. Dann bezeichnet man einen Geschwindigkeitsvektor im Einselement

$$\left. \frac{d\gamma(t)}{dt} \right|_{t=0} = \dot{\gamma}(0) \quad (221)$$

als *Generator* von L . Die Menge aller möglichen Geschwindigkeitsvektoren bildet den Tangentialraum $T_1 L$ von L in der Eins. Wiederum bildlich gesprochen kann $T_1 L$ als jene $\dim_{\text{top}} L$ -dimensionale Hyperebene aufgefasst werden, welche sich im Einselement an die Gruppenmannigfaltigkeit 'anschmiegt'. Die topologische

Dimension des Tangentialraumes ist dabei gleich der topologischen Dimension der Lie-Gruppe.

Aus der Entwicklung

$$\gamma(dt) = \gamma(0) + \dot{\gamma}(0)dt + o(dt^2) = \mathbb{1}_n + \dot{\gamma}(0)dt + o(dt^2), \quad |dt| \ll 1 \quad (222)$$

lässt sich anschaulich folgern, dass sich Gruppenelemente in hinreichender Nähe zur Einheitsmatrix durch die Summe der Einheitsmatrix und einem 'kleinen' Generator $\dot{\gamma}(0)dt$ annähern lassen. Aus diesem Grunde bezeichnet man die Generatoren oft auch als *infinitesimale Transformationen*, auch wenn die Generatoren im Allgemeinen keine Elemente der Lie-Gruppe selbst sind. Die Vorstellung, dass sich zumindest ein Teil der Lie-Gruppe durch Hintereinanderausführen kleiner Transformationen $\sim \mathbb{1}_n + \dot{\gamma}(0)dt$ in der Nähe der Eins rekonstruieren lässt, motiviert wiederum die Verwendung des Begriffs des Generators, welcher bereits aus der Theorie der endlichen Gruppen in einem verwandten, wenn auch doch unterschiedlichen Sinn bekannt ist.

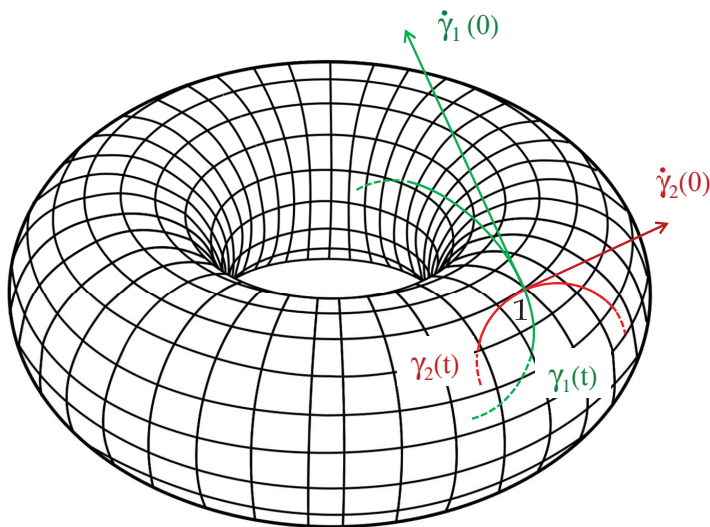


Abbildung 14: Sinnbildliche Darstellung von Geschwindigkeitsvektoren $\dot{\gamma}_{1,2}(0)$ im Einselement 1, welche den Tangentialraum T_1L einer Lie-Gruppe $L \sim U(1) \times U(1) \simeq T^2 \simeq S^1 \times S^1$ aufspannen.

Um diesen Sachverhalt durch Beispiele zu veranschaulichen, betrachten wir zuerst die unitäre Gruppe in einer Dimension

$$U(1) = \{z \in Mat(1, \mathbb{C}) \mid z^+ z = 1\} = \{z \in \mathbb{C} \mid \bar{z}z = 1\}, \quad (223)$$

welche also auch als die multiplikative Gruppe der komplexen Zahlen mit Betrag Eins erklärt werden kann. Die Gruppenmannigfaltigkeit ist der Einheitskreis in der (topologisch zweidimensionalen) komplexen Ebene und somit eindimensional. Die $U(1)$ spielt in der Quantenelektrodynamik als sogenannte $U(1)$ -Eichtheorie eine wichtige Rolle.

Als stetig differenzierbare Wege γ_α , welche zur Zeit $t = 0$ die Eins passieren, wählen wir

$$\gamma_\alpha : [-1, 1] \ni t \mapsto \gamma_\alpha(t) = e^{i\alpha t} \in U(1), \quad \alpha \in \mathbb{R}. \quad (224)$$

Offensichtlich gilt

$$\dot{\gamma}_\alpha(0) = i\alpha e^{i\alpha t} \Big|_{t=0} = i\alpha \in T_1U(1), \quad (225)$$

d.h. wie in Abbildung 15 dargestellt bildet der Raum der $U(1)$ -Generatoren, welchen wir von nun an auch mit $u(1)$ bezeichnen wollen, einen reellen eindimensionalen Vektorraum, welcher geometrisch gedacht seinen

Ursprung bei der 1 hat und durch die imaginäre Einheit als Basiselement aufgespannt wird: $u(1) = [i] = \mathbb{R}i$.

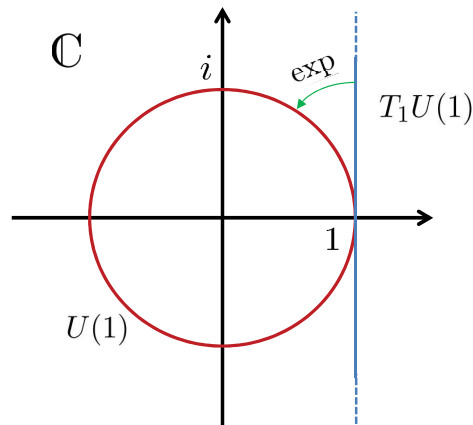


Abbildung 15: Der Generatorenraum $u(1)$ als Tangentialraum $T_1U(1)$ der eindimensionalen unitären Gruppe $U(1)$. Der Ursprung des reellen Vektorraumes $u(1)$ liegt in der 1 und wird durch die imaginäre Einheit i aufgespannt. Die Exponentialabbildung bildet Generatoren in $u(1)$ wieder in die Gruppenmannigfaltigkeit $U(1) \subset \mathbb{C}$ ab.

Eine zentrale, für spätere Betrachtungen sehr wichtige Beobachtung kann bereits an diesem ersten Beispiel illustriert werden: Das Exponentialbild eines $u(1)$ -Generators $i\alpha$ ist ein Element der Gruppe $U(1)$, es ist $\exp(i\alpha) = e^{i\alpha} \in U(1)$! Die Exponentialabbildung 'wickelt' den $u(1)$ (unendlich oft) auf die $U(1)$. α muss reell sein, $u(1)$ ist per Definition ein reeller Vektorraum. Ansonsten gälte mit $\alpha = \pi + i$ beispielsweise $e^{-1+i\pi} = -e^{-1} \notin U(1)$. Salopp kann man mengentheoretisch $\exp(u(1)) = U(1)$ schreiben.

Als weiteres Beispiel betrachten wir die zur $U(1)$ isomorphe spezielle orthogonale Gruppe in zwei reellen Dimensionen $SO(2)$. Die eindimensionale Gruppenmannigfaltigkeit dieser Gruppe liegt als Kreis eingebettet im topologisch vierdimensionalen reellen Vektorraum $Mat(2, \mathbb{R})$. Aus der Gruppendifinition folgt für

$$SO(2) \ni R = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad (226)$$

wegen

$$R^T R = \mathbb{1}_2 \rightarrow R^T = \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \frac{1}{\det R} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = R^{-1} \quad (227)$$

und $\det R = ad - bc = 1$ durch Koeffizientenvergleich sofort, dass sich wegen $a = d$ und $b = -c$ und somit $a^2 + b^2 = 1$ sämtliche R in der Form

$$\begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}, \quad \varphi \in [0, 2\pi) \quad (228)$$

darstellen lassen. Die Gruppe $O(2)$ enthält nebenbei bemerkt die zusätzlichen Matrizen der Form

$$\begin{pmatrix} \cos \varphi & \sin \varphi \\ \sin \varphi & -\cos \varphi \end{pmatrix} \in O^-(2) = \{R \in O(2) \mid \det R = -1\}, \quad \varphi \in [0, 2\pi). \quad (229)$$

Wiederum konstruieren wir den Tangentialraum bei der Eins $T_1SO(2) = so(2)$ mit Hilfe der Wegschar

$$\gamma_\alpha : [-1, 1] \ni t \mapsto \gamma_\alpha(t) = \begin{pmatrix} \cos(\alpha t) & -\sin(\alpha t) \\ \sin(\alpha t) & \cos(\alpha t) \end{pmatrix}, \quad \alpha \in \mathbb{R}. \quad (230)$$

Dann ist

$$\dot{\gamma}_\alpha(0) = \begin{pmatrix} 0 & -\alpha \\ \alpha & 0 \end{pmatrix} = -\alpha \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = -\alpha\epsilon, \quad (231)$$

d.h. die antisymmetrische Matrix ϵ spannt als Basisvektor den reellen Raum $so(2)$ auf. Wiederum zeigt sich das Wunder der Exponentialabbildung: Definiert man das Exponentialbild einer $n \times n$ -Matrix $X \in Mat(n, \mathbb{K})$ durch die auf ganz $Mat(n, \mathbb{K})$ konvergente Taylor-Reihe oder einen alternativen wohldefinierten Grenzwert

$$\exp X = e^X = \mathbf{1}_n + \sum_{k=1}^{\infty} \frac{X^k}{k!} = \lim_{n \rightarrow \infty} \left(\mathbf{1}_n + \frac{X}{n} \right)^n, \quad (232)$$

so ergibt sich

$$e^{-\alpha\epsilon} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & -\alpha \\ \alpha & 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} -\alpha^2 & 0 \\ 0 & -\alpha^2 \end{pmatrix} + \frac{1}{6} \begin{pmatrix} 0 & \alpha^3 \\ -\alpha^3 & 0 \end{pmatrix} + \dots = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}, \quad (233)$$

wiederum gilt die mengentheoretische Aussage $\exp(so(2)) = SO(2)$ (aber nicht $\exp(so(2)) = o(2)$).

Als vorerst letztes Beispiel betrachten wir die Generatoren der nicht-abelschen speziellen unitären Gruppe $SU(2)$, welche eine wichtige Rolle in der Spinphysik einnimmt. Die Gruppenmannigfaltigkeit ist dreidimensional, ebenso der Tangentialraum $su(2) = T_1SU(2)$ der Generatoren. Eine Basis von $su(2)$ erhalten wir durch die Geschwindigkeitsvektoren $\dot{\gamma}_{1,2,3}(0)$ der folgenden drei in der $SU(2)$ verlaufenden Wege (oder Trajektorien) $\gamma_{1,2,3}$ zur Zeit $t = 0$

$$\gamma_1(t) = \begin{pmatrix} \cos t & i \sin t \\ i \sin t & \cos t \end{pmatrix}, \quad \gamma_2(t) = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}, \quad \gamma_3(t) = \begin{pmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{pmatrix}. \quad (234)$$

Die Generatoren können mit Hilfe der Pauli-Matrizen $\sigma_{1,2,3}$ ausgedrückt werden

$$\dot{\gamma}_1(0) = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} = i\sigma_1, \quad \dot{\gamma}_2(0) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = i\sigma_2, \quad \dot{\gamma}_3(0) = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} = i\sigma_3. \quad (235)$$

Wiederum findet man durch explizites Nachrechnen, dass die Exponentialabbildung die $su(2)$ in die $SU(2)$ und sogar auf die $SU(2)$ abbildet. Mit dem Zahlentripel $\vec{\alpha} = (\alpha_1, \alpha_2, \alpha_3) \in \mathbb{R}^3$, $i\vec{\alpha} \cdot \vec{\sigma} = i\alpha_1\sigma_1 + i\alpha_2\sigma_2 + i\alpha_3\sigma_3$ und $\alpha = \sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2}$ ergibt sich

$$e^{i\vec{\alpha} \cdot \vec{\sigma}} = \mathbf{1}_2 \cos \alpha + i \frac{\vec{\alpha} \cdot \vec{\sigma}}{\alpha} \sin \alpha \in SU(2). \quad (236)$$

Nach dem Wunder der Exponentialabbildung lässt sich aber an diesem Beispiel ein weiteres Wunder beobachten: Der Kommutator zweier Generatoren ist wiederum ein Generator! Bekanntlich erfüllen die Pauli-Matrizen für $l, m, n = 1, 2, 3$ die Vertauschungsbeziehungen

$$[\sigma_l, \sigma_m] = 2i\epsilon_{lmn}\sigma_n \quad (237)$$

mit dem total antisymmetrischen Tensor ϵ in drei Dimensionen, welcher für paarweise gleiche Indizes wegen der Antisymmetrie verschwindet und ansonsten auf $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1 = -\epsilon_{213} = -\epsilon_{132} = -\epsilon_{321}$ normiert ist. Entsprechend gilt für die Basisgeneratoren $\dot{\gamma}_{1,2,3}(0)$

$$[i\sigma_l, i\sigma_m] = -2\epsilon_{lmn}(i\sigma_n), \quad (238)$$

woraus unmittelbar folgt, dass der Kommutator zweier beliebiger Generatoren als reelle Linearkombinationen der $\dot{\gamma}_{1,2,3}(0)$ wiederum ein Generator in der $su(2)$ ist. Die $su(2)$ ist unter der Kommutatorbildung als Verknüpfung abgeschlossen.

Generatorenräume als Lie-Algebren

Tatsächlich trifft die obige für die $SU(2)$ bzw. $su(2)$ gemachte Beobachtung bezüglich Abgeschlossenheit der Generatorenräume bezüglich Kommutatorbildung allgemein auf Matrix-Lie-Gruppen zu. Dies kann eingesehen werden, indem man in einer kontinuierlichen Lie-Gruppe $L \subseteq GL(n, \mathbb{K})$ zwei beliebig oft differenzierbare Wege $\alpha, \beta : \mathbb{R} \rightarrow L$ mit der Entwicklung

$$\alpha(t) = \mathbb{1}_n + Xt + Vt^2 \dots, \quad \beta(t) = \mathbb{1}_n + Yt + Wt^2 \dots, \quad X, Y, V, W \in Mat(n, \mathbb{K}) \quad (239)$$

und den Weg

$$\alpha(t)\beta(t)\alpha(t)^{-1}\beta(t)^{-1} = \mathbb{1}_n + (XY - YX)t^2 + \dots = \mathbb{1}_n + [X, Y]t^2 + \dots, \quad (240)$$

untersucht (Übung).

Die abstrakte Definition einer *Lie-Algebra* lautet nun wie folgt:

Definition 15 Ein Vektorraum \mathcal{L} über einem (beliebigen) Körper \mathbb{K} zusammen mit einer bilinearen Verknüpfungsabbildung $(X, Y) \mapsto [X, Y]$ (also einer \mathbb{K} -Algebra) heisst Lie-Algebra, wenn für alle $X, Y, Z \in \mathcal{L}$ gilt:

$$(AK) \quad [X, Y] = -[Y, X] \quad (\text{Antikommutativität})$$

$$(JI) \quad [X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0 \quad (\text{Jacobi-Identität})$$

$\mathcal{L}' \subset \mathcal{L}$ heisst Teilalgebra von \mathcal{L} , wenn $[X, Y] \in \mathcal{L}'$ für alle $X, Y \in \mathcal{L}'$. Eine Lie-Algebra \mathcal{L} heisst abelsch, wenn $[X, Y] = 0$ für alle $X, Y \in \mathcal{L}$.

Die Verknüpfung $[X, Y]$ ist nicht notwendigerweise ein durch ein vorgängig in \mathcal{L} definiertes Produkt XY bzw. YX gegebener Kommutator. Im vorliegenden Fall der Matrix-Lie-Algebren ist die Verknüpfung $[X, Y]$ aber durch den Kommutator

$$[X, Y] = XY - YX \quad (241)$$

gegeben, und es ist $\mathbb{K} = \mathbb{R}$. Es lässt sich sofort nachprüfen, dass der in Gleichung (241) definierte Kommutator die geforderte Jacobi-Identität erfüllt. Dieser Kommutator, welcher sowohl Matrizenprodukte wie auch eine Matrizendifferenz beinhaltet, ist zu unterscheiden vom Kommutator in der Gruppentheorie, welcher für zwei Elemente g und h einer Gruppe G durch die (einzige) Gruppenverknüpfung

$$[g, h] = ghg^{-1}h^{-1} \quad (242)$$

definiert wird. Ein solcher Kommutator taucht in Gleichung (240) auf.

Der Name 'Lie-Algebra' wurde im Jahre 1934 durch den Mathematiker Hermann Weyl eingeführt. Sophus Lie begründete seine Theorie der 'kontinuierlichen Gruppen' und fand die Abgeschlossenheit der 'infinitesimalen Transformationen' unter Kommutatorbildung im Jahre 1874 als ein krönendes Resultat seiner nervenzerrüttenden Forschungen.

Liesse man in $\mathcal{L} = T_1 L$ im Widerspruch zur ursprünglichen Konstruktion komplexe Skalare zu, so würde man die Matrix-Lie-Algebra *komplexifizieren*, ein Vorgang, welcher sich über die Exponentialabbildung auch auf die Lie-Gruppe selbst überträgt und zur Konstruktion einer eventuell grösseren Gruppe führt.

Ist nun $\{X_a, a = 1, \dots, n\} \subset \mathcal{L}$ eine Basis des reellen Vektorraums der Generatoren einer n -dimensionalen Lie-Gruppe L , so kann die durch den Kommutator erklärte Verknüpfung zweier Generatoren mit Hilfe von sogenannten Strukturkonstanten C_{ab}^c wie folgt geschrieben werden:

$$[X_a, X_b] = C_{ab}^c X_c, \quad C_{ab}^c = -C_{ba}^c. \quad (243)$$

Natürlich wird nach der Einsteinschen Summenkonvention über gleiche Indizes summiert. Verwendet man die *Jacobi-Identität*, so folgt daraus die Jacobi-Identität für die Strukturkonstanten

$$C_{ab}^e C_{ec}^d + C_{bc}^e C_{ea}^d + C_{ca}^e C_{eb}^d = 0. \quad (244)$$

Die Strukturkonstanten sind natürlich basisabhängig.

In der Physik ist es üblich, die in der mathematischen Literatur üblicherweise wie oben definierten Generatoren durch Multiplikation mit der imaginären Einheit i umzudefinieren. Dies hat den Vorteil, dass die Generatoren der orthogonalen Gruppen symmetrisch und die Generatoren der unitären Gruppen hermitesch werden und in vielen Fällen als Observable in quantenmechanischen Betrachtungen Eingang finden. Wie in den Gleichungen (237) und (238) angedeutet erhält die die Strukturkonstanten bezüglich einer Basis definierende Gleichung (243) eine leichte Modifikation. Mit 'physikalischen' Generatoren $\tilde{X}_a = \pm i X_a$ ergibt sich aus Gleichung (243)

$$[\tilde{X}_a, \tilde{X}_b] = \pm i C_{ab}^c \tilde{X}_c. \quad (245)$$

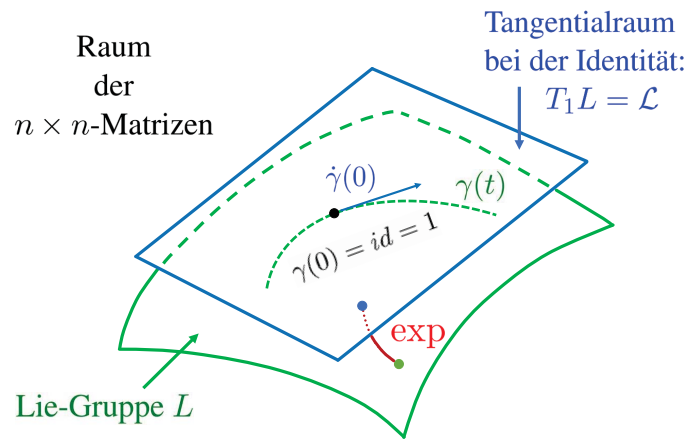


Abbildung 16: Sinnbild für den Generatorenraum bzw. die Lie-Algebra \mathcal{L} einer Matrix-Lie-Gruppe L als Tangentialraum T_1L an die Identität, welcher sich als Vektorraum aller Geschwindigkeitsvektoren $\dot{\gamma}(0)$ an die Identität von Wegen γ durch die Identität denken lässt. Die Exponentialabbildung bildet T_1L in die Gruppenmannigfaltigkeit ab, aber nicht zwingend *auf* diese; schliesslich sind in den meisten interessanten Fällen der Vektorraum T_1L und L nicht homöomorph.

Generatoren der SU(2) und der SO(3): Die Lie-Algebra su(2) bzw. so(3)

Wir führen Generatoren S_1, S_2 und S_3 für aktive Rotationen in SO(3) um die x^1 -, x^2 - und x^3 -Achse im $\mathbb{R}_\mathbb{R}^3$ mit Standardskalarprodukt ein: Eine Rotation um den Winkel δ um die x^3 -Achse wird beschrieben durch die Rotationsmatrix

$$R_\delta^3 = \begin{pmatrix} \cos \delta & -\sin \delta & 0 \\ \sin \delta & \cos \delta & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 - \delta^2/2! + \dots & -\delta + \delta^3/3! - \dots & 0 \\ \delta - \delta^3/3! + \dots & 1 - \delta^2/2! + \dots & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \delta \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + o(\delta^2) = \mathbf{1} + \delta S_3 + o(\delta^2) \simeq 1 + \delta S_3 \text{ für } \delta \ll 1, \quad (246)$$

entsprechend erhalten wir im allgemeinen Fall die Generatoren für Drehungen um die $x^{1,2,3}$ -Achsen

$$S_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & +1 & 0 \end{pmatrix}, \quad S_2 = \begin{pmatrix} 0 & 0 & +1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad S_3 = \begin{pmatrix} 0 & -1 & 0 \\ +1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (247)$$

welche mit dem vollständig antisymmetrischen Tensor $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1 = -\epsilon_{213} = -\epsilon_{132} = -\epsilon_{321}$ beziehungsweise $\epsilon_{lmn} = \frac{1}{2}(l-m)(m-n)(n-l)$ kompakt dargestellt werden können

$$(S_l)_{mn} = -\epsilon_{lmn}, \quad l, m, n = 1, 2, 3. \quad (248)$$

Die Generatoren erfüllen die Vertauschungsrelationen der abstrakten Drehimpulsalgebra

$$[S_l, S_m] = \epsilon_{lmn} S_n. \quad (249)$$

Ein Notationswechsel Mathematik \rightarrow Physik führt auf hermitesche Operatoren $\Sigma_l = iS_l$ bzw. $(\Sigma_l)_{mn} = -i\epsilon_{lmn}$

$$\Sigma_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & +i & 0 \end{pmatrix}, \quad \Sigma_2 = \begin{pmatrix} 0 & 0 & +i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}, \quad \Sigma_3 = \begin{pmatrix} 0 & -i & 0 \\ +i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (250)$$

mit

$$[\Sigma_l, \Sigma_m] = [iS_l, iS_m] = i\epsilon_{lmn}(iS_n) = i\epsilon_{lmn}\Sigma_n. \quad (251)$$

Diese Vertauschungsrelationen gelten analog für die *halben* Pauli-Matrizen. Mit

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (252)$$

ist nämlich

$$[\sigma_l, \sigma_m] = 2i\epsilon_{lmn}\sigma_n, \quad \left[\frac{\sigma_l}{2}, \frac{\sigma_m}{2}\right] = i\epsilon_{lmn}\frac{\sigma_n}{2}. \quad (253)$$

Durch explizite Rechnung überprüft man das 'Wunder der Exponentialabbildung':

exp(Generator) = Gruppenelement (Rotation)!

$$\exp(-i\alpha\Sigma_3) = \exp(\alpha S_3) = 1 + \alpha S_3 + \frac{\alpha^2}{2!} S_3^2 + \frac{\alpha^3}{3!} S_3^3 + \dots$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & -\alpha & 0 \\ +\alpha & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} \alpha^2 & 0 & 0 \\ 0 & \alpha^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \frac{1}{6} \begin{pmatrix} 0 & +\alpha^3 & 0 \\ -\alpha^3 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \dots$$

$$= \begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \in SO(3). \quad (254)$$

Als Verallgemeinerung erhält man die Rotationsmatrix für beliebige Rotation um eine Drehachse $\vec{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$ mit Drehwinkel $\alpha = |\vec{\alpha}| = (\alpha_1^2 + \alpha_2^2 + \alpha_3^2)^{1/2}$

$$R(\vec{\alpha}) = \exp(-i\vec{\alpha} \cdot \vec{S}) = \exp(\alpha_1 S_1 + \alpha_2 S_2 + \alpha_3 S_3). \quad (255)$$

Wir nutzen den Umstand aus, dass Generatoren der $SO(3)$ derselben abstrakte Lie-Algebra wie die Pauli-Matrizen angehören und betrachten als analoger Vorgang zur Rotation eines Vektors im euklidischen Raum die Drehung eines 'Spinors' im Spinor-Raum \mathbb{C}_C^2 . Wir bilden die spezielle unitäre Matrix

$$U(\vec{\alpha}) = \exp\left(-i\vec{\alpha} \cdot \frac{\vec{\sigma}}{2}\right), \quad (256)$$

die nach expliziter Rechnung (Übung!) gegeben ist durch

$$U(\vec{\alpha}) = \cos \frac{\alpha}{2} - i \frac{\vec{\alpha} \cdot \vec{\sigma}}{\alpha} \sin \frac{\alpha}{2} = \cos \frac{\alpha}{2} - i \hat{\alpha} \cdot \vec{\sigma} \sin \frac{\alpha}{2}, \quad \hat{\alpha} = \frac{\vec{\alpha}}{\alpha}. \quad (257)$$

Es gilt die (Strahl-)Darstellungseigenschaft

$$U(\vec{\gamma}) = U(\vec{\beta})U(\vec{\alpha}) \leftrightarrow R(\vec{\gamma}) = R(\vec{\beta})R(\vec{\alpha}). \quad (258)$$

Wir erläutern diese Eigenschaft durch ein Beispiel und betrachten zuerst eine Drehung um den Winkel ϑ um die x^2 - bzw. y -Achse im euklidischen Raum, dargestellt im Spinorraum durch

$$U_\vartheta = \exp(-i\vartheta\sigma_2/2) = \exp\left(\begin{array}{cc} 0 & -\vartheta/2 \\ \vartheta/2 & 0 \end{array}\right) = \left(\begin{array}{cc} \cos \vartheta/2 & -\sin \vartheta/2 \\ \sin \vartheta/2 & \cos \vartheta/2 \end{array}\right), \quad (259)$$

anschliessend eine Drehung um Winkel φ um x^3 - bzw. z -Achse

$$U_\varphi = \exp(-i\varphi\sigma_3/2) = \exp\left(\begin{array}{cc} -i\varphi/2 & 0 \\ 0 & i\varphi/2 \end{array}\right) = \left(\begin{array}{cc} e^{-i\varphi/2} & 0 \\ 0 & e^{i\varphi/2} \end{array}\right). \quad (260)$$

Diese Drehungen wenden wir auf einen in x^3 -Richtung ausgerichteten Spinor an und drehen diesen in Kugelkoordinaten in die ϑ - φ -Richtung

$$|\vartheta, \varphi\rangle = U_\varphi U_\vartheta \left(\begin{array}{c} 1 \\ 0 \end{array}\right) = U_\varphi \left(\begin{array}{c} \cos \vartheta/2 \\ \sin \vartheta/2 \end{array}\right) = \left(\begin{array}{c} \cos \frac{\vartheta}{2} e^{-i\varphi/2} \\ \sin \frac{\vartheta}{2} e^{i\varphi/2} \end{array}\right). \quad (261)$$

Zur Überprüfung der obigen Rechnungen berechnet man beispielsweise die Erwartungswerte der Spinkomponenten

$$\langle \vartheta, \varphi | \frac{\sigma_1}{2} | \vartheta, \varphi \rangle = \frac{1}{2} \sin \vartheta \cos \varphi, \quad \langle \vartheta, \varphi | \frac{\sigma_2}{2} | \vartheta, \varphi \rangle = \frac{1}{2} \sin \vartheta \sin \varphi, \quad \langle \vartheta, \varphi | \frac{\sigma_3}{2} | \vartheta, \varphi \rangle = \frac{1}{2} \cos \vartheta. \quad (262)$$

N.B.: Im Allgemeinen gilt

$$R(\vec{\alpha}) = \exp(\alpha_1 S_1 + \alpha_2 S_2 + \alpha_3 S_3) \neq \exp(\alpha_1 S_1) \exp(\alpha_2 S_2) \exp(\alpha_3 S_3). \quad (263)$$

Für $\vec{\alpha} = (0, 0, 2\pi)$ ist offensichtlich

$$R(\vec{\alpha}) = 1, \quad U(\vec{\alpha}) = -1! \quad (264)$$

Allgemeiner gilt

$$\alpha = |\vec{\alpha}| = 2\pi \rightarrow R(\vec{\alpha}) = 1, \quad U(\vec{\alpha}) = -1, \quad \alpha = |\vec{\alpha}| = 4\pi \rightarrow R(\vec{\alpha}) = 1, \quad U(\vec{\alpha}) = 1. \quad (265)$$

Eine Drehung um den Winkel 2π ändert das Vorzeichen des Spinors. Die Phase des Spinors ist quantenmechanisch nicht relevant, solange keine Selbstinterferenzeffekte eine Rolle spielen.

Die $SU(2)$ ist die (zweifache) universelle Überlagerungsgruppe der $SO(3)$: $SU(2)/\{\pm 1\} \cong SO(3)$. Die relativistische Verallgemeinerung der obigen Betrachtungen ($SL(2, \mathbb{C})/\{\pm 1\} \cong SO^+(1, 3)$) führt letztlich auch auf den (Majorana-)Spinor-Formalismus.

Generatoren der Lorentzgruppe $SO^+(1,3)$

Ein reiner passiver Lorentz-Boost $x'^{\mu} = \Lambda^{\mu}_{\nu} x^{\nu}$ der Raumzeitkoordinaten in x^1 -Richtung mit Geschwindigkeit $v_1 = \beta_1 c$ und Gammafaktor γ_1 lässt sich durch die Lorentz-Matrix

$$\Lambda(\beta_1) = \begin{pmatrix} \gamma_1 & -\gamma_1 \beta_1 & 0 & 0 \\ -\gamma_1 \beta_1 & \gamma_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \gamma_1 = \frac{1}{\sqrt{1-\beta_1^2}}, \quad \gamma_1^2 - \gamma_1^2 \beta_1^2 = 1 \quad (266)$$

darstellen, welche sich in erster Ordnung in β_1 durch

$$\Lambda(\beta_1) \simeq \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \beta_1 \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = 1 + \beta_1 N_1, \quad |\beta_1| \ll 1 \quad (267)$$

annähern lässt. Dabei ist N_1 also ein Generator für Boosts in x^1 -Richtung, er beschreibt die infinitesimale Abweichung der Raumzeitkoordinaten eines Ereignisses in einem infinitesimal in x^1 -Richtung geboosteten Inertialsystem IS' von den ursprünglichen Koordinaten im Inertialsystem IS .

Die volle Lorentztransformation ergibt sich wiederum durch Exponentiation des mit dem Boost-Parameter ξ_1 (der *Rapidity*) multiplizierten Generators

$$\exp(\xi_1 N_1) = \begin{pmatrix} +\cosh \xi_1 & -\sinh \xi_1 & 0 & 0 \\ -\sinh \xi_1 & +\cosh \xi_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \cosh(\xi_1) = \gamma_1. \quad (268)$$

Weitere Boost-Generatoren $N_{2,3}$ für die $x^{2,3}$ -Richtung folgen in völliger Analogie

$$N_2 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad N_3 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \quad (269)$$

die passiven Rotationsgeneratoren um die x^1 -, x^2 und x^3 -Achse sind gegeben durch

$$M_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & +1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 \end{pmatrix}, \quad M_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & +1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (270)$$

Zusammen spannen die obigen sechs Generatoren die reelle Lie-Algebra $so^+(1,3)$ auf, deren abstrakte Kommutationsrelationen mit Hilfe des total antisymmetrischen Tensors ϵ in drei Dimensionen beziehungsweise durch die $so(3)$ -Strukturkonstanten durch

$$[M_l, M_m] = -\epsilon_{lmn} M_n, \quad [N_l, N_m] = +\epsilon_{lmn} M_n, \quad [N_l, M_m] = -\epsilon_{lmn} N_n \quad (271)$$

dargestellt werden können. Die für die oben gewählte Matrizenbasis geltenden Vertauschungsrelationen in Gleichung (271) bleiben bei einer Ersetzung $N_{1,2,3} \rightarrow -N_{1,2,3}$ erhalten.

Die in den Kommutationsrelationen der Boost-Generatoren auftauchenden Rotationsgeneratoren spiegeln die Tatsache wieder, dass durch die Verknüpfung von Lorentz-Boosts Rotationen erzeugt werden können. Daher

erzeugen die Boosts im Gegensatz zu den Rotationsgeneratoren keine echte Untergruppe der eigentlichen orthochronen Lorentzgruppe; vielmehr erzeugen sie die ganze $SO^+(1, 3)$.

Die sogenannten *Casimir-Operatoren* $C_1 = \vec{M}^2 - \vec{N}^2 = M_1^2 + M_2^2 + M_3^2 - N_1^2 - N_2^2 - N_3^2$ und $C_2 = \vec{M} \cdot \vec{N} = M_1 N_1 + M_2 N_2 + M_3 N_3$ vertauschen mit allen Generatoren $N_{1,2,3}$ und $M_{1,2,3}$. Diese Operatoren gehören *nicht* zur Lie-Algebra $so^+(1, 3)$ der Lorentzgruppe, können aber in einer etwas abstrakteren Betrachtungsweise als Elemente der universellen einhüllenden Algebra $U(so^+(1, 3))$ aufgefasst werden.

Eine interessante Beobachtung im Zusammenhang mit den Vertauschungsrelationen (271) ergibt sich, wenn man mit Hilfe der Pauli-Matrizen σ_l die Matrizen

$$\tilde{M}_l = i \frac{\sigma_l}{2}, \quad \tilde{N}_l = i \tilde{M}_l = -\frac{\sigma_l}{2}, \quad l = 1, 2, 3 \quad (272)$$

einführt. Ausgehend von den wohlbekannten Vertauschungsrelationen der $su(2)$ -Lie-Algebra

$$\left[\frac{\sigma_l}{2}, \frac{\sigma_m}{2} \right] = i \epsilon_{lmn} \frac{\sigma_n}{2} \quad (273)$$

prüft man sofort nach, dass die Matrizen (272) den (271) völlig analogen Vertauschungsrelationen

$$[\tilde{M}_l, \tilde{M}_m] = -\epsilon_{lmn} \tilde{M}_n, \quad [\tilde{N}_l, \tilde{N}_m] = +\epsilon_{lmn} \tilde{M}_n, \quad [\tilde{N}_l, \tilde{M}_m] = -\epsilon_{lmn} \tilde{N}_n \quad (274)$$

genügen. Tatsächlich bilden die sechs *spurfreien* Matrizen $\sigma_{1,2,3}$ und $i\sigma_{1,2,3}$ eine Basis der reellen sechsdimensionalen Lie-Algebra $sl(2, \mathbb{C}) \simeq su(2) \oplus i \cdot su(2) \simeq \mathbb{C} \otimes su(2)$, also der Generatorenalgebra der Gruppe $SL(2, \mathbb{C})$. Offensichtlich sind die Lie-Algebren $so^+(1, 3)$ und $sl(2, \mathbb{C})$ isomorph; damit sind die Gruppen $SO^+(1, 3)$ und $SL(2, \mathbb{C})$ zwar *lokal* in einer Umgebung der Eins ($\mathbb{1}_4$ und $\mathbb{1}_2$) strukturell gleich, aber nicht global, da sich die $SL(2, \mathbb{C})$ ja wie früher beobachtet als universelle Überlagerungsgruppe der $SO^+(1, 3)$ topologisch von dieser unterscheidet.

Die Spurfreiheit der $SL(2, \mathbb{C})$ -Generatoren in $sl(2, \mathbb{C})$ ist eine direkte Konsequenz der die $SL(2, \mathbb{C})$ definierenden Determinantenbedingung $\det A = 1$ für $A \in SL(2, \mathbb{C})$:

$$SL(2, \mathbb{C}) \ni A = \begin{pmatrix} 1 + a_1 & a_3 \\ a_4 & 1 + a_2 \end{pmatrix} \rightarrow$$

$$\det A = 1 + a_1 + a_2 + a_1 a_2 - a_3 a_4 = 1 + a_1 + a_2 + o(a_i) = 1 : a_1 + a_2 = 0.$$

Man bemerke wiederum, dass in der physikalischen Literatur im Gegensatz zur mathematischen Tradition die Generatoren meist durch Multiplikation mit der imaginären Einheit $\pm i$ redefiniert werden, was in vielen Fällen auf praktische hermitesche Operatoren führt. Entsprechend führen wir an dieser Stelle die Operatoren

$$J_l = -i M_l, \quad K_l = -i N_l, \quad l = 1, 2, 3 \quad (275)$$

ein und erhalten aus den Kommutatoren (271) die renormierten Kommutatoren

$$[J_l, J_m] = i \epsilon_{lmn} J_n, \quad [K_l, K_m] = -i \epsilon_{lmn} J_n, \quad [K_l, J_m] = i \epsilon_{lmn} K_n = [J_l, K_m]. \quad (276)$$

Eine kompaktere Notation dieser Relationen ergibt sich durch die Indizierung der oben gefundenen Generatoren mit $\mu, \nu = 0, 1, 2, 3$ gemäss

$$J_{\mu\nu} = -J_{\nu\mu}, \quad \begin{cases} J_{lm} = \epsilon_{lmn} J_n & : l, m = 1, 2, 3, \\ J_{l0} = -J_{0l} = -K_l & : l = 1, 2, 3. \end{cases} \quad (277)$$

Dann lassen sich die Kommutationsrelationen der Generatoren durch die etwas umständliche, durch explizites Nachrechnen aber leicht überprüfbare Formel

$$[J_{\mu\nu}, J_{\rho\sigma}] = i(g_{\nu\rho}J_{\mu\sigma} - g_{\mu\rho}J_{\nu\sigma} + g_{\mu\sigma}J_{\nu\rho} - g_{\nu\sigma}J_{\mu\rho}) \quad (278)$$

darstellen. Beispielsweise gilt gemäss der letzten Relation in (276)

$$[J_1, K_3] = i\epsilon_{132}K_2 = -iK_2, \quad (279)$$

dies entspricht gemäss Gleichung (278) mit $\mu = 2$, $\nu = 3$, $\rho = 3$ und $\sigma = 0$

$$[J_1, K_3] = -[J_{23}, J_{30}] = -ig_{33}J_{20} = iJ_{20} = -iK_2. \quad (280)$$