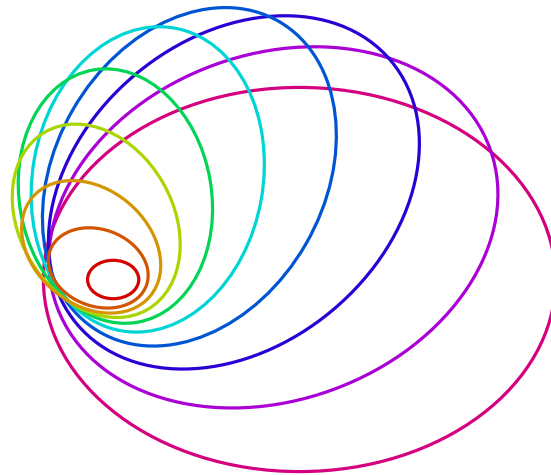


# Exkursionen in der Allgemeinen Mechanik

Aktivitäten aus der Übungsstunde zur  
Vorlesung *Allgemeine Mechanik* im  
Herbst 2021 an der ETH Zürich



SÉBASTIEN C. GARMIER

Version 1.1

21. Februar 2022

© 2021-2022 Sébastien C. Garmier.

Dieses Dokument sowie Teile davon sind urheberrechtlich geschützt und lizenziert unter der Creative-Commons-Lizenz „Namensnennung - Weitergabe unter gleichen Bedingungen 4.0 International“ (CC BY-SA 4.0).



Die Lizenz befindet sich hier:

<https://creativecommons.org/licenses/by-sa/4.0/deed.de>

Die neuste Version dieses Dokuments befindet sich hier:

<https://sebastiengarmier.ch/physics>

Bei Fragen, Rückmeldungen und Korrekturvorschlägen:

[kontakt@sebastiengarmier.ch](mailto:kontakt@sebastiengarmier.ch)

Titelbild: geometrisch ähnliche Bahnen des Kepler-Problems.

# Vorwort

Dies ist eine lose Sammlung von untereinander unabhängigen Aktivitäten zu Themen der klassischen Mechanik, speziellen Relativitätstheorie, sowie Lagrange- und Hamilton-Mechanik. Für jede Aktivität sind Lösungen, weiterführende Bemerkungen und passende Literatur aufgeführt.

Die Aktivitäten entwarf ich als Teil der wöchentlichen von mir geleiteten Übungsdoppellektion, welche die Vorlesung *Allgemeine Mechanik*<sup>1</sup> von PROF. DR. RENATO RENNER<sup>2</sup> im Herbstsemester 2021 an der ETH Zürich begleitete. Insgesamt wurden sieben Übungen bei verschiedenen Übungsassistenten angeboten und die Studierenden konnten ihre bevorzugte Übung auswählen. Die Gestaltung der Übungslektionen war den Übungsassistenten fast komplett selber überlassen, was sicherlich die Entstehung dieser Aktivitäten begünstigt hat.

Pro Woche wurde nebst anderen Inhaltspunkten in meiner Doppellektion eine Aktivität behandelt (mit Ausnahme von ein paar Wochen). Die Studierenden hatten in der Regel 15-20min Zeit, um vorzugsweise zu zweit oder in Gruppen eine *Teilaufgabe* der Aktivität zu bearbeiten. Die Teilaufgaben standen jeweils frei zur Auswahl. Anschliessend erklärte ich die Teilaufgaben jeweils kurz. Wenige Tage nach der Übungslektion veröffentlichte ich dann jeweils die Lösungen.

Ich verfolgte mit den Aktivitäten folgende Ziele:

- (i) Die Behandlung der Aktivitäten in der Übungslektion sollten zu einem vertieften Auseinandersetzen mit dem Stoff der Aktivität führen, und somit mit dem aktuellen Stoff der Vorlesung.

Die Teilaufgaben der Aktivität sind deshalb eher Verständnisfragen, Diskussionsfragen, Interpretationsfragen (im Sinn von physikalischer Interpretation) und Intuitionsfragen, statt Rechenaufgaben.

Zudem sind die Teilaufgaben bewusst nicht einfach; ich habe deshalb auch explizit nicht erwartet, dass ein Paar oder eine Gruppe alle Teilaufgaben innerhalb der gegebenen Zeit löst. Das Ziel war jeweils bereits nur die Diskussion einer Teilaufgabe.

- (ii) Zusammen mit den Lösungen sollen die Aktivitäten sinnvolle Beispiele, Anwendungen und auch Exkurse für die Themen der Vorlesung liefern. Deshalb habe ich auch die Lösung jeweils möglichst ausführlich geschrieben.
- (iii) Jede Aktivität soll ein interessantes und manchmal auch überraschendes Thema behandeln. In allen Themen gibt es sicherlich mehr zu entdecken, als man anfänglich vermutet.

Ich hoffe, dass die einte oder andere Aktivität auch ab und zu eine Studentin oder einen Studenten dazu bewegt hat, sich genauer ins Thema einzulesen, und so vielleicht weiter Konzepte zu entdecken.

Seit den Übungslektionen wurden Tippfehler in den Aktivitäten und Lösungen korrigiert, und jede Aktivität um einen Abschnitt mit weiterführenden Bemerkungen und passender Literatur ergänzt. Aktivität 9 zur *mechanischen Ähnlichkeit* wurde etwas umgeschrieben, weil die Version der Übungslektion zu verworren war. Die restlichen Aktivitäten wurden fast ohne Änderungen hier übernommen. Die Aktivitäten 7 zu *Ungleichungen für zeitartige 4er-Vektoren* und 11 zum *Noether-Theorem im Hamilton-Formalismus* wurden nicht in der Übungslektion behandelt; Aktivität 7 wurde interessierten Studierenden aber zur Verfügung gestellt, während 11 bisher unveröffentlicht war.

---

<sup>1</sup><http://www.vorlesungsverzeichnis.ethz.ch/Vorlesungsverzeichnis/lerneinheit.view?lerneinheitId=147764&semkez=2021W>

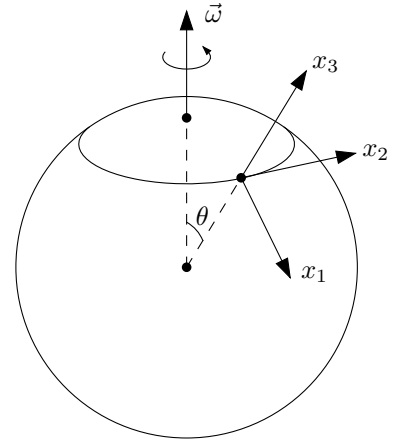
<sup>2</sup>Institut für Theoretische Physik, ETH Zürich, Schweiz.

# Inhaltsverzeichnis

Vorwort	2
1 Foucault'sches Pendel	4
2 Eigenschwingungen von gekoppelten Pendeln	7
3 Komplexer Ansatz für Oszillator mit allgemeiner Inhomogenität	11
4 Dzhani­bekov-Effekt	16
5 Der Gradient als Differentialform	20
6 Zeitreisen in die Vergangenheit?	26
7 * Ungleichungen für zeitartige 4er-Vektoren	30
8 Geodäten der Kugeloberfläche $S^2$	33
9 Mechanische Ähnlichkeit und drittes Keplersches Gesetz	37
10 Relativistischer Drehimpuls aus dem Noether-Theorem	39
11 * Noether-Theorem im Hamilton-Formalismus	43
Literaturverzeichnis	45

# 1 Foucault'sches Pendel

Wir betrachten das *Foucault'sche Pendel*: ein Pendel mit Fadenlänge  $\ell$ , das auf der Erde an einem Punkt  $P$  mit Polwinkel  $\theta$  aufgehängt ist (Breitengrad  $\pi/2 - \theta$ ). Wir betrachten das Pendel in einem Koordinatensystem, welches fest an der Erde befestigt und am Punkt  $P$  zentriert ist:  $x_1$  zeigt nach Süden,  $x_2$  nach Osten und  $x_3$  nach oben.



- (a) Vernachlässigen Sie vorerst die Erdrotation  $\vec{\omega}$ . Argumentieren Sie, dass für kleine Auslenkungen des Pendels die Bewegung der Pendelmasse approximativ auf die  $x_1$ - $x_2$ -Ebene beschränkt ist. Zeigen Sie, dass dann die Bewegungsgleichungen

$$\ddot{x}_1 + \kappa^2 x_1 = 0, \quad \ddot{x}_2 + \kappa^2 x_2 = 0$$

gelten und bestimmen Sie  $\kappa^2$ .

- (b) Jetzt nehmen wir die Erdrotation hinzu. Die Zentrifugalbeschleunigung können wir immer noch vernachlässigen, weil sie von der Ordnung  $\mathcal{O}(\omega^2)$  ist, und  $\omega T \ll 1$ , wobei  $T$  die Periode des Pendels (ohne Erdrotation) bezeichnet und  $\omega = |\vec{\omega}| \approx 1/24\text{h}$ . Die Coriolisbeschleunigung ist hingegen  $\mathcal{O}(\omega)$  und muss berücksichtigt werden:

$$\vec{a}_C = -2\vec{\omega} \times \dot{\vec{x}}.$$

Zeigen Sie, dass die Bewegungsgleichungen jetzt folgendermassen aussehen:

$$\ddot{x}_1 + \kappa^2 x_1 = 2\omega \cos \theta \dot{x}_2, \quad \ddot{x}_2 + \kappa^2 x_2 = -2\omega \cos \theta \dot{x}_1.$$

- (c) Führen Sie die komplexe Variable  $\xi := x_1 + ix_2$  ein. Zeigen Sie, dass die vorherigen Bewegungsgleichungen in der Form

$$\ddot{\xi} + 2i\omega \cos \theta \dot{\xi} + \kappa^2 \xi = 0$$

geschrieben werden können.

- (d) Lösen Sie die komplexen Bewegungsgleichungen mit einem geeigneten Ansatz für den Fall  $\omega \ll \kappa$ . Zeigen Sie, dass sich die Schwingungsrichtung des Pendels in der  $x_1$ - $x_2$ -Ebene mit der Kreisfrequenz

$$\Omega = \omega \cos \theta$$

dreht.

*Hinweis: Verwenden Sie, dass  $\sqrt{\omega^2 \cos^2 \theta + \kappa^2} = \kappa \left(1 + \frac{1}{2} \frac{\omega^2}{\kappa^2} \cos^2 \theta + \mathcal{O}\left(\frac{\omega^4}{\kappa^4}\right)\right)$ .*

## Lösung

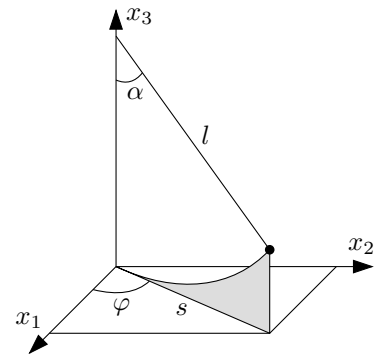
- (a) Wir können die Position des Pendels mit zwei Winkeln  $\alpha$  (Auslenkung) und  $\varphi$  (Drehung um die Vertikale) beschreiben, siehe die Abbildung rechts.

Die senkrecht nach unten gerichtete Gravitationskraft führt bloss zu einer rüctreibenden Kraft in  $\alpha$ -Richtung:

$$\ddot{\alpha} = -\frac{g}{l} \sin \alpha = -\frac{g}{l} \alpha + \mathcal{O}(\alpha^3), \quad \ddot{\varphi} = 0.$$

Es gilt  $s = l \sin \alpha = l\alpha + \mathcal{O}(\alpha^3)$  (vgl. Abbildung), also

$$\ddot{s} = l\ddot{\alpha} = -g\alpha + \mathcal{O}(\alpha^3) = -\frac{g}{l}s + \mathcal{O}(\alpha^3).$$



Wir nehmen nun an, dass  $\alpha \ll 1$  so klein ist, dass wir Terme von Ordnung  $\mathcal{O}(\alpha^2)$  und höher vernachlässigen können. Dabei finden wir auch, dass wir die  $x^3$ -Koordinate der Pendelmasse effektiv auf  $x^3 = 0$  setzen können, da

$$x^3 = l(1 - \cos \alpha) = l(1 - 1 + \mathcal{O}(\alpha^2)) = \mathcal{O}(\alpha^2).$$

Unsere Pendelmasse bewegt sich also für genügend kleine  $\alpha$  in der Ebene  $x^3 = 0$  und erfährt eine radiale Beschleunigung  $\ddot{s}$ .

Wir können diese Beschleunigung nun noch in ihre beiden Komponenten aufteilen:

$$\ddot{x}_1 = \cos \varphi \ddot{s} = -\frac{x_1}{s} \frac{g}{l} s = -\frac{g}{l} x_1$$

und analog

$$\ddot{x}_2 = -\frac{g}{l} x_2.$$

Wir finden somit, dass  $\kappa^2 = g/l$ .

- (b) Wir berechnen  $\vec{\omega} \times \dot{\vec{x}} = \dot{x}_1 \vec{\omega} \times \vec{e}_1 + \dot{x}_2 \vec{\omega} \times \vec{e}_2$ . Aus der Geometrie der Abbildung in der Aufgabenstellung folgt, dass

$$\dot{x}_1 \vec{\omega} \times \vec{e}_1 = \vec{e}_2 \dot{x}_1 \omega \sin(\pi/2 + \theta) = \vec{e}_2 \dot{x}_1 \omega \cos \theta$$

und

$$\dot{x}_2 \vec{\omega} \times \vec{e}_2 = -\vec{e}_1 \dot{x}_2 \omega \cos \theta - \vec{e}_3 \dot{x}_2 \omega \sin \theta.$$

Wenn wir nun noch das Minuszeichen in  $\vec{a}_C = -2\vec{\omega} \times \dot{\vec{x}}$  beachten, folgt die Aussage.

Alternativ können wir auch die Komponenten von  $\vec{\omega}$  im Koordinatensystem  $x^1, x^2, x^3$  aufschreiben und das Kreuzprodukt explizit ausrechnen:

$$-2\vec{\omega} \times \dot{\vec{x}} = -2\omega \begin{pmatrix} -\sin \theta \\ 0 \\ \cos \theta \end{pmatrix} \times \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ 0 \end{pmatrix} = \begin{pmatrix} 2\omega \cos \theta \cdot \dot{x}_2 \\ -2\omega \cos \theta \cdot \dot{x}_1 \\ 2\omega \sin \theta \cdot \dot{x}_2 \end{pmatrix}.$$

In beiden Methoden können wir die dritte Komponente der Coriolisbeschleunigung gegenüber der viel stärkeren Gravitationsbeschleunigung ignorieren.

- (c) Eine direkte Rechnung zeigt

$$\ddot{\xi} = \ddot{x}_1 + i\ddot{x}_2 = -\kappa^2(x_1 + ix_2) + 2\omega \cos \theta \dot{x}_2 - 2i\omega \cos \theta \dot{x}_1 = -\kappa^2 \xi - 2i\omega \cos \theta \dot{\xi}.$$

- (d) Wir versuchen den Ansatz  $\xi \propto e^{i\lambda t}$ . Einsetzen in die Differentialgleichung aus (c) und kürzen mit  $e^{i\lambda t}$  ergibt

$$-\lambda^2 - 2\omega \lambda \cos \theta + \kappa^2 = 0.$$

Die beiden Lösungen sind

$$\lambda_{\pm} = -\omega \cos \theta \pm \frac{1}{2} \sqrt{4\omega^2 \cos^2 \theta + 4\kappa^2} = -\omega \cos \theta \pm \kappa \left( 1 + \frac{1}{2} \frac{\omega^2}{\kappa^2} \cos^2 \theta + \mathcal{O}\left(\frac{\omega^4}{\kappa^4}\right) \right).$$

Wenn wir nur Terme bis und mit Ordnung  $\mathcal{O}(\omega/\kappa)$  behalten (weil  $\omega \ll \kappa$ ), dann erhalten wir die allgemeine Lösung

$$\xi = e^{-i\omega \cos \theta \cdot t} \cdot (A_1 e^{i\kappa t} + A_2 e^{-i\kappa t}),$$

wobei  $A_1, A_2 \in \mathbb{C}$ .

Der Term in Klammern beschreibt die harmonische Schwingung eines Pendels in zwei Dimensionen, ohne weitere äussere Einflüsse. Die komplexe Phase  $e^{-i\omega \cos \theta \cdot t}$  hat zur Folge, dass sich die Schwingungsebene mit der Zeit dreht, und zwar mit einer Kreisfrequenz  $\Omega = \omega \cos \theta$  um die negative  $x^3$ -Achse (mit Drehsinn gegeben durch die rechte Hand).

An den Polen ( $\theta = 0, \theta = \pi$ ) dreht sich die Schwingungsrichtung eines solchen Pendels also genau um  $2\pi$  alle 24 Stunden. Wir können z.B. am Nordpol auch sofort sehen, dass das Minuszeichen im Exponenten Sinn macht: die Erde dreht sich nach Osten, was am Nordpol Drehung um die  $x^3$ -Achse entspricht; die in  $x$ -Koordinaten wahrgenommen Drehung der Schwingungsebene ist also umgekehrt, um die negative  $x^3$ -Achse. Am Südpol ist die Ausrichtung der  $x^3$ -Achse gerade umgekehrt, aber dafür  $\cos \theta = -1$ . An den Breitengraden zwischen den Polen dauert eine Präzession des Pendels länger als 24 Stunden, am Äquator fällt sie sogar ganz aus. Ist  $\omega$  bekannt, lässt sich mit einem Foucaultschen Pendel der Breitengrad bestimmen. Die Präzession an sich zeigt auch bereits, dass sich die Erde dreht.

## Weiterführende Bemerkungen und Literatur

Die Aktivität basiert auf Aufgabe 3 in Kapitel §39 von [9].

Die Coriolisbeschleunigung (die in der Aktivität vorausgesetzt wurde) und die Zentrifugalbeschleunigung lassen sich aus der Bewegung des rotierenden Bezugssystems (hier  $x_1, x_2, x_3$ , befestigt an der Erde) relativ zu einem Inertialsystem (hier z.B. das Bezugssystem, in dem die Fixsterne ruhen <sup>1</sup>) herleiten. Siehe z.B. Kapitel 19-4 von [14] für eine eher intuitive Erklärung ausgehend von den Newtonschen Gleichungen und Kapitel §39 von [9] für eine Herleitung ausgehend von Lagrange-Mechanik.

**Vereinfachung durch komplexe Variablen** In der Aktivität haben wir ein reelles System zweier Differentialgleichungen zweiter Ordnung in eine einzige komplexe Differenzialgleichung zweiter Ordnung umgewandelt, indem wir die beiden Variablen als Real- und Imaginärteil einer komplexen Variablen auffassten.

Solch ein Ansatz kann hilfreich sein, wenn eine Rotation der Koordinaten  $x_1$  und  $x_2$  als Lösung erwartet wird, also  $x_1(t) = A \cos(\omega t)$ ,  $x_2(t) = A \sin(\omega t)$ . Denn solch eine Rotation kann komplex als  $x_1(t) + ix_2(t) = Ae^{i\omega t}$  geschrieben werden. In unserem Fall erwarten wir nicht bloss eine Rotation, sondern auch noch eine Schwingung (eine periodische Zu- und Abnahme von  $A$ ), aber der Ansatz funktioniert trotzdem.

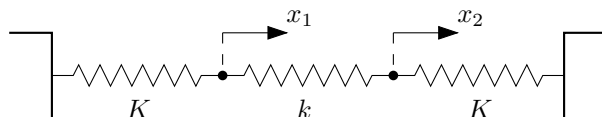
In Aktivität 3 werden wir ebenfalls zwei reelle Variablen zu einer einzigen komplexen Variable kombinieren, um das Problem zu vereinfachen. Nur werden es dort nicht zwei Koordinaten sein, sondern eine Koordinate und ihre Zeitableitung.

---

<sup>1</sup>Dies ist in sehr guter Näherung ein Inertialsystem.

## 2 Eigenschwingungen von gekoppelten Pendeln

Wir betrachten ein Federpendel mit zwei schwingenden Massen  $m$  ohne Gravitation. Die Federkonstanten sind  $k$  respektive  $K$  und es sei  $\sqrt{k} \ll \sqrt{K}$ . Die Auslenkungen aus der Ruhelage der Massen schreiben wir mit  $x_1$  und  $x_2$ .



- (a) Zeigen Sie, dass die folgende Bewegungsgleichung gilt:

$$\frac{d^2}{dt^2} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{m} \begin{pmatrix} -k - K & k \\ k & -k - K \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

- (b) Versuchen Sie, die beiden Eigenschwingungen (Lösungen, bei denen sich beide Massen mit der gleichen Frequenz bewegen) qualitativ zu finden und aufzuschreiben, ohne die Eigenwerte und Eigenvektoren der obigen Matrix zu berechnen.
- (c) Finden Sie die Frequenzen der Eigenschwingungen durch bestimmen der Eigenwerte der Matrix. Bestimmen Sie auch die Eigenvektoren und vergleichen Sie mit Ihren intuitiven Resultaten aus (b). Welche der Eigenschwingungen hat eine höhere Frequenz? Wieso macht das Sinn?
- (d) Erklären Sie, warum die allgemeine Lösung durch

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A \begin{pmatrix} 1 \\ 1 \end{pmatrix} \sin(\omega_1 \cdot t + \alpha) + B \begin{pmatrix} 1 \\ -1 \end{pmatrix} \sin(\omega_2 \cdot t + \beta), \quad A, B, \alpha, \beta \in \mathbb{R}.$$

gegeben ist, mit  $\omega_1^2 = K/m$  und  $\omega_2^2 = 2k/m + K/m$ .

- (e) Bei  $t = 0$  sollen sich beide Massen in der Ruheposition befinden ( $x_1(0) = x_2(0) = 0$ ). Masse 1 wurde kurz zuvor angestossen, sodass  $\dot{x}_1(0) = v$ , Masse 2 befindet sich bei  $t = 0$  in Ruhe. Zeigen Sie, dass Masse 1 nach endlicher Zeit beinahe stillsteht.

*Hinweis: Verwenden Sie*

$$\frac{1}{\omega_2} = \frac{1}{\omega_1 \sqrt{1 + 2k/K}} = \frac{1}{\omega_1} (1 + \mathcal{O}(k/K))$$

*und das Additionstheorem*

$$\sin \theta + \sin \phi = 2 \sin \left( \frac{\theta + \phi}{2} \right) \cos \left( \frac{\theta - \phi}{2} \right).$$



## Lösung

- (a) Auf Teilchen 1 wirken zwei Kräfte: die Federkraft  $-Kx_1$  der linken Feder und die Federkraft  $-kx_1+kx_2$  der mittleren Feder. Von den Vorzeichen überzeugen wir uns, indem wir uns vorstellen, dass Teilchen 1 und 2 der Reihe nach leicht aus der Ruhelage  $x_1 = x_2 = 0$  verschoben werden. Analog wirken auf Teilchen 2 die beiden Federkräfte  $-Kx_2$  und  $-kx_2+kx_1$  der rechten und mittleren Feder. Somit folgt

$$\frac{d^2}{dt^2} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{m} \begin{pmatrix} -k-K & k \\ k & -K-k \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} =: \frac{A}{m} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

- (b) Falls  $x_1 = x_2$ , so ist die mittlere Feder ungespannt und ihre Kraft verschwindet. Zudem üben die beiden Federn auf den Seiten dieselbe Kraft aus. Es sollte deshalb zu einer Oszillation mit  $x_1(t) = x_2(t) \propto \cos(\omega_1 \cdot t)$  kommen, mit irgendeiner Frequenz  $\omega_1$ , die durch  $K$  und  $m$  alleine bestimmt sein sollte, weil  $k$  hier keine Rolle spielt. Die beiden Massen schwingen sozusagen als eine einzige Masse, siehe (i) in der Abbildung unten.

Falls  $x_1 = -x_2$ , so ist die Situation perfekt symmetrisch. Wir vermuten deshalb, dass die beiden Massen nun entgegengesetzt schwingen:  $x_1(t) = -x_2(t) \propto \cos(\omega_2 \cdot t)$ , mit einer anderen Frequenz  $\omega_2$ , die auch von  $k$  abhängen sollte, da die mittlere Feder jetzt zur Schwingung beiträgt. Siehe (ii) in der Abbildung unten.



- (c) Das charakteristische Polynom der Matrix  $A$  aus (a) ist

$$\text{char}_A(\lambda) = (\lambda + k + K)^2 - k^2 = \lambda^2 + 2\lambda(k + K) + (k + K)^2 - k^2.$$

Die Nullstellen sind

$$\lambda_{1,2} = -(k + K) \pm \frac{1}{2} \sqrt{4(k + K)^2 - 4(k + K)^2 + 4k^2} = -(k + K) \pm k.$$

Die Eigenwerte von  $A/m$  sind also  $\lambda_{1,2}/m$ . Nun wissen wir, dass die komplexen Eigenfrequenzen die zweiten Wurzeln von  $\lambda_{1,2}$  sind, weil die Differentialgleichung aus (a) zweiter Ordnung ist. Die komplexen Eigenfrequenzen sind also

$$\pm i\omega_1, \quad \omega_1 = \sqrt{K/m}, \quad \pm i\omega_2, \quad \omega_2 = \sqrt{2k/m + K/m}.$$

Alle vier sind rein imaginär. Die vier Eigenschwingungen sind also reine Oszillationen (keine exponentielle Zu- oder Abnahme der Amplitude), mit den Frequenzen  $\omega_1$  und  $\omega_2$ .

Die Eigenschwingungen sind nun  $v_1 e^{\pm i\omega_1 t}$  und  $v_2 e^{\pm i\omega_2 t}$ , wobei  $v_1$  und  $v_2$  Eigenvektoren von  $A/m$  zu den Eigenwerten  $\lambda_1/m$  und  $\lambda_2/m$  sind. Mit (b) raten wir diese Vektoren als  $v_1 = (1, 1)^T$  und  $v_2 = (1, -1)^T$ . Tatsächlich ist dies der Fall, wie eine direkte Rechnung zeigt. Natürlich können wir  $v_1$  und  $v_2$  aber auch durch Lösen von linearen Gleichungssystemen finden.

Zum Schluss bemerken wir, dass  $\omega_1 < \omega_2$ . Dies macht Sinn, weil in der zweiten Eigenmode (Frequenz  $\omega_2$ ) alle drei Federn beteiligt sind, in der ersten (Frequenz  $\omega_1$ ) aber nur die beiden äusseren. Deshalb erwarten wir bei der zweiten Mode auch grössere wirkende Kräfte und somit eine grössere Schwingfrequenz.

- (d) Die allgemeine Lösung der Differentialgleichung ist eine Linearkombination der Eigenschwingungen aus (c):

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} (A_+ e^{i\omega_1 t} + A_- e^{-i\omega_1 t}) + \begin{pmatrix} 1 \\ -1 \end{pmatrix} (B_+ e^{i\omega_2 t} + B_- e^{-i\omega_2 t}), \quad A_{\pm}, B_{\pm} \in \mathbb{C}.$$

Wir interessieren uns aber nur für physikalische Lösungen, also solche mit reellen  $x_1$  und  $x_2$ . Damit die Imaginärteile der Ausdrücke in Klammern null sind, muss  $\bar{A}_- = A_+$  und  $\bar{B}_- = B_+$  gelten. Wir

schreiben nun  $A_+ = Ae^{i\phi_A}/2$  und  $B_+ = Be^{i\phi_B}/2$  mit  $A, B, \phi_A, \phi_B \in \mathbb{R}$  und erhalten

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cos(\omega_1 t + \phi_A) + B \begin{pmatrix} 1 \\ -1 \end{pmatrix} \cos(\omega_2 t + \phi_B) = A \begin{pmatrix} 1 \\ 1 \end{pmatrix} \sin(\omega_1 t + \alpha) + B \begin{pmatrix} 1 \\ -1 \end{pmatrix} \sin(\omega_2 t + \beta),$$

mit  $\alpha = \phi_A + \pi/2$  und  $\beta = \phi_B + \pi/2$ .

- (e)  $x_1(0) = x_2(0) = 0$  verlangt  $\alpha = \beta = 0$ .  $\dot{x}_1(0) = v$  und  $\dot{x}_2(0) = 0$  führt zu  $A\omega_1 = B\omega_2$  und  $A\omega_1 + B\omega_2 = v$ . Also,

$$A = \frac{v}{2\omega_1}, \quad B = \frac{v}{2\omega_2}.$$

Die Lösung zu den gegebenen Anfangswerten ist also

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{v}{2} \begin{pmatrix} \sin(\omega_1 t)/\omega_1 + \sin(\omega_2 t)/\omega_2 \\ \sin(\omega_1 t)/\omega_1 - \sin(\omega_2 t)/\omega_2 \end{pmatrix}.$$

Um den Ausdruck für  $x_1$  zu vereinfachen, verwenden wir den ersten Hinweis:

$$\frac{1}{\omega_2} = \frac{1}{\sqrt{2k/m + K/m}} = \frac{1}{\omega_1 \sqrt{1 + 2k/K}} = \frac{1}{\omega_1} \left( 1 - \frac{k}{K} + \mathcal{O}\left(\frac{k^2}{K^2}\right) \right) = \frac{1}{\omega_1} + \frac{1}{\omega_1} \mathcal{O}\left(\frac{k}{K}\right),$$

wobei wir Terme ab  $\mathcal{O}(k/K)$  vernachlässigen, weil  $\sqrt{k} \ll \sqrt{K}$ . Mit dieser Näherung ist

$$x_1 \approx \frac{v}{2\omega_1} (\sin(\omega_1 t) + \sin(\omega_2 t)) = \frac{v}{\omega_1} \sin\left(\frac{\omega_1 + \omega_2}{2} t\right) \cos\left(\frac{\omega_1 - \omega_2}{2} t\right),$$

wobei wir den zweiten Hinweis einsetzen. Wir bemerken, dass wir im Sinus und Kosinus  $\omega_2$  nicht mit  $\omega_1$  approximieren, da  $t$  potentiell sehr gross werden könnte.

$\dot{x}_1 \approx 0$  kommt vor, wenn der langsam variierende Kosinus-Term null ist, also zum ersten Mal bei

$$t_1 = \frac{\pi}{2} \frac{2}{|\omega_1 - \omega_2|}.$$

Nach der Zeit  $t_1$  hat die mittlere Feder also fast die ganze Energie der ersten Masse auf die zweite übertragen.

## Weiterführende Bemerkungen und Literatur

In der Lösung wurden folgende zwei Resultate aus der Theorie der Differentialgleichungen verwendet:

**Satz.** Betrachte ein lineares System von  $\ell$  Differentialgleichungen der Ordnung  $n$ , sprich

$$A_n(t) \frac{d^n}{dt^n} \vec{y}(t) + A_{n-1}(t) \frac{d^{n-1}}{dt^{n-1}} \vec{y}(t) + \dots + A_1(t) \frac{d}{dt} \vec{y}(t) + A_0(t) \vec{y}(t) = 0,$$

mit  $\vec{y}(t) \in \mathbb{C}^\ell$  und  $A_0(t), \dots, A_n(t) \in \text{Mat}_{\ell \times \ell}(\mathbb{C})$ .

Die Lösungen  $\vec{y}(t)$  dieses Gleichungssystems bilden einen Vektorraum der Dimension  $\ell \times n$ .

**Lemma.** Betrachte ein lineares System von  $\ell$  Differentialgleichungen der Ordnung  $n$  mit konstanten Koeffizienten und spezieller Form

$$\frac{d^n}{dt^n} \vec{y}(t) = A \vec{y}(t), \quad \vec{y}(t) \in \mathbb{C}^\ell, \quad A \in \text{Mat}_{\ell \times \ell}.$$

Falls  $\lambda \neq 0$  ein Eigenwert von  $A$  und  $\vec{v}_\lambda \neq 0$  der dazugehörige Eigenvektor ist, dann sind

$$\vec{y}(t) = \vec{v}_\lambda e^{\sqrt[n]{\lambda} t}, \quad \sqrt[n]{\lambda} \in \{n\text{-te Wurzeln von } \lambda\}. \quad (2.1)$$

$n$  linear unabhängige Lösungen des Differentialgleichungssystems.

Falls  $0$  ein Eigenvektor von  $A$  und  $\vec{v}_0 \neq 0$  der dazugehörige Eigenvektor ist, dann sind

$$\vec{y}(t) = \vec{v}_0 t^k, \quad k = 0, 1, \dots, n-1 \quad (2.2)$$

$n$  linear unabhängige Lösungen des Differentialgleichungssystems.

Zusammenfassend liefert jeder Eigenraum von  $A$  mit Dimension  $k$  so  $k \times n$  linear unabhängige Lösungen. Ist  $A$  diagonalisierbar, wird der ganze Lösungsraum des Differentialgleichungssystems durch solche Lösungen aufgespannt.

Lösungen der Form (2.1) heissen *Eigenschwingungen*;  $\sqrt[n]{\lambda}$  nennen wir die dazugehörige *komplexe Eigenfrequenz*. Ihr Imaginärteil bestimmt, wie schnell die Lösung oszilliert, ihr Realteil, ob und wie stark die Lösung im Betrag exponentiell zu- oder abnimmt. Manchmal werden (2.2) auch als Eigenschwingungen gezählt, mit Eigenfrequenz null.

In der Aktivität haben wir gerade ein Differentialgleichungssystem wie im Lemma, mit  $\ell = 2$  und  $n = 2$ . Weiter sind alle Eigenwerte ungleich null, also kommen nur Lösungen der Form (2.1) in Frage. Da wir zwei verschiedene Eigenwerte haben, können wir mit dem Lemma eine Basis des Lösungsraums finden.

**Beweis des Lemmas.** Angenommen,  $\lambda \neq 0$  ist ein Eigenwert von  $A$  mit Eigenvektor  $\vec{v}_\lambda \neq 0$ . Dann ist

$$\frac{d^n}{dt^n} (\vec{v}_\lambda e^{\sqrt[n]{\lambda}t}) = \vec{v}_\lambda \left( \sqrt[n]{\lambda} \right)^n e^{\sqrt[n]{\lambda}t} = \vec{v}_\lambda \lambda e^{\sqrt[n]{\lambda}t} = A \vec{v}_\lambda e^{\sqrt[n]{\lambda}t}.$$

Angenommen,  $0$  ist ein Eigenvektor von  $A$  mit Eigenwert  $\vec{v}_0 \neq 0$ . Dann ist

$$\frac{d^n}{dt^n} (\vec{v}_0 t^k) = 0 = A \vec{v}_0 = A \vec{v}_0 t^k,$$

wenn  $k < n$ . In beiden Fällen sind die  $n$  Funktionen linear unabhängig; im ersten Fall, weil Exponentialfunktionen mit unterschiedlichen Exponenten linear unabhängig sind, im zweiten Fall, weil Monome  $t^k$  mit verschiedenen  $k$  linear unabhängig sind.

Ein Eigenraum der Dimension  $k$  besitzt per Definition  $k$  linear unabhängige Eigenvektoren ungleich null, also folgt auch die zusammenfassende Aussage.  $A$  besitzt eine Basis aus Eigenvektoren, genau dann, wenn  $A$  diagonalisierbar ist.  $\square$

**Beweisskizze des Satzes.** Wir benötigen für den Beweis den *Satz von Piccard-Lindelöf*. Er besagt grob gesagt, dass ein System von  $\ell$  Differentialgleichungen erster Ordnung eine eindeutige Lösung besitzen, wenn  $\ell$  Anfangswerte für die Lösungsfunktionen angegeben werden und das Differentialgleichungssystem die sogenannte *Lipschitz-Bedingung* erfüllt. Nun können wir ein System von  $\ell$  Differentialgleichungen  $n$ -ter Ordnung stets in ein System von  $n \times \ell$  Differentialgleichungen erster Ordnung umschreiben, wobei die  $k$ -ten Ableitungen der Lösungsfunktionen für  $k = 1, \dots, n-1$  als neue Lösungsfunktionen aufgefasst werden. Die Anfangsbedingungen sind also die Anfangswerte der ursprünglichen  $\ell$  Lösungsfunktionen, zusammen mit den Anfangswerten ihrer  $k$ -ten Ableitungen ( $k = 1, \dots, n-1$ ).

Die möglichen Anfangswerte unseres Differentialgleichungssystems bilden somit einen Vektorraum der Dimension  $\ell \times n$ . Aber nun ist das System linear: das heisst, Linearkombinationen von Lösungen sind wieder Lösungen. Die Menge aller Lösungen ist somit ein Vektorraum. Da jede Lösung natürlicherweise eindeutige Anfangsbedingungen besitzt und wegen dem Satz von Piccard-Lindelöf auch umgekehrt jede Anfangsbedingung eine eindeutige Lösung ergibt (vorausgesetzt, die Lipschitz-Bedingung gilt), ist der Lösungsraum isomorph zum Vektorraum der Anfangsbedingungen. Der Lösungsraum hat also auch Dimension  $\ell \times n$ .

Dies funktioniert natürlich nur, wenn die Lipschitz-Bedingung erfüllt ist. Es kann gezeigt werden, dass diese für Differentialgleichungssysteme der betrachteten Form immer erfüllt ist.  $\square$

Zum Schluss bemerken wir, dass der Beweis einiges einfacher wird, wenn wir nur Matrizen  $A_0, \dots, A_n$  betrachten, die nicht von  $t$  abhängen. Dann kann die Lösung des Anfangswertproblems nämlich durch das Matrixexponential bestimmt werden. Die Dimension des Lösungsraums folgt dann aus den Eigenschaften des Matrixexponentials.

Siehe z.B. [12] für eine ausführliche Behandlung gewöhnlicher Differentialgleichungen, inklusive mathematischen Details, die wir hier ausgelassen haben.

### 3 Komplexer Ansatz für Oszillator mit allgemeiner Inhomogenität

Betrachten Sie die Bewegungsgleichung eines erzwungenen, eindimensionalen harmonischen Oszillators:

$$\ddot{x} + \omega^2 x = \frac{f(t)}{m}, \quad (3.1)$$

mit der Oszillatormasse  $m$  und einer allgemeinen, zeitabhängigen Kraft  $f(t)$ .

- (a) Zeigen Sie, dass (3.1) folgendermassen in eine neue ODE mit Inhomogenität umgeschrieben werden kann:

$$\dot{\xi} - i\omega\xi = \frac{f(t)}{m}, \quad \xi := \dot{x} + i\omega x. \quad (3.2)$$

Wieso lohnt sich diese Umformung?

*Hinweis: Auch für diese DG können wir Propagatoren einsetzen; ist der Propagator für  $\xi$  einfacher als für  $(x, \dot{x})$ ?*

- (b) Berechnen Sie den Propagator  $P(t, s)$  des homogenen Problems. Sie sollten  $P(t, s)\xi_s = e^{i\omega(t-s)}\xi_s$  erhalten. Setzen Sie dies in die Duhamel-Formel

$$\xi(t) = P(t, 0)\xi_0 + \int_0^t dt' P(t, t') \frac{f(t')}{m}$$

ein, um die Lösung  $\xi(t)$  aus den Anfangsbedingungen  $\xi_0$  zu erhalten.

*Hinweis: Um den Propagator zu finden, muss einfach die homogene Version von (3.2) gelöst werden.*

- (c) Sei nun  $\xi_0 = 0$ . Zeigen Sie, dass die Energie des Oszillators  $E(t) = \frac{m}{2}(\dot{x}^2(t) + \omega^2 x^2(t)) = \frac{m}{2}|\xi(t)|^2$  ist.  
 (d) Angenommen,  $f(t) = 0$ , falls  $t < 0$  oder  $t > T$ , mit einem  $T > 0$ . Zeigen Sie, dass bei einer Zeit  $t > T$  die Energie durch

$$E(t) = \frac{1}{2m} \left| \int_{-\infty}^{\infty} f(t') e^{-i\omega t'} dt' \right|^2 \quad (3.3)$$

gegeben ist.

Was ist die Bedeutung des Ausdrucks in den Betragsstrichen auf der rechten Seite? Macht das intuitiv Sinn?

*Hinweis: Solche Ausdrücke sind aus den Methoden der mathematischen Physik bekannt.*

- (e) Zusätzlich zu den Annahmen in (d) sei noch  $f(t) = f_0 = \text{konst.}$  und  $0 < T \ll 1/\omega$ . Zeigen Sie, dass

$$E = \frac{(f_0 T)^2}{2m} + \mathcal{O}(T^2 \omega^2).$$

Interpretieren Sie das Resultat.

## Lösung

- (a) Wir setzen  $\xi(t) := \dot{x} + i\omega x$ . (3.1) ist dann

$$\dot{\xi} = \ddot{x} + i\omega\dot{x} = -\omega^2 x + \frac{f(t)}{m} + i\omega\dot{x} = i\omega\xi + \frac{f(t)}{m},$$

was zu zeigen war.

Die neue Differentialgleichung ist wieder linear. Die homogene Version davon ist sogar autonom und der Propagator wird somit zu einem einfachen Exponential. Da die Gleichung für  $\xi$  aber nur erster Ordnung ist, ist der Propagator für  $\xi$  besonders einfach: er wird sich als Exponential einer  $1 \times 1$ -Matrix herausstellen, also einfach einer Zahl. Im Gegensatz dazu ist der Propagator für  $(x, \dot{x})$  das Exponential einer  $2 \times 2$ -Matrix und somit komplizierter.

- (b) Der Propagator kann an der Lösung der homogenen Gleichung abgelesen werden. Die homogene Gleichung ist  $\dot{\xi} = i\omega\xi$ , also ist  $\xi(t) = \xi_0 e^{i\omega t}$ . Wenn statt bei  $t = 0$  die Anfangsbedingungen zur Zeit  $s$  gegeben sind, dann finden wir  $\xi(t) = \xi_s e^{i\omega(t-s)}$ . Die Wirkung des Propagators ist also

$$P(t, s)\xi_s = e^{i\omega(t-s)}\xi_s.$$

Eingesetzt in die Duhamel-Formel ergibt dies

$$\xi(t) = e^{i\omega t}\xi_0 + \int_0^t dt' e^{i\omega(t-t')} \frac{f(t')}{m}.$$

- (c) Die Energie des Oszillators ist

$$E(t) = \frac{m}{2}\dot{x}^2 + \frac{m}{2}\omega^2 x^2.$$

Der erste Term ist die kinetische Energie, der zweite die potentielle. Dass der zweite Term die potentielle Energie ist, sehen wir, wenn wir die homogene Version von (3.1) etwas umschreiben:

$$m\ddot{x} = -m\omega^2 x = -\frac{d}{dx} \frac{m}{2}\omega^2 x^2 =: -\frac{d}{dx} U(x).$$

Also können wir  $U$  als potentielle Energie sehen. Zum Schluss bemerken wir, dass  $|\xi|^2 = \dot{x}^2 + \omega^2 x^2$ .

*Bemerkung:* Wir haben hier kein spezifisches physikalisches System angenommen. Die so aufgeschriebene Energie muss also nicht unbedingt die Energie sein, die dem System auf Basis der Newtonschen Mechanik zugeschrieben würde. Unsere „Energie“ ist also eine allgemeinere Grösse, die auch für Systeme definiert ist, die einfach nur Gleichung (3.1) gehorchen, aber sonst nichts mit Mechanik zu tun haben.

- (d) Wir haben unter diesen Annahmen

$$\xi(t) = \int_0^t dt' e^{i\omega t} e^{-i\omega t'} \frac{f(t')}{m},$$

also

$$E(t) = \frac{m}{2} \left| \int_0^t dt' e^{-i\omega t'} \frac{f(t')}{m} \right|^2.$$

Nun können wir die Integralgrenzen aber beliebig erweitern, da dort  $f$  ohnehin null ist:

$$E(t) = \frac{1}{2m} \left| \int_{-\infty}^{\infty} dt' e^{-i\omega t'} f(t') \right|^2$$

Das Integral ist gerade die *Fourier-Transformation*  $\hat{f}$  von  $f$ , ausgewertet an der Eigenfrequenz  $\omega$  des ungestörten Oszillators:

$$E = \frac{|\hat{f}(\omega)|^2}{2m}.$$

Somit trägt nur die Fourier-Komponente von  $f$ , die der Eigenfrequenz  $\omega$  des ungestörten Oszillators entspricht, zur Energie des Oszillators bei.

Dies macht intuitiv Sinn: Wir wissen nämlich, dass eine harmonische Kraft ( $f(t) \propto e^{i\Omega t}$ ) besonders effektiv Energie in den Oszillator übertragen kann, wenn  $\Omega \approx \omega$  (Resonanz). Dies erklärt, warum  $\hat{f}(\omega)$ , also der Anteil der Schwingung  $e^{i\omega t}$  in der Kraft  $f$ , einen wichtigen Beitrag zu  $E$  leistet. Dass es der einzige Anteil ist, ist mit dieser Intuition aber nicht sofort sichtbar.

- (e) Wir haben wieder  $t > T$ , also können wir das Integral auf den Bereich  $[0, T]$  einschränken. Also ist  $0 \leq t' \leq T \ll 1/\omega$ , das heisst,  $\omega t' \ll 1$ . Wir können dann die Exponentialfunktion im Integral in einer Taylorreihe schreiben und nur die ersten paar Terme betrachten:

$$e^{-i\omega t'} = 1 - i\omega t' + \mathcal{O}(\omega^2 t'^2).$$

Es folgt dann, dass

$$E = \frac{1}{2m} \left| \int_0^T dt' (1 - i\omega t' + \mathcal{O}(\omega^2 t'^2)) f_0 \right|^2 = \frac{f_0^2}{2m} \left| T - \frac{i\omega}{2} T^2 + \mathcal{O}(\omega^2 T^3) \right|^2 = \frac{(f_0 T)^2}{2m} (1 + \mathcal{O}(\omega^2 T^2)).$$

Wir behalten nur den ersten Term.  $\Delta p = f_0 T$  ist der gesamte Impuls, der durch die Kraft an den Oszillator in der kurzen Zeit  $T$  übertragen wurde. Aber dann ist  $E = \frac{\Delta p^2}{2m} = \Delta E_{\text{kin}}$  die kinetische Energie, die der Oszillator durch die kurze Einwirkung der Kraft gewonnen hat. Dies ist genau, was wir auch bei einem unendlich kurzen Kraftstoss erwarten würden. Das macht Sinn, weil wir ja angenommen haben, dass  $T \ll 1$ ; im Limes  $T\omega \rightarrow 0$  fallen alle höheren Terme in  $\omega T$  weg, und wir erhalten genau einen unendlich kurzen Kraftstoss.

*Bemerkung:* Weil die Kraft nur kurz wirkte, hat sich die Pendelmasse noch nicht sehr weit bewegt und die potentielle Energie ist immer noch näherungsweise null.

## Weiterführende Bemerkungen und Literatur

Die Aktivität basiert auf Kapitel §22 von [9].

**Exkurs: Das freie elektromagnetische Feld als Reihe von Oszillatoren** Der Ansatz  $\dot{\xi} = \dot{x} + i\omega x$  mag auf den ersten Blick willkürlich erscheinen. Eine etwas bessere Motivation dafür finden wir in der Elektrodynamik, wo das freie elektromagnetische Feld (keine Ladungen und Ströme) als Sammlung von unendlich vielen harmonischen Oszillatoren gesehen werden kann.

Im Vakuum sind das elektrische Feld  $\vec{E}$  und das magnetische Feld  $\vec{B}$  folgendermassen bestimmt:

$$\vec{E} = -\frac{1}{c} \frac{\partial \vec{A}}{\partial t}, \quad \vec{B} = \vec{\nabla} \times \vec{A}, \quad \square \vec{A} = 0, \quad \vec{\nabla} \cdot \vec{A} = 0.$$

$\vec{A}$  ist das *Vektorpotential* und  $\square = \frac{\partial^2}{c^2 \partial t^2} - \Delta$  ist der *d'Alambert-Operator*. Die allgemeinste Lösung der Gleichung  $\square \vec{A} = 0$  mit Nebenbedingung  $\vec{\nabla} \cdot \vec{A} = 0$  ist

$$\begin{aligned} \vec{A}(t, \vec{x}) &= \int \frac{d^3 k}{(2\pi)^{3/2}} \sum_{\lambda=1}^2 \left( a_{\lambda, \vec{k}} \vec{\varepsilon}_{\lambda, \vec{k}} \cdot e^{i(|\vec{k}|ct - \vec{k} \cdot \vec{x})} + \overline{a_{\lambda, \vec{k}} \vec{\varepsilon}_{\lambda, \vec{k}}} \cdot e^{-i(|\vec{k}|ct - \vec{k} \cdot \vec{x})} \right) \\ &=: \int \frac{d^3 k}{(2\pi)^{3/2}} \sum_{\lambda=1}^2 \vec{A}_{\lambda, \vec{k}}(t, \vec{x}), \end{aligned}$$

wobei  $a_{\lambda, \vec{k}} \in \mathbb{C}$  ist und  $\vec{\varepsilon}_{1, \vec{k}}, \vec{\varepsilon}_{2, \vec{k}}$  zwei orthogonale Einheitsvektoren in  $\mathbb{C}^3$  sind (sogenannte *Polarisationsvektoren*), die den Raum senkrecht zu  $\vec{k}$  aufspannen. Für eine genauere Herleitung siehe zum Beispiel die Kapitel §18 und §46 von [10] oder die Kapitel 18 und 20 von [15]<sup>1</sup>. Das freie Feld kann also als Superposition von ebenen Wellen  $\vec{A}_{\lambda, \vec{k}}$  mit verschiedenen Wellenvektoren  $\vec{k}$  und Polarisierungen  $\lambda$  gesehen werden. Die  $\vec{A}_{\lambda, \vec{k}}(t, \vec{x})$  werden auch *Eigenmoden* genannt; sie sind an sich bereits Lösungen für das freie elektromagnetische Feld.

Wir bemerken, dass jede Mode  $\vec{A}_{\lambda, \vec{k}}(t, \vec{x})$  die Differentialgleichung

$$\frac{\partial^2}{\partial t^2} \vec{A}_{\lambda, \vec{k}}(t, \vec{x}) = -|\vec{k}|^2 c^2 \vec{A}_{\lambda, \vec{k}}(t, \vec{x})$$

eines harmonischen Oszillators erfüllt. Wir können uns deshalb fragen, ob das elektromagnetische Feld als unendliche Ansammlung von ungekoppelten, harmonischen Oszillatoren, einer pro Mode, beschrieben werden kann. Solch eine Beschreibung wäre nicht nur elegant, sondern auch im Rahmen der *Quantenmechanik*

<sup>1</sup>Sowohl [10] wie auch [15] behandeln Elektrodynamik, aber [15] ist deutlich einfacher und intuitiver, dafür ist [10] viel ausführlicher.

interessant: denn wir wissen sehr gut, wie ein harmonischer Oszillator *quantisiert* wird; somit könnten wir gleich auch das elektromagnetische Feld quantisieren.

Um zu zeigen, dass dies der Fall ist, sind weitere Schritte notwendig:

- (i) Wir müssen zeigen, dass jede Mode  $\vec{A}_{\lambda,\vec{k}}$  tatsächlich ein harmonischer Oszillator ist. Dazu müssen wir pro Mode eine Auslenkungsvariable  $q_{\lambda,\vec{k}}(t)$ , eine Impulsvariable  $p_{\lambda,\vec{k}}(t)$ , sowie eine Hamiltonfunktion  $H_{\lambda,\vec{k}}(q_{\lambda,\vec{k}}, p_{\lambda,\vec{k}})$  finden, sodass die Hamiltonschen Bewegungsgleichungen dieses Systems genau diejenigen eines harmonischen Oszillators sind, d.h.:

$$H_{\lambda,\vec{k}} = \frac{1}{2} \left( p_{\lambda,\vec{k}}^2 + \omega_{\lambda,\vec{k}}^2 q_{\lambda,\vec{k}}^2 \right),$$

wobei  $\omega_{\lambda,\vec{k}}$  die Frequenz des jeweiligen Oszillators ist.

- (ii) Sinnvollerweise sollte sich die Gesamtenergie  $H$  des elektromagnetischen Felds ausschliesslich aus den Energien  $H_{\lambda,\vec{k}}$  der einzelnen Oszillatoren zusammensetzen und zudem in den  $H_{\lambda,\vec{k}}$  linear sein, weil die Oszillatoren ungekoppelt sind.

Mit den Moden  $\vec{A}_{\lambda,\vec{k}}$  haben wir bereits Kandidaten für die Auslenkungsvariablen. Zum Beispiel erfüllt aber auch  $a_{\lambda,\vec{k}} \exp(i|\vec{k}|ct) + \bar{a}_{\lambda,\vec{k}} \exp(-i|\vec{k}|ct)$  die Oszillatorgleichung und würde in Frage kommen.

Wir beginnen mit dem Punkt (ii). Aus der Elektrodynamik wissen wir, dass die Energie des elektromagnetischen Felds in Gausschen Einheiten durch

$$H(t) = \frac{1}{8\pi} \int d^3x \left( \vec{E}(t, \vec{x})^2 + \vec{B}(t, \vec{x})^2 \right) = \int d^3k \sum_{\lambda} \frac{|\vec{k}|^2}{2\pi} |a_{\lambda,\vec{k}}|^2$$

gegeben ist. Die rechte Seite folgt nach längerer Rechnung. Die Gesamtenergie des Felds zerfällt also in Beiträge aus den einzelnen Moden. Dies ist längst nicht trivial, da wir wegen den Quadraten  $\vec{E}^2$  und  $\vec{B}^2$  Mischterme zwischen den Moden erwarten würden; sie streichen sich aber alle weg. Punkt (ii) könnte also erfüllt werden.

Wir möchten nun für Punkt (i)  $q_{\lambda,\vec{k}}$  und  $p_{\lambda,\vec{k}}$  so bestimmen, dass  $H_{\lambda,\vec{k}} = |\vec{k}|^2 |a_{\lambda,\vec{k}}|^2 / 2\pi$  (wir wählen die Koeffizienten in der Linearkombination so, dass  $H_{\lambda,\vec{k}}$  als Energie pro Wellenvektor und Polarisierung aufgefasst wird). Wir benötigen

$$\frac{1}{2} \left( p_{\lambda,\vec{k}}^2 + \omega_{\lambda,\vec{k}}^2 q_{\lambda,\vec{k}}^2 \right) = \frac{|\vec{k}|^2}{2\pi} |a_{\lambda,\vec{k}}|^2.$$

An dieser Stelle wird ein Ansatz wie in der Aktivität nützlich:

$$B(t) \frac{|\vec{k}|}{\sqrt{2\pi}} a_{\lambda,\vec{k}} = \frac{1}{\sqrt{2}} p_{\lambda,\vec{k}} + i \frac{1}{\sqrt{2}} \omega_{\lambda,\vec{k}} q_{\lambda,\vec{k}}$$

mit einer komplexen Funktion  $B(t)$  von Betrag 1 (diese Funktion ist notwendig, weil die rechte Seite von  $t$  abhängt). Dieser Ansatz ist motiviert dadurch, dass dann

$$\frac{|\vec{k}|^2}{2\pi} |a_{\lambda,\vec{k}}|^2 \underbrace{|B(t)|^2}_{=1} = \frac{1}{2} \left( p_{\lambda,\vec{k}}^2 + \omega_{\lambda,\vec{k}}^2 q_{\lambda,\vec{k}}^2 \right),$$

wie verlangt.

Zum Vergleich: in der Aktivität hatten wir  $H = \frac{m}{2} (\dot{x}^2 + \omega^2 x^2)$ , wobei  $\sqrt{m}x$  die Rolle der Auslenkungsordinate und  $\sqrt{m}\dot{x}$  die Rolle der Impulsordinate übernahm. Bis auf andere Faktoren entsprach der Ansatz  $\xi(t) = \dot{x} + i\omega x$  dort dem Ansatz  $B(t)a_{\lambda,\vec{k}} \propto p_{\lambda,\vec{k}} + i\omega_{\lambda,\vec{k}} q_{\lambda,\vec{k}}$  hier.

Nun müssen gemäss (i) für unsere Oszillatoren auch Oszillatorgleichungen gelten:

$$\frac{dq_{\lambda,\vec{k}}}{dt} = \frac{\partial H_{\lambda,\vec{k}}}{\partial p_{\lambda,\vec{k}}} = p_{\lambda,\vec{k}}, \quad \frac{d^2 q_{\lambda,\vec{k}}}{dt^2} = \frac{dp_{\lambda,\vec{k}}}{dt} = -\frac{\partial H_{\lambda,\vec{k}}}{\partial q_{\lambda,\vec{k}}} = -\omega_{\lambda,\vec{k}}^2 q_{\lambda,\vec{k}}.$$

Somit wird klar, dass  $B(t) = \exp(i\omega_{\lambda,\vec{k}}t + \delta_{\lambda,\vec{k}})$ <sup>2</sup>. Nehmen wir nun von unserem Ansatz den Imaginärteil, so folgt

$$\frac{|\vec{k}|}{\sqrt{\pi}} \frac{1}{2i} \left( a_{\lambda,\vec{k}} e^{i\omega_{\lambda,\vec{k}}t + i\delta_{\lambda,\vec{k}}} - \bar{a}_{\lambda,\vec{k}} e^{-i\omega_{\lambda,\vec{k}}t - i\delta_{\lambda,\vec{k}}} \right) = \omega_{\lambda,\vec{k}} q_{\lambda,\vec{k}}.$$

<sup>2</sup> $q_{\lambda,\vec{k}}(t) = \sin(\omega_{\lambda,\vec{k}}t + \delta_{\lambda,\vec{k}})$  und  $p_{\lambda,\vec{k}}(t) = w_{\lambda,\vec{k}} \cos(\omega_{\lambda,\vec{k}}t + \delta_{\lambda,\vec{k}})$ , verwende dann die Eulersche Formel. OBdA. haben wir die Auslenkung gleich 1 gesetzt; die Wahl der Auslenkung entspricht nämlich nur einer Umbenennung der Variablen. Genauso enthält  $\delta_{\lambda,\vec{k}}$  keine physikalische Information und könnte auf null gesetzt werden; wir behalten uns diese Freiheit aber noch vor.

Es wird klar, dass  $|\vec{k}| \propto \omega_{\lambda, \vec{k}}$ , wobei aber die Proportionalitätskonstante frei wählbar ist, solange sie die richtigen Einheiten besitzt. Somit haben wir (i) und (ii) erfüllt. Wir können nun noch unsere verbleibenden Freiheiten aufbrauchen, um die Auslenkungs- und Impulsvariablen so sinnvoll wie möglich zu wählen.

Es ist sinnvoll, dass der harmonische Oszillator einer Mode mit derselben Frequenz wie die Mode schwingt; deshalb wählen wir  $\omega_{\lambda, \vec{k}} := |\vec{k}|c$ . Mit der Wahl  $\delta_{\lambda, \vec{k}} = \pi/2$  erhalten wir die einfache Definition

$$q_{\lambda, \vec{k}}(t) := \frac{1}{\sqrt{4\pi c^2}} \left( a_{\lambda, \vec{k}} \cdot e^{i|\vec{k}|ct} + \overline{a_{\lambda, \vec{k}}} \cdot e^{-i|\vec{k}|ct} \right).$$

Und somit

$$p_{\lambda, \vec{k}}(t) = -\frac{i}{\sqrt{4\pi c^2}} |\vec{k}|c \left( a_{\lambda, \vec{k}} \cdot e^{i|\vec{k}|ct} - \overline{a_{\lambda, \vec{k}}} \cdot e^{-i|\vec{k}|ct} \right),$$

was auch am Realteil unseres Ansatzes abgelesen werden kann. Zusammen mit

$$H = \int d^3k \sum_{\lambda} H_{\lambda, \vec{k}}, \quad H_{\lambda, \vec{k}} = \frac{1}{2} \left( p_{\lambda, \vec{k}}^2 + |\vec{k}|^2 c^2 q_{\lambda, \vec{k}}^2 \right)$$

fassen die letzten beiden Gleichungen die Tatsache zusammen, dass das elektromagnetische Feld als Sammlung von unendlich vielen, ungekoppelten, harmonischen Oszillatoren aufgefasst werden kann, mit einem Oszillator pro Mode.

An dieser Stelle könnten wir mit der Quantisierung des elektromagnetischen Felds beginnen; es sei aber dazu auf weiterführende Literatur verwiesen: z.B. die ersten paar Kapitel von [11]. Viele fortgeschrittene Bücher über *Quantenfeldtheorie* behandeln meistens auch die Quantisierung des freien elektromagnetischen Felds, so z.B. [4].

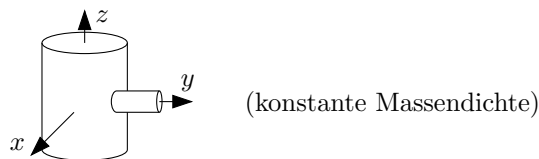


## 4 Dzhanibekov-Effekt

Wir betrachten einen asymmetrischen, starren Körper, der in Schwerelosigkeit (kein externes Drehmoment) um die mittlere Hauptträgheitsachse dreht. Diese Drehung ist instabil und schon bei winzigen Störungen kommt es zum sogenannten *Dzhanibekov-Effekt*<sup>1</sup>. Die Bewegung beginnt stabil, wird aber schnell instabil.

Wir möchten hier verstehen, auf welcher Zeitskala die Bewegung instabil wird.

- (a) Betrachten Sie den starren Körper in der Abbildung. Argumentieren Sie, dass seine Hauptträgheitsachsen etwa die Koordinatenachsen sind und  $I_x > I_y > I_z$  gilt.



- (b) Wir betrachten die vereinfachten Anfangsbedingungen  $\omega_y(0) = \omega_0$ ,  $\omega_x(0) = \varepsilon_x$ ,  $\omega_z(0) = \varepsilon_z$ , mit  $\omega_0 \gg |\varepsilon_x|, |\varepsilon_z|$ . Inwiefern beschreibt dies die Anfangsbedingungen im Video?
- (c) Die Euler-Gleichungen sind

$$I_x \dot{\omega}_x = (I_y - I_z) \omega_y \omega_z, \quad I_y \dot{\omega}_y = (I_z - I_x) \omega_z \omega_x, \quad I_z \dot{\omega}_z = (I_x - I_y) \omega_x \omega_y.$$

Wir interessieren uns nur für den Anfang der Bewegung, wo  $\omega_y(t) \sim \omega_0$  und  $\omega_x(t), \omega_z(t) \sim |\varepsilon_x|, |\varepsilon_z|$ . („ $\sim$ “ bedeutet hier „in der Grössenordnung von“)

Zeigen Sie, dass mit diesen Annahmen und den Anfangsbedingungen aus (b) bis und mit Ordnung  $\varepsilon_{x,z}/\omega_0$  folgt, dass

$$\omega_y(t) = \omega_0 = \text{konst.}, \quad \ddot{\omega}_{x,z} = \frac{(I_y - I_z)(I_x - I_y)}{I_x I_z} \omega_0^2 \cdot \omega_{x,z} =: J^2 \cdot \omega_{x,z}, \quad J^2 > 0.$$

- (d) Schliessen Sie daraus, dass zumindest am Anfang der Bewegung

$$\vec{\omega}(t) = \left( \varepsilon_x \cosh(Jt) + A \sinh(Jt), \omega_0, \varepsilon_z \cosh(Jt) + B \sinh(Jt) \right)^T$$

gilt. Die Konstanten  $A$  und  $B$  müssen nicht genau bestimmt werden. Was bedeutet das für die Stabilität der Rotation um die  $y$ -Achse?

- (e) Welches ist die erste Ordnung von  $t$ , in der  $\omega_x$  oder  $\omega_z$  von ihrem Anfangswert abweichen? Was bestimmt, wie lange die Bewegung zu Beginn stabil aussieht?

*Hinweis: Betrachten Sie den Graphen von  $\cosh$  und  $\sinh$  oder genauer die Taylorreihen. Berechnen Sie explizit die Konstanten  $AJ$  und  $BJ$ .*

<sup>1</sup>Siehe [https://www.youtube.com/watch?v=L2o9eB1\\_Gzw](https://www.youtube.com/watch?v=L2o9eB1_Gzw)

## Lösung

- (a) Wir betrachten Hauptachsen durch den Schwerpunkt, da die Drehung eines kräftefreien Körpers am natürlichsten um seinen Schwerpunkt beschrieben wird.

Die  $y$ -Achse ist eine Hauptachse. Denn für jeden Massenpunkt im Körper gibt es auch einen, der durch die  $y$ -Achse gespiegelt auf der anderen Seite des Körpers liegt; Somit folgt, dass der gesamte Drehimpuls des Körpers bei einer Rotation um die  $y$ -Achse ebenfalls entlang der  $y$ -Achse zeigt; aber wenn  $\vec{L} \parallel \vec{\omega}$ , haben wir es gerade mit einer Hauptachse zu tun.

Die Hauptachse mit dem grössten Hauptträgheitsmoment maximiert das Integral

$$\int d^2y \rho(\vec{y}^2) \vec{y}_\perp^2,$$

wobei  $\vec{y}_\perp^2$  das Quadrat des vertikalen Abstands vom Punkt  $\vec{y}$  zur betrachteten Achse ist. Dies ist intuitiv der Fall, wenn der Körper um die  $x$ -Achse dreht, weil dann die beiden Zylinder des Körpers senkrecht von der Achse in drei Richtungen weg zeigen und somit möglichst viel Masse möglichst weit Weg von der Drehachse ist. Es gilt also  $I_x > I_y$ .

Schliesslich müssen alle Hauptträgheitsachsen senkrecht aufeinander stehen, weshalb die  $z$ -Achse die dritte Hauptachse sein muss. Sie ist zugleich auch die mit dem kleinsten Hauptträgheitsmoment, weil die Massenverteilung dann am nächsten an der Drehachse ist, was das obige Integral minimiert. Wir haben also auch noch  $I_y > I_z$ .

- (b) Im Video dreht der Körper zu Beginn fast genau um die  $y$ -Achse; leichte Störungen haben zur Folge, dass die Drehachse nicht genau die  $y$ -Achse ist. Somit konzentriert sich die anfängliche Winkelgeschwindigkeit fast gänzlich auf die  $y$ -Achse, mit vergleichsweise kleinen Beiträgen entlang den beiden anderen Achsen. Dies entspricht der gegebenen mathematischen Form der Anfangsbedingungen.
- (c) Die Euler-Gleichung für  $\omega_y$  wird durch Multiplikation mit  $\omega_0^{-2}$  zu

$$\omega_0^{-2} I_y \dot{\omega}_y = (I_z - I_x) \frac{\omega_x \omega_z}{\omega_0 \omega_0}.$$

Auf der rechten Seite steht ein Term der Ordnung  $\varepsilon^2/\omega_0^2$ , den wir vernachlässigen. Somit folgt  $\dot{\omega}_y = 0$ , also  $\omega_y(t) = \omega_y(0) = \omega_0$ .

Die restlichen beiden Euler-Gleichung werden nun zu

$$\dot{\omega}_x = \frac{I_y - I_z}{I_x} \omega_0 \omega_z, \quad \dot{\omega}_z = \frac{I_x - I_y}{I_z} \omega_0 \omega_x.$$

Wir leiten beide nach der Zeit ab und setzen die eine in die andere ein:

$$\ddot{\omega}_x = \frac{(I_y - I_z)(I_x - I_y)}{I_x I_z} \omega_0^2 \omega_x.$$

Dieselbe Gleichung erhalten wir auch für  $\omega_z$  statt  $\omega_x$ . Die Ausdrücke in den Klammern sind alle positiv, genauso wie die Hauptträgheitsmomente selber. Also ist

$$\ddot{\omega}_{x,z} = J^2 \cdot \omega_{x,z}, \quad J^2 := \omega_0^2 \frac{(I_y - I_z)(I_x - I_y)}{I_x I_z} > 0.$$

- (d) Die Differentialgleichungen für  $\omega_x$  und  $\omega_z$  können mittels sinh und cosh gelöst werden (vgl. harmonischer Oszillator, wo die Lösungen aber aus sin und cos bestehen):

$$\omega_x(t) = C \cosh(Jt) + A \sinh(Jt), \quad \omega_z(t) = D \cosh(Jt) + B \sinh(Jt),$$

mit Konstanten  $A, B, C, D$ . Da  $\sinh(0) = 0$  und  $\cosh(0) = 1$ , muss  $C = \varepsilon_x$  und  $D = \varepsilon_z$ , damit die Anfangsbedingungen eingehalten werden. Wir erhalten so die gesuchte Form:

$$\vec{\omega}(t) = \left( \varepsilon_x \cosh(Jt) + A \sinh(Jt), \omega_0, \varepsilon_z \cosh(Jt) + B \sinh(Jt) \right)^T,$$

mit Konstanten  $A$  und  $B$ , die später noch festzulegen sind. Diese Lösung gilt natürlich gemäss unseren Annahmen nur für kleine  $t$ , solange  $|\omega_x(t)| \sim |\varepsilon_x|$  und  $|\omega_z(t)| \sim |\varepsilon_z|$ .

Die Lösung ist stabil, genau dann, wenn  $\omega_x$  und  $\omega_z$  beschränkt bleiben. Aber dies ist nur genau dann der Fall, wenn  $\varepsilon_x = -A$  und  $\varepsilon_z = -B$ ; denn

$$a \cosh(x) + b \sinh(x) = a \frac{e^x + e^{-x}}{2} + b \frac{e^x - e^{-x}}{2} = \frac{1}{2} \left[ (a+b)e^x + (a-b)e^{-x} \right].$$

Wenn wir die Euler-Gleichungen betrachten, können wir schnell qualitativ sehen, dass  $A$  durch  $\varepsilon_z$  bestimmt ist und  $B$  durch  $\varepsilon_x$ . Für Stabilität müssen  $\varepsilon_x$  und  $\varepsilon_z$  also genau aufeinander abgestimmt sein, was eher unwahrscheinlich ist. Die Bewegung ist also fast garantiert instabil.

(e) Wir haben

$$\cosh(Jt) = 1 + \mathcal{O}(J^2 t^2), \quad \sinh(Jt) = Jt + \mathcal{O}(J^3 t^3).$$

Wir erhalten so

$$\omega_x(t) = \varepsilon_x + AJt + \mathcal{O}(J^2 t^2), \quad \omega_z(t) = \varepsilon_z + BJt + \mathcal{O}(J^2 t^2).$$

Die Abweichung vom Anfangszustand geschieht also in erster Ordnung von  $t$ . Die Rate der Abweichung ist durch  $AJ$  beziehungsweise  $BJ$  gegeben.

Als wir die erste und dritte Euler-Gleichung (beides Gleichungen erster Ordnung) abgeleitet haben (und sie somit zu Gleichungen zweiter Ordnung wurden), entstanden je ein „künstlicher“ Freiheitsgrad. Diese werden durch die Konstanten  $A$  und  $B$  widerspiegelt. Natürlich müssen aber auch die ursprünglichen Euler-Gleichungen gelten. Indem wir unsere gefundenen Lösungen dort einsetzen, können wir  $AJ$  und  $BJ$  ablesen.

Einsetzen der Lösung in die erste Euler-Gleichung ergibt (mit  $\cosh' = \sinh$ ,  $\sinh' = \cosh$ )

$$\varepsilon_x J \sinh(Jt) + AJ \cosh(Jt) = \frac{I_y - I_z}{I_x} \omega_0 (\varepsilon_z \cosh(Jt) + B \sinh(Jt)).$$

Da  $\sinh$  und  $\cosh$  linear unabhängig sind, müssen die entsprechenden Terme einzeln gleich sein. So können wir  $AJ$  und  $BJ$  bestimmen:

$$AJ = \varepsilon_z \frac{I_y - I_z}{I_x} \omega_0,$$

$$BJ = \varepsilon_x J^2 \frac{I_x}{I_y - I_z} \frac{1}{\omega_0} = \varepsilon_x \frac{I_x - I_y}{I_z} \omega_0.$$

Wir bekommen sinnvollerweise dasselbe, wenn wir statt in die erste in die dritte Euler-Gleichung einsetzen.

Wegen der Ungleichung  $I_i + I_j \geq I_k$  ( $ijk$  irgendwelche Permutationen von  $x, y, z$ ) können die drei Hauptträgheitsmomente nicht allzu weit auseinander liegen: die Differenz zweier Momente ist stets kleiner als das dritte. Aus diesem Grund sind die Brüche  $(I_y - I_z)/I_x$  und  $(I_x - I_y)/I_z$  in den Ausdrücken für  $AJ$  und  $BJ$  etwa von der Grössenordnung 1.

In einer Zeit  $\Delta t$  (immer noch kurz) verändert sich  $\omega_x$  also um  $\Delta \omega_x \approx AJ \Delta t \sim \varepsilon_z \omega_0 \Delta t$ . Mit anderen Worten:  $\omega_x$  hat sich um den Bruchteil  $\varepsilon_z \Delta t$  von  $\omega_0$  verändert. Analoges gilt für  $\omega_z$  mit  $\varepsilon_x$  statt  $\varepsilon_z$ .

$\varepsilon_x^{-1}$  und  $\varepsilon_z^{-1}$  sind also die Zeitskalen, auf denen die Rotation um die  $y$ -Achse instabil wird.

## Weiterführende Bemerkungen und Literatur

**Numerische Simulation** Wir konnten analytisch die Entwicklung für kleine Zeiten bestimmen und sahen, dass die Instabilität der Rotationsachse vorerst nur sehr langsam wächst, wenn die anfänglichen Störungen  $\varepsilon_x$  und  $\varepsilon_z$  klein sind. Mit numerische Simulationen können wir die Drehung auch über längere Zeiten bestimmen: Abbildung 4.1 zeigt das Resultat einer numerischen Simulation mit `python`, mit  $I_x = 2$ ,  $I_y = 1.5$ ,  $I_z = 1$ ,  $\vec{\omega}(0) = (0.001, 10, 0)$ . Durch numerische Integration von

$$\dot{R}(t) = R(t)\Omega(t), \quad \Omega(t) = \begin{pmatrix} 0 & -\omega_z(t) & \omega_y(t) \\ \omega_z(t) & 0 & -\omega_x(t) \\ -\omega_y(t) & \omega_x(t) & 0 \end{pmatrix}$$

können wir die Rotationsmatrix  $R(t)$  des Körpers berechnen. In der Simulation wird deutlich, wie die Rotation um die mittlere Achse zwischen schlagartigen Richtungswechseln für vergleichsweise lange Zeit stabil bleibt.

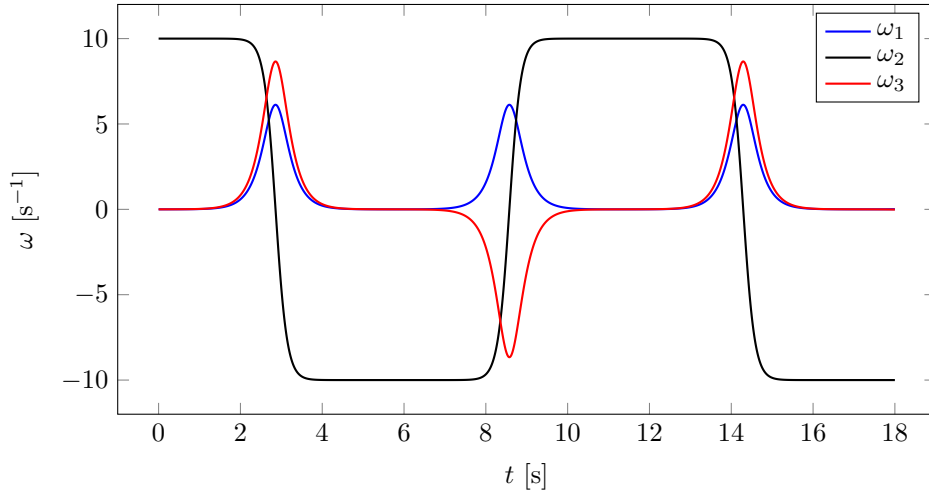


Abbildung 4.1: Numerische Simulation des Dzhanibekov-Effekts.

**Analytische Lösung** Es ist selbst für lange Zeiten eine analytische Lösung möglich, mittels sogenannten *elliptischen Funktionen*. Wir möchten nur erwähnen, dass so insbesondere die Zeit zwischen zwei Richtungswechsel (Periode der Bewegung) gefunden werden kann. Sie beträgt

$$T = 4K \sqrt{\frac{I_1 I_2 I_3}{(I_1 - I_2)(L^2 - 2EI_3)}}, \quad K = \int_0^{\pi/2} \frac{du}{\sqrt{1 - k^2 \sin^2 u}}, \quad k^2 = \frac{(I_2 - I_3)(2EI_1 - L^2)}{(I_1 - I_2)(L^2 - 2EI_3)},$$

wobei  $E = \frac{1}{2}(I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2)$  die Gesamtenergie und  $L^2 = I_1^2\omega_1^2 + I_2^2\omega_2^2 + I_3^2\omega_3^2$  das Quadrat des Gesamtdrehimpulses sind; weiter gilt die Formel nur, wenn  $I_1 > I_2 > I_3$  (wie bei uns der Fall) und wenn  $L^2 > 2EI_2$  (ebenfalls der Fall); ansonsten muss die Formel durch Ersetzen von gewissen Indizes angepasst werden. Wir erhalten so  $T \approx 11.4$  s, was auch dem numerischen Resultat entspricht (die Zeit zwischen den beiden Maxima von  $\omega_3$  können wir in 4.1 als ungefähr 11.5 s ablesen).

Siehe z.B. Kapitel §37 in [9] für eine detaillierte Erklärung der analytischen Lösung.

# 5 Der Gradient als Differentialform

Wir betrachten ein *Skalarfeld*  $\psi : M \rightarrow \mathbb{R}$  auf der Raumzeit  $M$  (der Menge aller Ereignisse). In einem Minkowski-Koordinatensystem wird dieses Feld zu einer Funktion  $\psi : \mathbb{R}^4 \rightarrow \mathbb{R}$  der vier Koordinaten:  $x^\mu = (x^0, x^1, x^2, x^3) \mapsto \psi(x^\mu)$ .

Der *Gradient*  $d\psi(E)$  von  $\psi$  am Ereignis  $E$  mit Koordinaten  $x^\mu$  hat im gewählten Koordinatensystem die Komponenten  $d\psi_\nu(x^\mu) := (\partial_0\psi(x^\mu), \dots, \partial_3\psi(x^\mu))$ , wobei  $\partial_\nu$  die Ableitung nach der  $\nu$ -ten Koordinate ist.

Sei  $x'^\nu(x^\mu) = (\Lambda^{-1})^\nu_\mu x^\mu$  eine Lorentz-Transformation mit Transformationsmatrix  $\Lambda^{-1}$ . Im neuen Koordinatensystem ist das Skalarfeld gegeben durch die neue Funktion  $\psi'(x'^\nu) = \psi(x^\mu(x'^\nu))$ .

## Analytischer Teil

(a) Zeigen Sie, dass

$$\frac{\partial x'^\nu}{\partial x^\mu} = (\Lambda^{-1})^\nu_\mu, \quad \frac{\partial x^\mu}{\partial x'^\nu} = \Lambda^\mu_\nu.$$

(b) Zeigen Sie mit Hilfe der mehrdimensionalen Kettenregel, dass der Gradient  $d\psi$  im neuen Koordinatensystem die folgenden Komponenten hat:

$$d\psi'_\nu = \Lambda^\mu_\nu d\psi_\mu.$$

Vergleichen Sie dieses Transformationsgesetz mit dem eines 4er-Vektors. Ist der Gradient  $d\psi$  ein 4er-Vektor?

*Hinweis: Koordinatendifferenzen  $\Delta x^\mu$  sind z.B. die Komponenten eines 4er-Vektors (zumindest wenn wir nur lineare  $x'^\nu(x^\mu)$  betrachten).*

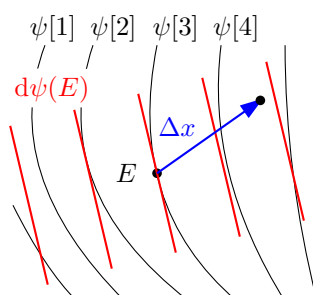
## Geometrischer Teil

(c) Erklären Sie, dass sich  $\psi : M \rightarrow \mathbb{R}$  als die Untermengen  $\psi[r] := \{E \in M \mid \psi(E) = r\} \subset M$  für  $r \in \mathbb{R}$  beschreiben lässt. Solch eine Beschreibung heißt „geometrisch“, weil sie keine Koordinaten voraussetzt. Erklären Sie, dass in einer 2D-Raumzeit diese Mengen zu den „Höhenlinien“ des Skalarfelds werden.

(d) Wir können  $\psi(x^\mu)$  um die Koordinaten  $x_0^\mu$  entwickeln:  $\psi(x_0^\mu + \Delta x^\mu) - \psi(x_0^\mu) = d\psi_\nu \Delta x^\nu + \mathcal{O}((\Delta x^\mu)^2)$ . Dies funktioniert in jedem wohldefinierten Koordinatensystem.

Argumentieren Sie ohne Verwendung von (a) und (b), dass die Zahl  $\psi(x_0^\mu + \Delta x^\mu) - \psi(x_0^\mu) \approx d\psi_\nu \Delta x^\nu$  unabhängig vom Koordinatensystem ist, in dem  $\psi$ ,  $\Delta x$  und  $d\psi$  ausgedrückt werden. Was bedeutet das für die Transformation der Komponenten von  $d\psi$  im Vergleich zu den Komponenten von  $\Delta x$ ?

(e) Erklären Sie, dass  $d\psi(E)$  geometrisch als „Stapel von Hyperflächen“ parallel zu  $\psi[\psi(E)]$  gesehen werden kann, wobei ein geometrischer Abstandsvektor  $\Delta x$  bei  $E$  genau  $d\psi(E)_\nu \Delta x^\nu$  Hyperflächen des Stapels durchstößt; siehe Abbildung unten. Wie können wir den Fall  $d\psi_\nu \Delta x^\nu < 0$  ebenfalls abdecken?



hier:  $\psi(E) = 3$  und  
 $d\psi_\nu \Delta x^\nu = 1.75$

# Lösung

## Analytischer Teil

- (a) Ableiten von  $x'^\nu = (\Lambda^{-1})^\nu_\mu x^\mu$  nach  $x^\rho$  (der  $\rho$ -ten Koordinate) ergibt

$$\frac{\partial x'^\nu}{\partial x^\rho} = \frac{\partial}{\partial x^\rho} ((\Lambda^{-1})^\nu_\mu x^\mu) = (\Lambda^{-1})^\nu_\mu \frac{\partial x^\mu}{\partial x^\rho} = (\Lambda^{-1})^\nu_\mu \delta^\mu_\rho = (\Lambda^{-1})^\nu_\rho.$$

Zudem ist

$$\Lambda^\rho_\nu x'^\nu = \Lambda^\rho_\nu (\Lambda^{-1})^\nu_\mu x^\mu = \delta^\rho_\mu x^\mu = x^\rho,$$

wobei die zweitletzte Gleichheit folgt, weil  $\Lambda \Lambda^{-1} = \text{id}$ , ausgedrückt als Matrixmultiplikation ( $\Lambda^\mu_\nu$  ist der Eintrag in der  $\mu$ -ten Zeile und  $\nu$ -ten Spalte der Matrix  $\Lambda$ ). Wenn wir dies jetzt nach  $x'^\sigma$  ableiten, erhalten wir analog zur obigen Ableitung

$$\frac{\partial x^\rho}{\partial x'^\sigma} = \Lambda^\rho_\sigma.$$

Nun können wir noch die Indizes umbenennen, um die Ausdrücke in der Aufgabenstellung zu erhalten.

- (b) Wir haben

$$d\psi'_\mu = \frac{\partial \psi}{\partial x'^\mu}.$$

Jetzt können wir aber  $\psi$  auch als Funktion der ungestrichenen Koordinaten sehen:  $\psi(x^\nu)$ , die über  $x^\nu(x'^\nu)$  wiederum Funktionen der gestrichenen Koordinaten sind. Anwendung der mehrdimensionalen Kettenregel ergibt somit (das Matrixprodukt in der Kettenregel wird durch die Einsteinsche Summenkonvention kompakt beschrieben):

$$d\psi'_\mu = \frac{\partial \psi}{\partial x'^\mu} = \frac{\partial \psi}{\partial x^\nu} \frac{\partial x^\nu}{\partial x'^\mu} = d\psi_\nu \Lambda^\nu_\mu = \Lambda^\nu_\mu d\psi_\nu.$$

Weil Ausdrücke in Indexnotation ja Summen über Produkte einzelner Komponenten (Zahlen) sind, können wir die Reihenfolge von Faktoren wie hier im letzten Schritt vertauschen.

Für die Komponenten eines 4er-Vektors  $V$  gilt ein anderes Transformationsgesetz:

$$V'^\nu = (\Lambda^{-1})^\nu_\mu V^\mu.$$

Man kann sich das folgendermassen merken: 4er-Vektorkomponenten  $V^\mu$  transformieren unter Lorentz-Transformationen (und allgemeiner unter linearen Koordinatentransformationen) wie Koordinatendifferenzen  $\Delta x^\mu$ . Der Vergleich zeigt, dass  $d\psi$  kein 4er-Vektor ist.

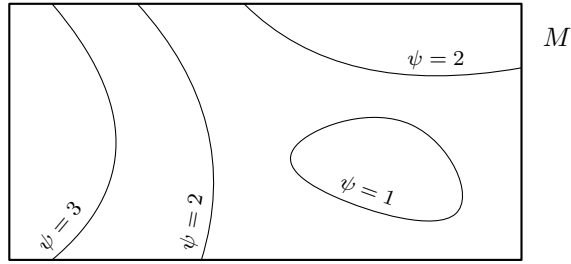
*Bemerkung: Objekte, deren Komponenten wie die Komponenten eines Gradienten transformieren, sind sogenannte Differentialformen. Wir schreiben die Komponenten einer Differentialform im Gegensatz zu 4er-Vektorkomponenten mit einem tiefgestellten Index; so wird die Einsteinsche Summenkonvention natürlicher.*

*Differentialformen können auch als Dualvektoren zu 4er-Vektoren verstanden werden: eine Differentialform  $\omega$  mit Komponenten  $\omega_\mu$  wirkt nämlich linear auf einen 4er-Vektor  $V$  via  $\omega[V] := \omega_\mu V^\mu$ .*

## Geometrischer Teil

- (c) Die gesamte Information über die Funktion  $\psi : M \rightarrow \mathbb{R}$  ist bekanntlich durch alle Ereignisse zusammen mit ihren Funktionswerten gegeben. Wir können diese Information etwas kompakter schreiben, indem wir alle Ereignisse  $E \in M$  mit einem bestimmten Funktionswert  $\psi(E) = r$  zusammen in einer Menge festhalten; so kommen wir auf die Mengen  $\psi[r] := \{E \in M \mid \psi(E) = r\}$ . Diese beschreiben  $\psi$  vollständig.

Wenn  $M$  zweidimensional ist, können wir uns vorstellen, dass  $\psi(P)$  die Höhe der Landschaft auf einer Landkarte  $M$  angibt als Funktion des Punkts  $P \in M$ . Dann wären  $\psi[r]$  die „Höhenkurven“ der Landkarte (siehe Abbildung unten).

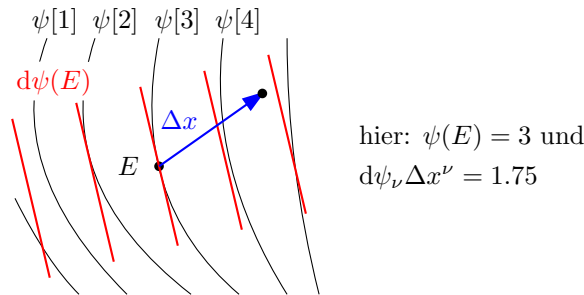


- (d)  $\psi(x_0^\mu)$  und  $\psi(x_0^\mu + \Delta x^\mu)$  sind beides Zahlen (Skalare), also unabhängig vom Koordinatensystem. Deshalb ist auch  $\psi(x_0^\mu + \Delta x^\mu) - \psi(x_0^\mu)$  koordinatenunabhängig, und im Rahmen der Approximation also auch  $d\psi_\nu \Delta x^\nu$  (die sogenannte *Kontraktion* von  $d\psi$  und  $\Delta x$ ).

Nun sind die Komponenten  $\Delta x^\nu$  aber koordinatenabhängig. Damit  $d\psi_\nu \Delta x^\nu$  koordinatenunabhängig ist, müssen die Komponenten  $d\psi_\nu$  also auch koordinatenabhängig sein, mit einem Transformationsgesetz, das genau der Transformation von  $\Delta x^\nu$  entgegenwirkt.

*Bemerkung:* Das folgt auch direkt aus (a) und (b) auf analytischem Weg.

- (e) Wir schreiben der Einfachheit halber  $\Delta\psi := \psi(x_0^\mu + \Delta x^\mu) - \psi(x_0^\mu)$ . Wir können  $\Delta\psi$  geometrisch folgendermassen ablesen: wir beginnen beim Ereignis  $E$ , welches zu den Koordinaten  $x_0^\mu$  gehört, und bewegen uns entlang des Vektors  $\Delta x$ , der zu den Komponenten  $\Delta x^\mu$  gehört, bis wir an der Spitze angekommen sind; dabei zählen wir, wie viele „Höhenkurven“ wir überschreiten, um  $\Delta\psi$  zu erhalten (für ein genaues Resultat müssen wir nicht nur die „ganzen“ Höhenkurven zählen, die  $\pm 1$  auseinanderliegen, sondern auch die Bruchteile von Höhenkurven dazwischen). Siehe blauer Pfeil und schwarze Höhenkurven in der Abbildung unten. Da  $M$  vierdimensional ist, sind  $\psi[r]$  natürlich keine Höhenkurven, sondern dreidimensionale Hyperflächen, aber die wären etwas schwierig zum zeichnen.



Jetzt möchten wir aber verstehen, wie  $d\psi$  geometrisch zu interpretieren ist. Wir beginnen dazu mit der Tatsache, dass  $d\psi$  über  $d\psi_\nu \Delta x^\nu$  eine Approximation an  $\Delta\psi$  liefert. Weiter ist diese Approximation linear in  $\Delta x$ .

Dementsprechend können wir uns vorstellen, dass  $d\psi_\nu \Delta x^\nu$ , die lineare Approximation von  $\Delta\psi$ , geometrisch auf analoge Weise entsteht: indem wir zählen, wie viele Höhenkurven von  $\Delta x$  durchstoßen werden; nur dürfen wir nicht die Höhenkurven von  $\psi$  dafür verwenden, sondern die lineare Approximation davon. Linearität verlangt, dass diese approximativen Höhenkurven Geraden sind, mit jeweils gleichem Abstand zueinander. Damit sie  $\Delta\psi$  approximieren, müssen sie zudem in  $E$  parallel zu den echten Höhenkurven von  $\psi$  sein, und im richtigen Abstand voneinander liegen. Diese linear approximierten Höhenkurven sind rot in der Abbildung eingezeichnet. Sie liefern die geometrische Bedeutung von  $d\psi$ .

Genauso wie wir echte Höhenkurven in eine Richtung positiv und in die andere negativ zählen, können wir dies auch mit den approximierten Höhenkurven tun. So können wir den Fall  $d\psi_\nu \Delta x^\nu < 0$  behandeln. Ausserdem werden die parallelen Höhenkurven von  $d\psi$  in vier Dimensionen zu einem Stapel von parallelen, dreidimensionalen Hyperebenen.

*Bemerkung:* Allgemeine Differentialformen, die nicht unbedingt Gradienten von Funktionen sind, können auch als Stapel von Hyperebenen gesehen werden. Man kann eine geometrische Addition und Skalarmultiplikation definieren, wie dies auch für Vektoren möglich ist; darauf soll hier aber nicht genauer eingegangen werden.

## Weiterführende Bemerkungen und Literatur

Die geometrische Interpretation von Differentialformen als Stapel von Hyperflächen stammt aus [1].

**Exkurs: Einführung in die Differentialgeometrie am Beispiel der Raumzeit** Wir haben gesehen, wie wir uns Differentialformen als Stapel von Hyperflächen vorstellen können; analog zu wie wir Vektoren als Pfeile sehen können. Diese geometrischen Interpretationen sind elegant, weil sie koordinatenunabhängig sind. Die geometrischen Betrachtungen in der Aktivität waren aber nicht ganz rigoros. Wir möchten hier zeigen, wie Vektoren und Differentialformen möglichst koordinatenunabhängig, also möglichst geometrisch, definiert werden können. Da diese Konzepte zur Differentialgeometrie gehören, nutzen wir die Gelegenheit und liefern eine kurze Einführung in die Differentialgeometrie anhand der Raumzeit.

Wir beginnen dazu mit der Raumzeit  $M$  als Menge von Ereignissen und nehmen bloss an, dass  $M$  eine *glatte Mannigfaltigkeit* ist. Grob gesagt besteht eine glatte Mannigfaltigkeit aus einer Menge  $M$  mit folgenden Zusatzstrukturen:

- (i) *Topologie*. Diese gibt uns einen Begriff von „Nachbarschaft“ zwischen Ereignissen in  $M$ . Konkret definiert die Topologie, welche Teilmengen von  $M$  als *offen* gelten. Sie ermöglicht uns auch, *stetige Abbildungen* von und nach  $M$  zu definieren.
- (ii) *Dimension*.  $M$  soll „in der Nähe“ von jedem  $E \in M$  wie ein Ausschnitt des vierdimensionalen euklidischen Raums  $\mathbb{R}^4$  aussehen. Grob gesagt verlangen wir deshalb, dass wir um jedes  $E \in M$  zumindest lokal ein stetiges Koordinatensystem mit vier Koordinaten finden können; aber in der Regel sind sehr viele Koordinatensysteme möglich. Mit einer weiteren eher technischen Eigenschaft wird die Raumzeit so zu einer *vierdimensionalen topologischen Mannigfaltigkeit*.
- (iii) *Glatte Struktur*. Wir möchten, dass die Koordinatensysteme nicht nur stetig, sondern beliebig oft differenzierbar (*glatt*) sind. Deshalb behalten wir grob gesagt nur Koordinatensysteme, sodass die *Koordinatentransformationsfunktionen*  $x^\mu(x^\nu)$  glatt sind. Indem wir Abbildungen  $M \rightarrow M$ ,  $M \rightarrow \mathbb{R}^n$ ,  $\mathbb{R}^n \rightarrow M$  mit Koordinaten (oder deren Inverse) kombinieren, erhalten wir stets Abbildungen zwischen euklidischen Räumen, die *Koordinatendarstellungen* der ursprünglichen Abbildungen. Wir nennen die ursprünglichen Abbildungen *differenzierbar*, *glatt*, usw., wenn es ihre Koordinatendarstellungen sind. Die glatte Struktur liefert uns also auch ein Begriff von Differenzierbarkeit für Abbildungen von und nach  $M$ .

Mit zwei weiteren technischen Eigenschaften wird die Raumzeit so zu einer *glatten Mannigfaltigkeit*.

In der Physik verlangen wir meistens noch eine *Metrik*; wir werden sie hier aber nicht brauchen. Auch haben wir in dieser kurzen Übersicht die meisten Details ausgelassen und verweisen auf die Literatur. Für eine anschauliche Einführung in die Topologie, siehe z.B. [8]. Für eine einfache Einführung in glatte Mannigfaltigkeiten, siehe z.B. [5]. Für eine ausführlichere Behandlung der Differentialgeometrie, siehe z.B. [7].

Wir können jetzt unter anderem glatte Abbildung von und nach  $M$  betrachten. Die beiden einfachsten Arten solcher Abbildungen sind *Kurven*  $\gamma : \mathbb{R} \rightarrow M$  und *Skalarfelder*  $f : M \rightarrow \mathbb{R}$ . Wir haben in der Aktivität gesehen, wie wir eine Differentialform an einem Ereignis  $E$  intuitiv als *Gradienten*  $df(E)$  eines Skalarfelds  $f$  am Ereignis  $E$  sehen können. Analog können wir einen Vektor am Ereignis  $E$  intuitiv als *Geschwindigkeitsvektor*  $v_{\gamma,E}$  bei  $E$  einer Kurve  $\gamma$  durch  $E$  sehen. Wir werden diese Ideen nun präzisieren, beginnend mit Vektoren.

Im euklidischen Raum können wir den Geschwindigkeitsvektor einer Kurve einfach durch Ableiten nach dem Kurvenparameter finden. Dies funktioniert, weil der euklidische Raum *affin* ist, und wir Punkte subtrahieren können. In der Raumzeit können wir Ereignisse ohne weiteres nicht subtrahieren. Als nächstbestes können wir aber statt Ereignisse vielleicht Objekte, die an Ereignissen definiert sind, subtrahieren; dafür bieten sich die oben erwähnten Skalarfelder an, denn die haben ja reellwertige Funktionswerte an jedem Ereignis. Wir kommen so zum Konzept der *Richtungsableitung* eines Skalarfelds  $f$  entlang der Kurve  $\gamma$  am Ereignis  $E = \gamma(\lambda_0)$ :

$$\frac{d}{d\lambda}(f \circ \gamma)(\lambda) \Big|_{\lambda=\lambda_0} =: v_{\gamma,\gamma(\lambda_0)}(f).$$

Natürlich ist an  $f$  nichts speziell; der Wert der Richtungsableitung auf der linken Seite eignet sich also nicht für die Definition des Geschwindigkeitsvektors  $v_{\gamma,E}$ . Wir können  $v_{\gamma,E}$  aber als die *Richtungsableitung an sich*



sehen; sie ist definiert durch die Wirkung, die sie auf *alle möglichen*  $f$  hat. Dies haben wir auf der rechten Seite bereits geschrieben. Der Vektor  $v_{\gamma, E}$  ist also ein Ableitungsoperator für Skalarfelder bei  $E$ , eine lineare Abbildung vom Raum aller Skalarfelder (wobei zwei Skalarfelder mit gleichen Richtungsableitungen bei  $E$  als identisch gesehen werden können) in die reellen Zahlen. Es bleibt zu zeigen, dass diese Richtungsableitungen auch einen Vektorraum bilden:

- (i) Wir erhalten die um  $s$  skalierte Richtungsableitung, wenn wir die Parametrisierung der entsprechenden Kurve um  $s$  strecken.
- (ii) Für die Summe zweier Richtungsableitungen können wir aus den beiden Kurven eine dritte Kurve konstruieren, die lokal um  $E$  betrachtet Ereignis  $E$  die „Summe“ der ersten beiden Kurven ist. Dies geht, wenn wir ein lokales Koordinatensystem um  $E$  wählen, mit Ursprung bei  $E$ , und in diesen Koordinaten die Koordinatendarstellungen der Kurven addieren. Es lässt sich zeigen, dass die resultierende Richtungsableitung nicht von der Wahl des Koordinatensystems abhängt. Wir finden so auch, dass der Vektorraum vierdimensional ist.

Der vierdimensionale Vektorraum der Geschwindigkeitsvektoren (und somit überhaupt der Vektoren) am Ereignis  $E$  wird *Tangentialraum von  $M$  bei  $E$*  genannt.

Wir kommen nun noch zu Differentialformen, beziehungsweise Gradienten von Skalarfeldern. Im euklidischen Raum ist der Gradient die Ableitung des Skalarfelds in alle möglichen Richtungen des Raums. Dies ist wohldefiniert und einfach zu berechnen, weil wir Distanzen zwischen Punkten im Raum bestimmen können; in unserer glatten Mannigfaltigkeit geht das jedoch nicht. Statt das Skalarfeld am Ereignis  $E$  entlang Richtungen des Raums abzuleiten, können wir es entlang von Kurven durch  $E$  ableiten, und den Gradienten  $df(E)$  so definieren:

$$\left. \frac{d}{d\lambda}(f \circ \gamma)(\lambda) \right|_{\lambda=\lambda_0} =: (df(E))(\gamma),$$

wobei  $E = \gamma(\lambda_0)$ . Der Gradient von  $f$  bei  $E$  ist also ein Differentialoperator, der Kurven durch  $E$  auf reelle Zahlen abbildet (wobei Kurven, die zur selben Richtungsableitung von  $f$  führen, als identisch gesehen werden können). Auch hier müssen wir noch zeigen, dass die so definierten Gradienten (und somit Differentialformen) einen Vektorraum bilden. Dies ist erstaunlicherweise einfacher als für Vektoren: die lineare Struktur folgt von der Tatsache, dass wir Skalarfelder natürlicherweise skalieren und addieren können. Wenn wir aber die Dimension dieses Vektorraums bestimmen wollen, müssen wir indirekt auf Koordinaten zugreifen: entweder erkennen wir, dass der Raum der Differentialformen isomorph zum *Dualraum* der Vektoren ist, oder wir versuchen explizit eine Basis für Differentialformen zu finden. Zweiteres erfordert ein Koordinatensystem, um Skalarfelder zu finden, die lokal um  $E$  linear unabhängige Gradienten haben. In beiden Fällen finden wir, dass der Vektorraum auch vierdimensional ist. Der vierdimensionale Vektorraum der Differentialformen am Ereignis  $E$  wird *Kotangentialraum von  $M$  bei  $E$*  genannt.

Zusammenfassend können wir Vektoren und Differentialformen also (fast) geometrisch mittels Kurven und Skalarfeldern auf  $M$  definieren. Auch sind die beiden Konzepte in vielen Hinsichten dual. Wir kommen dabei aber nicht ganz um Koordinaten herum: Koordinaten treten grob gesagt auf, sobald wir Dimensionen bestimmen müssen. Auch bei der Definition von  $M$  werden Koordinaten hauptsächlich für die Definition der Dimension von  $M$  benötigt.

Von hier aus können wir viele weitere Strukturen definieren:

- (i) Häufig werden nicht Vektoren *und* Differentialformen so konstruiert, sondern nur ein Vektorraum von beiden. Der jeweils andere Vektorraum wird dann als *Dualraum* des ersten eingeführt.
- (ii) Mittels *Tensorprodukten*, *symmetrischen Produkten* und *antisymmetrischen Produkten* können wir aus dem Tangential- und Kotangentialraum weitere Vektorräume an Ereignissen der Raumzeit definieren.
- (iii) Jedes Ereignis hat eigene Tangential- und Kotangentialräume. Wir können jedoch *Vektorfelder* und *Differentialformfelder* definieren, die jedem Ereignis ein Vektor aus dem dortigen Tangentialraum bzw. eine Differentialform aus dem dortigen Kotangentialraum zuweist. Dasselbe gilt für allgemeinere *Tensorfelder*. Mittels *Faserbündel* können wir Konzepte wie Stetigkeit oder Glätte auf diese allgemeinen Felder übertragen.
- (iv) Eine *Metrik* liefert unter anderem einen Isomorphismus zwischen dem Tangential- und dem Kotangentialraum eines Ereignisses. Eine glatte Mannigfaltigkeit mit einer Metrik heißt *Riemannsche Mannigfaltigkeit*.

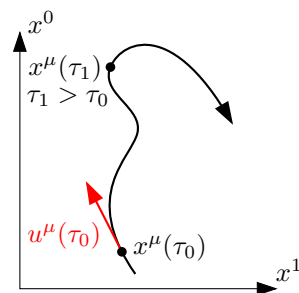
- (v) Um Grössen wie Vektoren, Differentialformen und Tensoren an verschiedenen Ereignissen zu vergleichen, sind weitere Strukturen auf der Raumzeit notwendig, wie zum Beispiel einen *parallelen Transport*.

Für eine rigorose Behandlung der Differentialgeometrie, siehe z.B. [7] (Glatte Mannigfaltigkeiten) und [6] (Riemannsche Mannigfaltigkeiten). Für ihre Bedeutung in der Physik der allgemeinen Relativitätstheorie, siehe [1].

## 6 Zeitreisen in die Vergangenheit?

Wir betrachten die Weltlinie eines Beobachters aus der Sicht eines Inertialsystems  $K$ . Die Koordinaten (in  $K$ ) des Beobachters sind dann Funktionen seiner Eigenzeit:  $x^\mu(\tau)$ . Die Komponenten der momentanen 4er-Geschwindigkeit sind  $u^\mu(\tau) = dx^\mu(\tau)/d\tau$ .

- (a) Wir nehmen an, dass die Weltlinie stetig differenzierbar ist und wie in der Abbildung gegeben ist; insbesondere soll sie in der  $ct$ - $x$ -Ebene verlaufen. Erklären Sie, dass sich der Beobachter aus  $K$  gesehen zuerst in Richtung Zukunft bewegt, dann aber in die Vergangenheit. Argumentieren Sie intuitiv, dass diese Weltlinie für den Beobachter unmöglich ist.



*Hinweis: Lichtstrahlen bewegen sich im Diagramm auf Diagonalen. Sie können hier annehmen, dass die Lichtgeschwindigkeit vom Beobachter nie erreicht werden kann. Nehmen Sie zudem an, dass die 4er-Geschwindigkeit während der ganzen Reise nie null ist, also dass ihre Komponenten nie alle gleichzeitig null sind.*

Wir werden nun zeigen, dass die Lichtgeschwindigkeit nie erreicht werden kann und folglich Zeitreisen in die Vergangenheit wie in (a) nicht möglich sind.

- (b) Betrachten Sie  $(u_\mu u^\mu)(\tau) = (u^0(\tau))^2 - \sum_{k=1}^3 (u^k(\tau))^2$  und zeigen Sie, dass dieser Ausdruck  $> 0$  ist, falls sich der Beobachter in  $K$  gerade mit weniger als Lichtgeschwindigkeit bewegt<sup>1</sup>, und  $= 0$ , wenn er sich gerade mit Lichtgeschwindigkeit bewegt.
- (c) Wir können nun zu jedem Eigenzeitpunkt  $\tau$  ein Inertialsystem  $\tilde{K}(\tau)$  finden, in dem der Beobachter momentan ruht (ein sogenanntes *momentanes Ruhesystem*). Argumentieren Sie, dass für die momentanen 4er-Geschwindigkeitskomponenten in  $\tilde{K}(\tau)$  gilt:  $\tilde{u}^\mu(\tau) = (1, 0, 0, 0)$ .
- (d) Argumentieren Sie, dass  $\tilde{u}^\mu(\tau) = (\Lambda(\tau)^{-1})^\mu_\nu u^\nu(\tau)$ , wobei  $(\Lambda(\tau)^{-1})^\mu_\nu$  die Lorentz-Transformation von  $K$  nach  $\tilde{K}(\tau)$  ist. Schlüssen Sie daraus, dass  $(u_\mu u^\mu)(\tau) = 0$  nicht vorkommen kann.

*Hinweis:  $(u_\mu u^\mu)(\tau)$  ist Lorentz-invariant.*

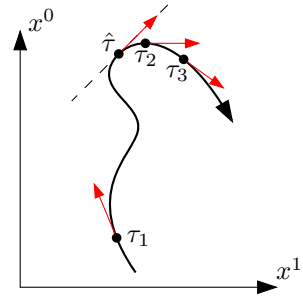
Zum Schluss verallgemeinern wir unser Resultat.

- (e) Betrachten Sie eine allgemeine Weltlinie, die zuerst mit weniger als Lichtgeschwindigkeit in die Zukunft führt, später aber in die Vergangenheit. Insbesondere sei sie nicht auf die  $ct$ - $x$ -Ebene beschränkt. Zeigen Sie, dass solch eine Weltlinie für Beobachter unmöglich sind.
- (\*f) Warum war die Annahme weiter oben, dass die 4er-Geschwindigkeitskomponenten nie alle null sein können, sinnvoll?

<sup>1</sup>Mit (c) und (d) lässt sich später zeigen, dass genauer  $u^\mu u_\mu = 1$ .

## Lösung

- (a) Wir betrachten die Weltlinie an den Eigenzeiten  $\tau_1$  und  $\tau_3$ , siehe Abbildung rechts. Die 4er-Geschwindigkeit zeigt bei  $\tau_1$  nach oben, sprich,  $u^0(\tau_1) > 0$ ; und da  $x^\mu(\tau_1 + \Delta\tau) = x^\mu(\tau_1) + u^\mu \Delta\tau + \mathcal{O}(\Delta\tau^2)$ , bewegt sich der Beobachter an diesem Ereignis in Richtung Zukunft. Umgekehrt ist es beim Ereignis mit Eigenzeit  $\tau_3$ : dort ist  $u^0(\tau_3) < 0$  und der Beobachter bewegt sich in Richtung Vergangenheit.



Nun hängt die 4er-Geschwindigkeit stetig von der Eigenzeit ab. Es muss also einen Eigenzeitpunkt  $\tau_2$  geben, wo sich der Beobachter weder in die Zukunft noch in die Vergangenheit bewegt. Genau genommen folgt dies mit dem Zwischenwertsatz für  $u^0$ . Per Annahme ist  $u^1(\tau_2) \neq 0$ .

Aber auch die geometrische Darstellung der 4er-Geschwindigkeit hängt stetig von  $\tau$  ab. Es muss also auch ein Eigenzeitpunkt  $\hat{\tau}$  zwischen  $\tau_1$  und  $\tau_2$  geben, wo die 4er-Geschwindigkeit um genau  $45^\circ$  geneigt ist und sich der Beobachter mit Lichtgeschwindigkeit bewegt.

Aber die Lichtgeschwindigkeit kann er nie erreichen. Somit ist die ganze Weltlinie unphysikalisch.

*Bemerkungen:*

- (i) Die Bezeichnungen „Richtung Zukunft“ und „Richtung Vergangenheit“ beziehen sich stets auf ein Inertialsystem, hier auf  $K$ . Für eine Weltlinie, die sich aber mit weniger als Lichtgeschwindigkeit Richtung Zukunft oder Vergangenheit bewegt, stimmen die Bedeutungen in jedem Inertialsystem überein.
  - (ii) Dass die 4er-Geschwindigkeit bei  $\tau_1$  steiler als  $45^\circ$  nach oben zeigt, bedeutet, dass sich der Beobachter mit weniger als Lichtgeschwindigkeit bewegt, siehe (b). Wenn wir in der Praxis wie hier eine Zeitreise versuchen möchten, dann ist dies am Anfang der Reise sicherlich der Fall.
  - (iii) Obiges Argument zeigt auch, dass sich der Beobachter nicht mit Überlichtgeschwindigkeit bewegen kann, wenn er sich vorher mit weniger als Lichtgeschwindigkeit bewegt hat. Denn dazu müsste er zuerst Lichtgeschwindigkeit erreichen.
  - (iv) Das eher geometrische Argument für die Existenz von  $\hat{\tau}$  kann mit dem Resultat von (b) rigoros geführt werden, siehe Bemerkung dort.
- (b) Für Unterlichtgeschwindigkeit ist  $(u^0)^2 > \sum_{k=1}^3 (u^k)^2$ , denn

$$c^2 > \sum_{k=1}^3 \left( \frac{dx^k}{dt} \right)^2 = \sum_{k=1}^3 \left( \frac{dx^k}{d\tau} \right)^2 \cdot \left( \frac{d\tau}{dt} \right)^2 \quad \xrightarrow{d\tau/dt \neq 0} \quad c^2 \cdot \left( \frac{dt}{d\tau} \right)^2 > \sum_{k=1}^3 \left( \frac{dx^k}{d\tau} \right)^2.$$

Also ist  $u^\mu u_\mu > 0$ . Dass sogar  $u^\mu u_\mu = 1$ , wird noch folgen.

Wir können alternativ auch eher geometrisch argumentieren: für kleine  $\Delta t$  muss  $|\vec{v}| \approx |\Delta \vec{x}| / \Delta t < c$ , also  $c \Delta t > |\Delta \vec{x}|$ . Da die 4er-Geschwindigkeit tangential zur Weltlinie ist, folgt nun  $u^0 > |\vec{u}|$ , bzw.  $(u^0)^2 > \sum_{k=1}^3 (u^k)^2$ . Dieses Argument wird genau, wenn  $\Delta t \rightarrow 0$ .

Für Lichtgeschwindigkeit müssen wir in den obigen Ausdrücken  $>$  mit  $=$  ersetzen und erhalten das gewünschte Resultat.

*Bemerkung:* Bei  $\tau_1$  haben wir  $(u^0)^2 > \sum_{k=1}^3 (u^k)^2$ , da dort die Geschwindigkeit kleiner als  $c$  ist. Bei  $\tau_2$  ist dann  $(u^0)^2 = 0 < \sum_{k=1}^3 (u^k)^2$ ; es gilt  $<$  statt  $\leq$ , weil Gleichheit  $u^\mu(\tau_2) \equiv 0$  bedeuten würde. Mit dem Zwischenwertsatz folgt, dass es ein  $\hat{\tau}$  geben muss, mit  $(u^0)^2 = \sum_{k=1}^3 (u^k)^2$ , also ist die Weltlinie dort lichtschnell. Dies zeigt den letzten Teil von (a) rigoros.

- (c) Die Eigenzeit  $\tau$  ist so definiert, dass  $d\tilde{x}^0/d\tau = 1$  im momentanen Ruhesystem (das Inertialsystem, in dem der Beobachter momentan ruht). Das heisst: eine Uhr, die der möglicherweise beschleunigte Beobachter mitträgt (diese definiert die Eigenzeit), soll so eingestellt sein, dass sie gleich schnell tickt, wie eine inertielle Uhr, die sich momentan mit gleicher Geschwindigkeit wie der Beobachter bewegt. Im Momentanen Ruhesystem  $\tilde{K}$  ist also per Definition  $d\tilde{x}^0/d\tau = 1$ . Zudem bewegt sich der Beobachter momentan nicht, also  $d\vec{\tilde{x}}/d\tau = d\vec{\tilde{x}}/d\tilde{x}^0 \cdot d\tilde{x}^0/d\tau = 0$ . Sprich,  $\tilde{u}^\mu = (1, 0, 0, 0)$ .

*Bemerkungen:*

- (i) Dies gilt in geometrischen Einheiten. Wenn wir stattdessen die Faktoren  $c$  beibehalten, kommt je nach Einheitenkonvention  $u^\mu = (c, 0, 0, 0)$  oder  $u^\mu = (1, 0/c, 0/c, 0/c) = (1, 0, 0, 0)$  heraus.
  - (ii) Dass der Beobachter überhaupt ein momentanes Ruhesystem hat, ist eine Folge davon, dass er ein Beobachter ist. Denn um beobachten zu können, muss er ein Bezugssystem/Koordinatensystem mit einer Zeitachse und drei Raumachsen um sich herum konstruieren können; und wenn dies funktioniert, können wir ein momentanes Ruhesystem finden, indem wir einfach jegliche Beschleunigung vom vorher konstruierten Bezugssystem entfernen.
- (d) Aus (c) folgt  $\tilde{u}^\mu \tilde{u}_\mu = 1$  im momentanen Ruhesystem. Wir betrachten die Lorentz-Transformation vom momentanen Ruhesystem nach  $K$  (d.h. die Inverse der Transformation in der Aufgabenstellung). Dann können wir wie beschrieben die Komponenten  $u^\mu$  der 4er-Geschwindigkeit in  $K$  berechnen. Da das Quadrat eines 4er-Vektors Lorentz-invariant ist, folgt  $u^\mu u_\mu = \tilde{u}^\mu \tilde{u}_\mu = 1$ . Wenn wir nun aber den Eigenzeitpunkt  $\hat{\tau}$  betrachten, haben wir gesehen, dass  $u^\mu u_\mu = 0$ ; an  $\tilde{u}^\mu \tilde{u}_\mu = 1$  ändert dies aber nichts. Somit müsste an diesem Ereignis  $0 = 1$  gelten, ein Widerspruch.

Es folgt, dass ein lichtschneller Beobachter kein momentanes Ruhesystem besitzen kann, also kann er kein wirklicher Beobachter sein. Beobachter können sich also nicht mit Lichtgeschwindigkeit bewegen. Schliesslich folgt, dass Zeitreisen wie in (a) nicht möglich sind.

*Bemerkungen:*

- (i) Es folgt auch, dass Eigenzeit nur für Weltlinien mit einer Geschwindigkeit kleiner als Lichtgeschwindigkeit definiert sind. Und genauso sind 4er-Geschwindigkeiten nur für solche definiert. Wir können aber allgemeinere Kurven  $x^\mu(\lambda)$  eines allgemeineren Parameters betrachten, wenn  $\lambda$  keine Zeit ist.
  - (ii) Ist  $u^\mu$  definiert, haben wir  $u^\mu u_\mu = 1$ , wie in (b) behauptet.
- (e) Wie in (a) muss es einen Eigenzeitpunkt  $\tau_2$  mit  $u^0 = 0$  geben. Mit der Bemerkung von (b) folgt dann, dass es auch einen Eigenzeitpunkt  $\hat{\tau}$  gibt, wo sich der Beobachter mit Lichtgeschwindigkeit bewegt, was gemäss (d) nicht möglich ist.
- (\*f)  $u^\mu = 0$  impliziert  $u_\mu u^\mu = 0$ , was gemäss (d) unmöglich ist.

Wir können dies auch leicht anders mittels (c) und (d) sehen:  $\tilde{u}^\mu = (1, 0, 0, 0)$  im momentanen Ruhesystem bedeutet, dass  $u^\mu$  nicht komplett null sein kann, weil Lorentz-Transformationen linear und invertierbar sind, also einen trivialen Nullraum besitzen.

*Bemerkung:* Die soeben motivierte Annahme wurde genau genommen nur in (a), (b) und somit (e) verwendet. (c) und (d) waren von der Annahme unabhängig und wir konnten diese Resultate verwenden, um die Annahme zu motivieren.

*Bemerkung:* Wir haben Vergangenheitsreisen von Beobachtern in der speziellen Relativitätstheorie ausgeschlossen. Es gibt jedoch andere Arten, wie solche Zeitreisen potentiell möglich sind, vor allem im Rahmen der allgemeinen Relativitätstheorie. Es wird aber stark davon ausgegangen, dass selbst die unmöglich sind.

## Weiterführende Bemerkungen und Literatur

Teilaufgabe (\*f) sowie die Annahme, die damit motiviert wird, war nicht Teil der ursprünglichen Aktivität und wurde der Vollständigkeit halber später ergänzt.

Weder die Aktivität noch die folgende Bemerkung basieren speziell auf einer Quelle. Folgende Bücher behandeln aber die spezielle Relativitätstheorie: [2] (eher intuitiv, mit vielen Illustrationen) und [17] (eher ausführlich).

**Alternative Argumente für die Unerreichbarkeit der Lichtgeschwindigkeit** Neben dem hier gezeigten Argument, existieren eine Reihe einfacherer Argumente, die zeigen, dass Beobachter Lichtgeschwindigkeit nie erreichen können:

- (i) Angenommen, ein inertialer Beobachter bewegt sich mit Lichtgeschwindigkeit. Sein Ruhesystem ist dann ein Inertialsystem und gemäss Relativitätsprinzip muss sich Licht darin mit Lichtgeschwindigkeit ausbreiten. Dies ist nicht vereinbar damit, dass das Ruhesystem des Beobachters auch ein Ruhesystem für das Licht sein müsste. Es folgt, dass solch ein Beobachter nicht existieren kann, und dass Licht kein Ruhesystem haben kann.
- (ii) Die Energie eines massebehafteten Körpers (z.B. eines Beobachters) strebt gegen unendlich, wenn sich der Körper in einem Inertialsystem der Lichtgeschwindigkeit annähert. Mit endlicher Energie kann Lichtgeschwindigkeit somit nicht erreicht werden.
- (iii) Die Matrix eines Lorentz-Boosts hat divergierende Einträge, wenn die Boostgeschwindigkeit Lichtgeschwindigkeit erreicht. Gleichermassen werden Zeitdilatation und Längenkontraktion singular. Dies ist ein Anzeichen dafür, dass die momentanen Ruhesysteme von lichtschnellen Beobachtern (wenn sie denn existieren), sicherlich keine herkömmlichen Inertialsysteme sein können. Somit kann ein lichtschneller Beobachter sicherlich kein herkömmlicher Beobachter sein.

Wir haben in der Aktivität ein eher technisches Argument verwendet, weil wir so verschiedene Aspekte von 4er-Geschwindigkeit, momentanen Ruhesystemen und Lorentz-Transformationen anwenden konnten.

# 7 \* Ungleichungen für zeitartige 4er-Vektoren

Ein zeitartiger 4er-Vektor  $v$  mit Komponenten  $v^\mu$  heisst *zukunftsgerichtet*, falls  $v^0 > 0$ <sup>1</sup>.

Wir schreiben kurz  $(a, b) := \eta_{\mu\nu} a^\mu b^\nu$  für das Skalarprodukt zweier 4er-Vektoren. Zur Erinnerung:  $v$  ist *zeitartig*, falls  $(v, v) > 0$ , *raumartig*, falls  $(v, v) < 0$ , und *lichtartig*, falls  $(v, v) = 0$ .

## Mathematischer Teil

- (a) Zeigen Sie, dass zeitartige 4er-Vektoren  $v$  und  $w$  die *umgekehrte Cauchy-Schwarz-Ungleichung* erfüllen:

$$(v, w)^2 \geq (v, v) \cdot (w, w), \quad =, \text{ gdw. } v \propto w.$$

*Hinweis: Wechseln Sie in ein Inertialsystem, in dem  $v^\mu = (v^0, 0, 0, 0)$ .*

- (b) Zeigen Sie, dass zeitartige, zukunftsgerichtete 4er-Vektoren  $v$  und  $w$  die *umgekehrte Dreiecksungleichung* erfüllen:

$$\sqrt{(v+w, v+w)} \geq \sqrt{(v, v)} + \sqrt{(w, w)}, \quad =, \text{ gdw. } v = \lambda w, \lambda > 0.$$

*Hinweis: Verwenden Sie wieder das Inertialsystem aus (a), sowie die in (a) bewiesene Ungleichung.*

- (c) Beweisen Sie die allgemeinere umgekehrte Dreiecksungleichung:

$$\sqrt{\left( \sum_{i=1}^N v_i, \sum_{i=1}^N v_i \right)} \geq \sum_{i=1}^N \sqrt{(v_i, v_i)},$$

wobei alle  $v_i$  zeitartige, zukunftsgerichtete 4er-Vektoren sind. Gleichheit gilt, wenn alle  $v_i$  in die gleiche Richtung zeigen.

## Physikalischer Teil

- (d) Betrachten Sie ein Gas aus  $N$  relativistischen Teilchen mit 4er-Impulsen  $p_i$  und Ruhemassen  $m_i$ ,  $i = 1, \dots, N$ . Zeigen Sie, dass die Gesamtmasse des Systems folgende Ungleichung erfüllt:

$$M \geq \sum_{i=1}^N m_i, \quad =, \text{ gdw. alle Teilchen mit derselben Geschwindigkeit bewegt sind.}$$

*Hinweis: Es ist  $m_i^2 = (p_i, p_i)$  und  $M^2 = (p, p)$ , wobei  $p$  der gesamte 4er-Impuls ist.*

- (e) Seien  $A$  und  $B$  zwei Ereignisse mit zeitartigem 4er-Abstand. Zeigen Sie, dass die unbeschleunigte Weltlinie zwischen  $A$  und  $B$  von allen möglichen Weltlinien zwischen  $A$  und  $B$  am meisten Eigenzeit benötigt (allgemeines Zwillingsparadoxon).

*Hinweis: Weiten Sie die umgekehrte Dreiecksungleichung von Summen auf Integrale von 4er-Vektoren aus.*

<sup>1</sup>Man kann zeigen, dass  $v^0 > 0$  unter Transformationen  $\Lambda \in L^\uparrow$  invariant ist. Die „Zukunfts-Gerichtetheit“ eines zeitartigen 4er-Vektors macht also auch unabhängig von Koordinaten Sinn, solange wir keine Zeitspiegelung vornehmen.

# Lösung

## Mathematischer Teil

- (a) **Beweis.** Wir wechseln in ein Inertialsystem, in dem  $v^\mu = (v^0, 0, 0, 0)$ . Solch ein Inertialsystem können wir durch eine Rotation der Koordinaten und anschließenden Boost in  $x^1$ -Richtung immer finden. Wir haben dann  $(v, w)^2 = (v^0)^2(w^0)^2$  und  $(v, v) = (v^0)^2$ . Auch ist  $(w, w) = (w^0)^2 - |\vec{w}|^2$ , wobei  $\vec{w}$  die Raumkomponenten von  $w$  sind. Aber da  $|\vec{w}| \geq 0$ , ist  $(w, w) \leq (w^0)^2$ , und die Ungleichung folgt.

Gleichheit gilt, wenn  $|\vec{w}| = 0$ . Dann ist im gewählten Inertialsystem  $w^\mu = (w^0, 0, 0, 0)$  und somit  $v \propto w$ .  $\square$

- (b) **Beweis.** Es ist  $(v+w, v+w) = (v, v) + (w, w) + 2(v, w)$ . Wir verwenden wieder dasselbe Inertialsystem wie in (a). Dort sehen wir, dass  $(v, w) = v^0 w^0 > 0$ , da  $v^0 > 0$  und  $w^0 > 0$ . Nach (a) folgt  $\sqrt{(v, w)^2} \geq \sqrt{(v, v)(w, w)}$ , also

$$(v+w, v+w) \geq (v, v) + (w, w) + 2\sqrt{(v, v)(w, w)} = \left(\sqrt{(v, v)} + \sqrt{(w, w)}\right)^2.$$

Also ist

$$\sqrt{(v+w, v+w)} \geq \sqrt{(v, v)} + \sqrt{(w, w)}.$$

Gleichheit tritt genau dann ein, wenn  $(v, w)^2 = (v, v)(w, w)$ , also wenn  $v \propto w$ , nach (a). Aber dies bedeutet wegen  $v^0 > 0$  und  $w^0 > 0$  sogar  $v = \lambda w$ , mit  $\lambda > 0$ .  $\square$

- (c) **Beweis.** Wir bemerken als erstes, dass die Summe zweier zeitartigen, zukunftsgerichteten 4er-Vektoren wieder zeitartig und zukunftsgerichtet ist. Somit gilt nach der Dreiecksungleichung

$$\sqrt{\sum_{i=1}^N (v_i, v_i)} \geq \sqrt{(v_1, v_1)} + \sqrt{\left(\sum_{i=2}^N v_i, \sum_{i=2}^N v_i\right)}.$$

Die Ungleichung folgt nun mit endlicher Induktion.

Wir betrachten noch die Bedingung für Gleichheit. Dazu müssen insbesondere im letzten Schritt  $v_{N-1}$  und  $v_N$  in die gleiche Richtung zeigen. Aber die Sortierung der  $v_i$  ist im Beweis willkürlich. Deshalb ist eine notwendige Bedingung für Gleichheit, dass die  $v_i$  paarweise in dieselbe Richtung zeigen, bzw. alle in dieselbe Richtung zeigen. Aber dann sind alle  $v_i$  positive Vielfache voneinander und wir sehen schnell, dass die Ungleichung zu einer Gleichung wird.  $\square$

## Physikalischer Teil

- (d) **Beweis.** Der gesamte 4er-Impuls des Systems ist  $p = \sum_{i=1}^N p_i$ , wobei  $p_i$  der 4er-Impuls des  $i$ -ten Teilchens ist. Dann ist mit der allgemeinen Dreiecksungleichung

$$M = \sqrt{(p, p)} = \sqrt{\left(\sum_{i=1}^N p_i, \sum_{i=1}^N p_i\right)} \geq \sum_{i=1}^N \sqrt{(p_i, p_i)} = \sum_{i=1}^N m_i.$$

Gleichheit tritt ein, wenn alle  $p_i$  in dieselbe Richtung zeigen. Somit zeigen auch alle 4er-Geschwindigkeiten  $u_i$  in dieselbe Richtung, weil  $m_i > 0$ . Die Normierung  $(u_i, u_i) = 1$  legt diese 4er-Geschwindigkeiten aber bereits komplett fest; alle Teilchen haben somit dieselbe 4er-Geschwindigkeit und somit auch dieselbe Geschwindigkeit.

Somit ist die Masse eines Systems grösser oder gleich der Summe der Teilmassen; dies steht im Gegensatz zur klassischen Physik, wo Masse immer additiv ist. Die behandelten Ungleichungen lassen sich auch auf lichtartige 4er-Vektoren erweitern (siehe weiterführende Berkeungen unten); wir können also auch Systeme betrachten, die Licht enthalten. Zum Beispiel kann ein System aus Lichtpulsen eine Masse besitzen, obwohl die einzelnen Pulse masselos sind.  $\square$

- (e) **Beweis.** Wir argumentieren hier eher intuitiv.

OBdA. geschieht  $A$  vor  $B$ . Der 4er-Abstand von  $A$  nach  $B$  sei  $\Delta x$ . Für einen inertial bewegten Beobachter, der sich von  $A$  nach  $B$  bewegt, vergeht dabei die Eigenzeit  $\Delta\tau = \sqrt{(\Delta x, \Delta x)}$ .



Wir betrachten eine zweite, allgemeine Weltlinie von  $A$  nach  $B$ , gegeben durch  $x(\tau')$ , wobei  $\tau'$  die Eigenzeit entlang dieser Weltlinie ist. Dabei vergeht trivialerweise die Eigenzeit  $\Delta\tau' = \int_{\tau'_A}^{\tau'_B} d\tau'$ . Nun gilt für die 4er-Geschwindigkeit  $u$  der Weltlinie  $\sqrt{(u, u)} = 1$ , also

$$\Delta\tau' = \int_{\tau'_A}^{\tau'_B} \sqrt{(u, u)} d\tau' = \int_{\tau'_A}^{\tau'_B} \sqrt{\left(\frac{dx}{d\tau'}, \frac{dx}{d\tau'}\right)} d\tau' = \int_A^B \sqrt{(dx, dx)}.$$

Das Integral auf der rechten Seite kann als Grenzwert einer Riemann-Summe gesehen werden: vor dem Grenzwert haben wir statt  $dx$  endliche 4er-Abstände  $\Delta x_i$ . Diese sind zeitartig und zukunftsgerichtet. Bevor der Grenzwert genommen wird, gilt also die allgemeine Dreiecksungleichung; im Grenzwert kann die Ungleichung bloss eine Ungleichung bleiben, also folgt

$$\Delta\tau' = \int_A^B \sqrt{(dx, dx)} \leq \left[ \left( \underbrace{\int_A^B dx}_{=\Delta x}, \underbrace{\int_A^B dx}_{=\Delta x} \right) \right]^{1/2} = \sqrt{(\Delta x, \Delta x)} = \Delta\tau.$$

□

## Weiterführende Bemerkungen

Diese Aktivität wurde im Unterricht nicht behandelt. Sie setzt auch eher auf mathematische Intuition als physikalische, und ist etwas schwieriger als andere Aktivitäten.

**Erweiterung um lichtartige 4er-Vektoren** Die umgekehrte Cauchy-Schwarz-Ungleichung gilt immer noch, wenn  $v$  und/oder  $w$  lichtartig sind.

**Beweis.** Die linke Seite der Ungleichung ist stets  $\geq 0$ .

Sind  $v$  und  $w$  lichtartig, so ist die rechte Seite null und die Ungleichung gilt. Gleichheit gilt, wenn  $(v, w) = v^0 w^0 - v^1 w^1 = 0$ , wobei wir die Koordinaten so gedreht haben, dass  $v^\mu = (v^0, v^1, 0, 0)$ ; hier haben wir  $|v^0| = |v^1|$ . Gleichheit kann also nur gelten, wenn auch  $|w^0| = |w^1|$ , also  $w^\mu = (w^0, w^1, 0, 0)$  in diesen Koordinaten. Zudem muss  $\text{sgn}(v^0) \cdot \text{sgn}(w^0) = \text{sgn}(v^1) \cdot \text{sgn}(w^1)$ , was  $v \propto w$  bedeutet.

Sei nun  $v$  zeitartig und  $w$  lichtartig. Wir wechseln in ein Inertialsystem, wo  $v^\mu = (v^0, 0, 0, 0)$ . Falls  $w = 0$ , gilt Gleichheit trivialerweise. Sei OBdA.  $w \neq 0$ ; es folgt sofort  $v \not\propto w$ . Die rechte Seite der Ungleichung ist wieder null; wir müssen nur noch zeigen, dass die linke Seite  $> 0$  ist. Tatsächlich:  $(v, w) = v^0 w^0 \neq 0$ , da  $v^0 \neq 0$  und  $w^0 \neq 0$ . □

Auch die umgekehrte Dreiecksungleichung in (c) gilt, wenn  $v$  und/oder  $w$  lichtartig sind, solange beide zukunftsgerichtet sind.

**Beweis.** Seien beide Vektoren lichtartig (aber nicht null, weil sie zukunftsgerichtet sind). Dann ist die rechte Seite der Ungleichung null. Analog zum Beweis in der Aktivität müssen wir nur  $(v, w) \geq 0$  zeigen; die Ungleichung folgt dann mit der soeben bewiesenen umgekehrten Cauchy-Schwarz-Ungleichung. Aber

$$(v, w) = \underbrace{v^0 w^0}_{=|\vec{v}||\vec{w}|} - \vec{v} \cdot \vec{w} \geq 0,$$

wobei die Ungleichung aus der herkömmlichen Cauchy-Schwarz-Ungleichung folgt. Gleichheit tritt bei  $(v, w) = 0$ , also  $v \propto w$  ein.

Sei nun  $v$  zeitartig und  $w$  lichtartig. Auch hier ist

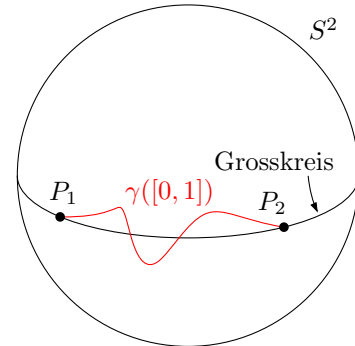
$$(v, w) = v^0 w^0 - v^1 w^1 \geq 0,$$

wobei wir in ein Inertialsystem mit  $v^\mu = (v^0, v^1, 0, 0)$  gewechselt sind und verwendet haben, dass  $v^0 > |v^1|$ . Die Ungleichung folgt wie oben. Gleichheit tritt ein, wenn  $(v, w) = 0$ , also  $v \propto w$  (was aber hier ohnehin nicht passieren kann). □

Schliesslich kann auch die allgemeine umgekehrte Dreiecksungleichung auf lichtartige, zukunftsgerichtete 4er-Vektoren erweitert werden. Die notwendigen Anpassungen im Beweis von (c) sind aber trivial.

# 8 Geodäten der Kugeloberfläche $S^2$

Wir haben zwei Punkte  $P_1$  und  $P_2$  auf der zweidimensionalen Oberfläche der Kugel mit Radius 1,  $S^2$ , gegeben. Wir möchten zeigen, dass die kürzeste Verbindungslinie zwischen  $P_1$  und  $P_2$ , welche gänzlich auf der Oberfläche verläuft, das kürzere Segment des *Grosskreises* durch  $P_1$  und  $P_2$  ist.



Ein Kreis auf  $S^2$  heisst *Grosskreis*, wenn er  $S^2$  in zwei Halbkugeln teilt. Auf der Erde sind dies z.B. der Äquator und alle Längengrade.

Wir verwenden Kugelkoordinaten  $(\theta, \phi)$  die wir später noch geseheit drehen werden.

- (a) Wir parametrisieren eine allgemeine Kurve auf  $S^2$  von  $P_1$  nach  $P_2$  durch eine Funktion  $\gamma : [0, 1] \rightarrow S^2$ ,  $\lambda \mapsto P(\theta(\lambda), \phi(\lambda))$ , wobei  $P(\theta, \phi)$  der Punkt mit Koordinaten  $\theta$  und  $\phi$  ist. Zeigen Sie, dass die Länge dieser Kurve durch folgendes Funktional gegeben ist:

$$L[\gamma] := \int_0^1 \ell \, d\lambda, \quad \ell := \sqrt{\theta'^2 + \phi'^2 \sin^2 \theta}.$$

Um die kürzeste Verbindung zu finden, können wir die Euler-Lagrange-Gleichungen (E-L) lösen, die der Variation  $\delta L[\gamma] = 0$  entsprechen. Dies stellt sich aber als ziemlich schwierig heraus (Bonusfrage: wieso?). Stattdessen betrachten wir ein einfacheres Funktional

$$\tilde{L}[\gamma] := \int_0^1 \tilde{\ell} \, d\lambda, \quad \tilde{\ell} := \theta'^2 + \phi'^2 \sin^2 \theta.$$

Wir werden in (e) sehen, dass  $\delta \tilde{L}[\gamma] = 0 \Rightarrow \delta L[\gamma] = 0$ ; somit können wir Geodäten auch mit  $\tilde{L}$  bestimmen. Dies werden wir jetzt tun.

Zur Erinnerung: die Euler-Lagrange-Gleichungen äquivalent zu  $\delta \tilde{L}[\gamma] = 0$  sind

$$\frac{\partial \tilde{\ell}}{\partial \phi} = \frac{d}{d\lambda} \frac{\partial \tilde{\ell}}{\partial \phi'}, \quad \frac{\partial \tilde{\ell}}{\partial \theta} = \frac{d}{d\lambda} \frac{\partial \tilde{\ell}}{\partial \theta'}.$$

- (b) Zeigen Sie, dass  $\phi' \sin^2 \theta = K = \text{konst.}$
- (c) Zeigen Sie mit Hilfe der E-L-Gleichung für  $\theta$ , dass  $\theta'' = K^2 \cos \theta \sin^{-3} \theta$ .
- (d) Angenommen,  $(\theta(\lambda), \phi(\lambda))$  löst die Differentialgleichungen von (b) und (c), und verbindet  $P_1$  mit  $P_2$ ; sprich, das entsprechende  $\gamma([0, 1])$  ist eine Verbindungslinie mit extremaler Länge. Argumentieren Sie, dass wir die Koordinaten bei festgehaltenem  $\gamma([0, 1])$  so drehen können, dass  $\theta(0) = \pi/2$ ,  $\theta'(0) = 0$ . Folgern Sie, dass  $\theta(\lambda) = \pi/2 = \text{konst.}$  und  $\gamma([0, 1])$  somit auf dem Äquator liegt, einem Grosskreis. Beweisen Sie somit die Aussage in der Einleitung.

*Hinweis:*  $\gamma([0, 1])$  ist ein geometrisches Objekt.  $\gamma([0, 1])$  sollte also das Bild der Lösungskurve sein, egal wie die Koordinaten gedreht werden.

- (e) Beweisen Sie, dass  $\delta \tilde{L}[\gamma] = 0 \Rightarrow \delta L[\gamma] = 0$ . Sprich, dass eine Lösung  $\gamma$  der E-L-Gleichungen für  $\tilde{L}[\gamma]$  auch die E-L-Gleichungen für  $L[\gamma]$  löst.

*Hinweis:*  $\tilde{\ell} = \ell^2$ . Wegen  $\frac{\partial \tilde{\ell}}{\partial \lambda} = 0$  ist  $\theta' \frac{\partial \tilde{\ell}}{\partial \theta'} + \phi' \frac{\partial \tilde{\ell}}{\partial \phi'} - \tilde{\ell} = \tilde{\ell} = \text{konst.}$

## Lösung

- (a) Das infinitesimale Linienelement ist  $dL = \sqrt{d\theta^2 + d\phi^2 \sin^2 \theta} = d\lambda \sqrt{\theta'^2 + \phi'^2 \sin^2 \theta}$ . Somit ist die Länge

$$L[\gamma] = \int_0^1 \ell \, d\lambda, \quad \ell = \sqrt{\theta'^2 + \phi'^2 \sin^2 \theta}.$$

*Bemerkung:* Die Extremierung des Funktionals  $L[\gamma]$  kann bloss die Form der Verbindungskurve festlegen, nicht aber die Parametrisierung. Denn  $L[\gamma]$  ist unter Reparametrisierung invariant, da dies die Länge unverändert belässt. Das heisst, wenn  $\gamma$  eine Lösung ist, so ist es auch  $\gamma' = \gamma \circ f$ , wobei  $f : [0, 1] \rightarrow [0, 1]$  eine stetige, monoton steigende Bijektion ist. Also haben die E-L-Gleichungen, welche aus  $\delta L[\gamma] = 0$  folgen, eine überabzählbare Menge an Lösungen, selbst wenn wir die Randbedingungen (Anfangs- und Endpunkt) wie hier vorgeben. Folglich ist es schwierig, die E-L-Gleichungen zu lösen.

Aus diesem Grund wechseln wir zum anderen Funktional  $\tilde{L}[\gamma]$ . Die Idee ist, dass  $\tilde{L}[\gamma]$  dem eigentlichen Längenfunktional  $L[\gamma]$  entspricht, wenn  $\tilde{\ell} = \text{konst.}$  mit einem bestimmten Wert. Wir werden noch sehen, dass  $\tilde{\ell}$  tatsächlich konstant ist und durch die Variation gerade den richtigen Wert erhält. Somit können wir statt  $L[\gamma]$  auch  $\tilde{L}[\gamma]$  variieren. Zudem fixiert uns die Variation von  $\tilde{L}[\gamma]$  bereits eine praktische Parametrisierung und vereinfacht so die Rechnung.

- (b) Weil  $\partial \tilde{\ell} / \partial \phi = 0$ , ist gemäss E-L-Gleichung für  $\phi$

$$\frac{\partial \tilde{\ell}}{\partial \phi'} = 2\phi' \sin^2 \theta = \text{konst.} =: 2K.$$

- (c) Die E-L-Gleichung für  $\theta$  ist

$$\frac{\partial \tilde{\ell}}{\partial \theta} = \frac{d}{d\lambda} \frac{\partial \tilde{\ell}}{\partial \theta'}.$$

Also ist

$$\frac{\partial \tilde{\ell}}{\partial \theta} = 2\phi'^2 \cos \theta \sin \theta = 2K^2 \cos \theta \sin^{-3} \theta \stackrel{\text{E-L}}{=} \frac{d}{d\lambda} \frac{\partial \tilde{\ell}}{\partial \theta'} = \frac{d}{d\lambda} (2\theta') = 2\theta''.$$

- (d) Das Bild  $C := \gamma([0, 1]) = P(\theta([0, 1]), \phi([0, 1])) \subset S^2$  der Lösungskurve muss komplett von Koordinaten unabhängig sein, wenn es die kürzeste Verbindungslinie zwischen  $P_1$  und  $P_2$  beschreiben soll; denn die kürzeste Verbindungslinie ist sicherlich nur durch die Geometrie von  $S^2$  sowie durch die beiden Punkte  $P_1$  und  $P_2$  bestimmt.

Wir können deshalb unser Koordinatensystem beliebig drehen, die Gleichungen aus (b) und (c) lösen, und werden als Bild der Lösungskurve dieselbe Teilmenge  $C \subset S^2$  erhalten. Ohne Beschränkung der Allgemeinheit können wir die Koordinaten also so ausrichten, dass  $\theta(0) = \pi/2$  und  $\theta'(0) = 0$  (Start am Äquator und parallel zum Äquator, der durch die Wahl der Koordinaten definiert wird). Dann folgt aus (c) aber, dass  $\theta(\lambda) = \pi/2$  für alle  $\lambda$ . Die Verbindungslinie liegt also komplett auf dem Äquator, einem Grosskreis.

Wenn jetzt (e) stimmt, dann ist die kürzeste Verbindung zwischen  $P_1$  und  $P_2$  der kleinste Grosskreisbogen zwischen den Punkten. Meistens ist dieser Kreisbogen eindeutig. Wenn aber  $P_1$  und  $P_2$  antipodal sind, d.h. genau gegenüberliegen, dann gibt es unendlich viele kürzeste Verbindungslinien.

- (e)  $\delta \tilde{L}[\gamma] = 0$  impliziert die E-L-Gleichung

$$\frac{\partial \tilde{\ell}}{\partial \theta} = \frac{d}{d\lambda} \frac{\partial \tilde{\ell}}{\partial \theta'}.$$

Nun ist  $\tilde{\ell} = \ell^2$ , also entspricht die letzte Gleichung

$$2\ell \frac{\partial \ell}{\partial \theta} = \frac{d}{d\lambda} \left( 2\ell \frac{\partial \ell}{\partial \theta'} \right) = 2 \frac{d\ell}{d\lambda} \frac{\partial \ell}{\partial \theta'} + 2\ell \frac{d}{d\lambda} \frac{\partial \ell}{\partial \theta'}. \quad (8.1)$$

Da  $\partial \tilde{\ell} / \partial \lambda = 0$ , ist gemäss Hinweis

$$0 = \frac{d}{d\lambda} \left( \theta' \frac{\partial \tilde{\ell}}{\partial \theta'} + \phi' \frac{\partial \tilde{\ell}}{\partial \phi'} - \tilde{\ell} \right) = \frac{d}{d\lambda} \left( 2\theta'^2 + 2\phi'^2 \sin^2 \theta - \tilde{\ell} \right) = \frac{d}{d\lambda} \tilde{\ell}.$$

Aber dann ist auch

$$\frac{d}{d\lambda}\ell = \frac{d}{d\lambda}\sqrt{\tilde{\ell}} = 0.$$

Damit die Kurve überhaupt von  $P_1$  nach  $P_2$  kommen kann, muss  $\tilde{\ell} > 0$  und somit  $\ell > 0$ .

Also fällt in (8.1) der Term mit  $d\ell/d\lambda$  weg und wir können insgesamt durch  $2\ell$  teilen:

$$\frac{\partial\ell}{\partial\theta} = \frac{d}{d\lambda}\frac{\partial\ell}{\partial\theta'}.$$

Dies ist nun genau die E-L-Gleichung für  $\theta$ , welche  $\delta L[\gamma] = 0$  entspricht.

Für  $\phi$  funktioniert dasselbe Argument ebenfalls. Wir sehen also, dass eine Lösung der E-L-Gleichungen für  $\tilde{L}[\gamma]$  ebenfalls eine Lösung für  $L[\gamma]$  ist. Sprich, wir können Geodäten auch durch Extremierung des Funktionals  $\tilde{L}[\gamma]$  finden.

*Bemerkung: Der Unterschied in den Funktionalen liegt wie oben bemerkt in der Parametrisierung.  $L[\gamma]$  legt die Parametrisierung nicht fest,  $\tilde{L}[\gamma]$  hingegen schon: sie ist über  $d\ell/d\lambda = 0$  und die Randwerte  $\gamma(0) = P_1$ ,  $\gamma(1) = P_2$  gegeben. Eine Parametrisierung mit „konstanter Geschwindigkeit“  $\ell = \text{konst.}$  wird auch affin genannt.  $\tilde{L}[\gamma]$  führt somit natürlicherweise zu affinen Parametrisierungen.*

## Weiterführende Bemerkungen und Literatur

Geodäten auf Oberflächen sind eines der vielen Themen der Differentialgeometrie von *Riemannschen Mannigfaltigkeiten* (grob gesagt Mannigfaltigkeiten mit Metrik, siehe auch weiterführende Bemerkungen in Aktivität 5). Für eine ausführliche Behandlung siehe z.B. [7] (Grundlagen von Mannigfaltigkeiten) und [6] (Riemannsche Mannigfaltigkeiten).

**Intuition hinter dem Funktional  $\tilde{\ell}$**  Die Intuition hinter dem Längen-Funktional  $\ell$  ist klar:  $\ell\Delta\lambda$  ist die approximative Länge eines Kurvensegments zwischen zwei Punkten, die um die kleine Parameterdifferenz  $\Delta\lambda$  auseinanderliegen. Diese Bedeutung wird genau, wenn  $\Delta\lambda \rightarrow 0$ .

Die Intuition hinter  $\tilde{\ell} = \ell^2$ , dem sogenannten *Energie-Funktional*, ist hingegen weniger offensichtlich. Wir möchten sie hier aufzeigen. Zwei Intuitionen sind möglich:

- (i) Wenn wir  $\lambda$  als Zeit interpretieren, dann ist  $\tilde{\ell}$  bis auf den Faktor  $m/2$  die kinetische Energie eines Teilchens der Masse  $m$ , deshalb der Name.

Wenn das Teilchen frei ist, dann ist dies auch bereits die Lagrange-Funktion des Teilchens. Die Kurve, die  $\tilde{L}$  extremiert, kann also als Flugbahn eines freien Teilchens gesehen werden, dessen Bewegung auf die Oberfläche der Kugel (oder irgendeines anderen Körpers) beschränkt ist.

Im dreidimensionalen Raum bewegt sich das Teilchen in einer geraden Linie, also wird sich das Teilchen auf der Kugel *lokal* auf einer Linie bewegen, die der Projektion einer Geraden auf die Kugeloberfläche entspricht. So erhalten wir grob gesagt die Grosskreise, die wir in der Aktivität analytisch gefunden haben. (Auf der Kugel ist der *ganze* Grosskreis die Projektion einer geraden Linie im Raum; für andere Oberflächen sind Geodäten aber nicht immer ganzheitlich Projektionen einer geraden Linie auf die Oberfläche, sondern nur lokal.)

- (ii) Die kürzeste Verbindungslinie zwischen zwei Punkten auf einer Oberfläche kann intuitiv auch durch Spannen eines Gummibands, welches komplett in der Oberfläche zwischen den beiden Punkten verläuft, gefunden werden, weil sich das Gummiband zusammenzieht und somit seine Länge minimiert.

Wir können ein Gummiband als eine Kette aus  $N$  Federn betrachten und später den *Kontinuumsfall*  $N \rightarrow \infty$  betrachten. Die potentielle Energie der Federkette ist

$$U = \sum_{i=1}^N \frac{K(N)}{2} |\Delta\vec{x}_i|^2,$$

wobei  $K(N)$  die Federkonstante einer einzelnen Feder ist.  $\Delta\vec{x}_i$  ist die Auslenkung der Ruhelage der  $i$ -ten Feder.

Damit die Energie bei  $N \rightarrow \infty$  nicht null wird, obwohl  $\Delta \vec{x}_i$  ungefähr linear mit  $N$  abnimmt und die Anzahl Terme linear in  $N$  zunehmen, muss  $K(N) = K_0 N$ , mit einer Konstante  $K_0$ . Wir kommen auf dasselbe, wenn wir verlangen, dass die totale Federkraft der Kette gleich bleibt (dort müssen wir beachten, dass nur die letzte Feder effektiv eine Kraft auf den Endpunkt ausübt). Also ist

$$U = \frac{K_0}{2} N \sum_{i=1}^N |\Delta \vec{x}_i|^2 = \frac{K_0}{2} \sum_{i=1}^N \left| \frac{\Delta \vec{x}_i}{\Delta \lambda} \right|^2 \Delta \lambda,$$

wobei wir  $\Delta \lambda = 1/N$  eingeführt haben.

Nun können wir die Summe als Riemann-Summe interpretieren, und erhalten im Limes

$$U = \frac{K_0}{2} \int \left( \frac{d\vec{x}}{d\lambda} \right)^2 d\lambda = \frac{K_0}{2} \tilde{L}.$$

Die Energie des Gummibands ist also bis auf einen konstanten Faktor das Energie-Wirkungsintegral  $\tilde{L}$ ; deshalb der Name.

Wenn wir nun das Gummiband loslassen, wird es schnell seine endgültige Form annehmen; solch ein Prozess können wir also *adiabatisch* approximieren. Nun wissen wir aus der Thermodynamik, dass während einem adiabatischen Prozess die Entropie unverändert bleibt, und deshalb die *innere Energie* des Systems minimiert wird. Die innere Energie ist in unserem Fall einfach  $U$ .

Wir sehen also, dass das Energie-Wirkungsintegral  $\tilde{L}$  auch durch ein Gummiband zwischen den betrachteten Punkten motiviert werden kann, welches seine innere Energie und somit seine eigene Länge minimiert, wenn es losgelassen wird.

# 9 Mechanische Ähnlichkeit und drittes Keplersches Gesetz

Angenommen, ein physikalisches System bestehe aus  $N$  Punktteilchen mit kartesischen Koordinaten  $\vec{x}_i$ ,  $i = 0, \dots, N$ , und werde durch eine zeitunabhängige Lagrange-Funktion beschrieben:

$$L = T - U, \quad T = \sum_{i=1}^N \frac{1}{2} m_i \left( \frac{d\vec{x}_i}{dt} \right)^2, \quad U = U(\vec{x}_1, \dots, \vec{x}_N).$$

Wir nehmen zudem an, dass  $U$  eine spezielle Form hat, nämlich

$$U(\alpha\vec{x}_1, \dots, \alpha\vec{x}_N) = \alpha^k \cdot U(\vec{x}_1, \dots, \vec{x}_N), \quad \text{für eine Zahl } k \text{ und alle Zahlen } \alpha > 0.$$

- (a) Erklären Sie, dass die Lagrange-Funktion  $L' := \lambda \cdot L$ ,  $\lambda \neq 0$ , genau dieselben Lösungen für die Bewegung des Systems liefert, wie  $L$ . Wir sagen,  $L'$  und  $L$  sind *äquivalente Lagrange-Funktionen*.
- (b) Wir betrachten nun die Koordinatentransformation  $\vec{x}'_i = \alpha^{-1} \vec{x}_i$  für  $\alpha > 0$ , kombiniert mit einer Skalierung der Zeit:  $t' = \beta^{-1} t$  für  $\beta > 0$ .

Zeigen Sie, dass die Lagrange-Funktion  $L'(\vec{x}'_1, \dots, d\vec{x}'_1/dt', \dots)$  ausgedrückt in den neuen Koordinaten und der neuen Zeit, äquivalent ist zur Lagrange-Funktion  $L(\vec{x}_1, \dots, d\vec{x}_1/dt, \dots)$ , ausgedrückt in den alten Koordinaten und der alten Zeit <sup>1</sup>, wenn  $\alpha^2 \beta^{-2} = \alpha^k$ , sprich  $\beta = \alpha^{1-k/2}$ .

*Hinweis: Geschwindigkeiten skalieren bei der Transformation um  $\alpha^{-1}/\beta^{-1}$ .*

- (c) Angenommen, wir haben eine mögliche Lösung der Bewegungsgleichungen gefunden. Wir skalieren die Dimensionen der Bahnkurven der Teilchen dann um  $\alpha \neq 0$  und erhalten neue Bahnkurven, die die Bewegungsgleichungen aber nicht mehr unbedingt lösen.

Argumentieren Sie, dass die neuen Bahnkurven wieder Lösungen der Bewegungsgleichungen sind, wenn wir sie zudem um den Faktor  $\beta$  langsamer abspielen, wobei  $\beta$  in (c) durch  $\alpha$  und  $k$  gegeben ist. Die alte und neue Bahn heißen zueinander *mechanisch ähnlich*.

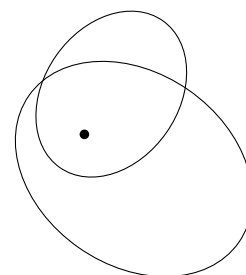
*Hinweis: Wenn Sie die Bahnen skalieren, können Sie die Koordinaten umgekehrt skalieren, sodass die neuen Bahnen in den neuen Koordinaten numerisch gleich aussehen wie die alte Bahn in den alten Koordinaten.*

- (d) Betrachten Sie den Fall eines Teilchens im Gravitationspotential  $U(\vec{x}) = -GM/|\vec{x}|$ . Für dieses Potential ist  $k = -1$ .

Zeigen Sie, dass für zwei geometrisch ähnliche Ellipsenbahnen  $A$  und  $B$  (siehe Abbildung rechts) für die Perioden  $T$  und grosse Halbachsen  $a$  gilt, dass

$$\frac{T_A}{T_B} = \left( \frac{a_A}{a_B} \right)^{3/2} \quad (\text{schwache Version des 3. Kepler-Gesetzes})$$

*Hinweis: Verwenden Sie neben (c) auch, dass dieses Potential rotationssymmetrisch ist.*



Bahnen, die durch Drehung und Skalierung ineinander übergehen, heißen *geometrisch ähnlich*.

<sup>1</sup>Genauer gesagt:

$$L'(\vec{x}'_1, \dots, d\vec{x}'_1/dt', \dots) := L(\vec{x}_1(\vec{x}'_1), \dots, d\vec{x}_1/dt(d\vec{x}'_1/dt'), \dots) \stackrel{!}{=} \lambda \cdot L(\vec{x}_1, \dots, d\vec{x}_1/dt, \dots)$$

muss erfüllt sein für ein  $\lambda \neq 0$ .

## Lösung

- (a) Die Euler-Lagrange-Gleichungen für  $L'$  sind

$$\frac{\partial L'}{\partial \vec{x}_i} = \frac{d}{dt} \frac{\partial L'}{\partial \dot{\vec{x}}_i} \quad \Rightarrow \quad \lambda \frac{\partial L}{\partial \vec{x}_i} = \lambda \frac{d}{dt} \frac{\partial L}{\partial \dot{\vec{x}}_i}.$$

Den Faktor  $\lambda$  können wir wegen Linearität der Ableitung nach aussen ziehen. Da  $\lambda > 0$ , können wir auf beiden Seiten durch  $\lambda$  teilen. So erhalten wir die Euler-Lagrange-Gleichungen für  $L$ .

- (b) Unter Verwendung des Hinweises ist

$$T \left( \frac{d\vec{x}'_1}{dt'}, \dots \right) := T \left( \frac{d\vec{x}_1}{dt}, \dots \right) = \sum_{i=1}^N \frac{m_i}{2} \left( \frac{d\vec{x}_1}{dt} \right)^2 = \sum_{i=1}^N \frac{m_i}{2} \frac{\alpha^2}{\beta^2} \left( \frac{d\vec{x}'_i}{dt'} \right)^2.$$

Zudem ist

$$U'(\vec{x}'_1, \dots) := U(\vec{x}_1, \dots) = U(\alpha \vec{x}'_1, \dots) = \alpha^k U(\vec{x}'_1, \dots).$$

Unter Beobachtung der Fussnote in der Aufgabenstellung ist also  $L'$  äquivalent zu  $L$ , wenn  $\alpha^k = \alpha^2/\beta^2$ , sprich  $\beta = \alpha^{1-k/2}$ .

- (c) Seien  $\vec{x}_i(t)$  die ursprünglichen Bahnen (also Lösungen der Euler-Lagrange-Gleichungen). Die skalierten Bahnen sind  $\vec{x}'_i(t) := \alpha \vec{x}_i(t)$ . Die skalierten Bahnen unterscheiden sich physikalisch von den ursprünglichen und sind nicht unbedingt Lösungen der Euler-Lagrange-Gleichungen. Die Idee ist jetzt, dass wir die skalierten Bahnen mit den Erkenntnissen aus (b) noch weiter anpassen können, indem wir sie auch in der Zeit skalieren, sodass sie wieder Lösungen der Bewegungsgleichungen werden. Wir spielen die Bahnen also zusätzlich um  $\beta$  langsamer ab:  $\vec{x}'_i(t') := \alpha \vec{x}_i(t'/\beta)$ , wobei  $\beta$  wie in (b) durch  $\alpha$  und  $k$  bestimmt ist.

Wir können nun noch eine Koordinaten- und Zeittransformation vornehmen: wenn wir die Koordinaten mit  $\alpha^{-1}$  skalieren, und die Zeit mit  $\beta^{-1}$ , dann haben die neuen, verlangsamten Bahnen in den neuen Koordinaten/Zeit dieselbe funktionelle Form, wie die alten Bahnen in den alten Koordinaten/Zeit.

Nur gehorchen die alten Bahnen in den alten Koordinaten/Zeit der Lagrange-Funktion  $L$ , während die neue Bahnen in neuen Koordinaten/Zeit der Lagrange-Funktion  $L'$  gehorchen müsste, um eine physikalische Lösung zu sein.

Wir haben in (b) gesehen, dass  $L$  und  $L'$  in den jeweiligen Koordinaten/Zeiten äquivalent sind, wenn die Koordinaten- und Zeittransformation gemäss  $\beta = \alpha^{1-k/2}$  aufeinander abgestimmt sind. Aber dies ist hier gerade der Fall und da die alten Bahnen Lösungen waren, sind es die neuen auch.

- (d) Aus  $k = -1$  folgt  $\beta = \alpha^{3/2}$ . Wir beginnen mit einer Ellipsenbahn  $A$  und skalieren sie um den Faktor  $\alpha$ , um die Bahn  $B$  zu erhalten. Gleichzeitig muss die Bahn  $B$  um den Faktor  $\beta$  langsamer verlaufen als  $A$ , damit sie eine echte Bahn ist. Sprich, die Periode von  $B$  ist um  $\beta$  grösser als die Periode von  $A$ . Wenn  $a_A$  und  $a_B$  die grossen Halbachsen der Bahnen sind, dann ist  $\alpha = a_B/a_A$ . Es folgt

$$\frac{T_B}{T_A} = \beta = \alpha^{3/2} = \left( \frac{a_B}{a_A} \right)^{3/2}.$$

Wenn wir davon den Kehrwert nehmen, erhalten wir gerade die Gleichung in der Aufgabenstellung.

In einem zweiten Schritt können wir  $B$  nun noch beliebig um das Massenzentrum drehen. Da das Potential rotationssymmetrisch ist, muss die so erhaltene Bahn immer noch eine Lösung der Bewegungsgleichungen sein, mit genau derselben Periode.

## Weiterführende Bemerkungen und Literatur

Die Aktivität basiert auf Kapitel §10 von [9]. Dort werden auch weitere Anwendungen von mechanischer Ähnlichkeit behandelt. So kann zum Beispiel elegant gezeigt werden, dass die Periode eines harmonischen Oszillators nicht von der Amplitude abhängt.

In der Übungsstunde wurde eine andere Version der Aktivität verwendet; da diese Version nicht sehr klar war, wurde die Aktivität nachträglich angepasst.

# 10 Relativistischer Drehimpuls aus dem Noether-Theorem

Wir betrachten ein relativistisches Teilchen mit Masse  $m$  und Eigenzeit  $\tau$ . Seine Bewegung ist z.B. gegeben durch das folgende Wirkungsintegral:

$$S = \int_{\tau_1}^{\tau_2} L d\tau, \quad L = \frac{m}{2} \eta_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau}.$$

Wir betrachten den Fluss  $\phi_\lambda : x^\mu \mapsto \Lambda(\lambda)^\mu{}_\nu x^\nu$ , mit  $\Lambda(\lambda) \in L_+^\uparrow$  eine Lorentz-Matrix.

- (a) Zeigen Sie, dass  $L\left(\frac{d}{d\tau}\phi_\lambda x^\mu\right) = L\left(\frac{dx^\mu}{d\tau}\right)$ , also dass  $\phi_\lambda$  eine Symmetrie ist.

*Hinweis: Lorentz-Transformationen verändern die quadrierte Länge eines 4er-Vektors nicht.*

- (b) Wir schreiben  $A^\mu{}_\nu := \frac{\partial}{\partial \lambda} \Lambda(\lambda)^\mu{}_\nu \Big|_{\lambda=0}$ . Zeigen Sie mit dem Noether-Theorem, dass

$$M = m \eta_{\mu\nu} \frac{dx^\mu}{d\tau} A^\nu{}_\rho x^\rho = \eta_{\mu\nu} p^\mu A^\nu{}_\rho x^\rho$$

erhalten ist, wobei  $p^\mu = m \frac{dx^\mu}{d\tau}$  der 4er-Impuls des Teilchens ist.

*Hinweis: Das Theorem besagt, dass  $\frac{\partial L}{\partial(dx^\nu/d\tau)} \frac{\partial}{\partial \lambda} \phi_\lambda x^\nu \Big|_{\lambda=0}$  erhalten ist.*

- (c) Folgern Sie aus  $\Lambda(\lambda) \in L_+^\uparrow = \left\{ \Lambda \in \text{GL}(4) \mid \Lambda^T \eta \Lambda = \eta, \det \Lambda = 1, \Lambda^0{}_0 \geq 1 \right\}$  und  $\Lambda(0) = \text{id}$ , dass

$$A^T \eta + \eta A = 0, \quad \text{tr } A = 0, \quad A^0{}_0 = 0 \quad \text{bzw.} \quad A = \begin{pmatrix} 0 & a^1 & a^2 & a^3 \\ a^1 & 0 & -b^3 & b^2 \\ a^2 & b^3 & 0 & -b^1 \\ a^3 & -b^2 & b^1 & 0 \end{pmatrix}, \quad \text{für beliebige } a^i, b^j \in \mathbb{R}$$

*Hinweis: Verwenden Sie  $\det(id + \lambda A + \mathcal{O}(\lambda^2)) = 1 + \lambda \text{tr } A + \mathcal{O}(\lambda^2)$ .*

- (d) Folgern Sie aus (b) und (c), dass der Tensor  $M^{\mu\nu} = p^\mu x^\nu - p^\nu x^\mu$  erhalten ist. Diskutieren Sie, welche spezifischen Symmetrien  $\phi_\lambda$  zur Erhaltung der einzelnen Komponenten führen. Erklären Sie, wieso es sinnvoll ist,  $M^{\mu\nu}$  als Drehimpuls des Teilchens zu sehen.

*Hinweis: Betrachten Sie z.B. den Boost in x-Richtung, d.h.*

$$\Lambda(\lambda) = \begin{pmatrix} \cosh \lambda & \sinh \lambda & 0 & 0 \\ \sinh \lambda & \cosh \lambda & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

*oder die Rotation um die z-Achse, d.h.*

$$\Lambda(\lambda) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos \lambda & -\sin \lambda \\ 0 & 0 & \sin \lambda & \cos \lambda \end{pmatrix}$$

- (e) Können wir dies auf  $N$  Teilchen erweitern? Was bräuchten wir noch in der Lagrange-Funktion, damit die Drehimpulserhaltung interessant wird?



## Lösung

- (a) Wir haben  $\frac{d}{d\tau}\phi_\lambda x^\mu = \frac{d}{d\tau}[\Lambda(\lambda)^\mu{}_\nu x^\nu] = \Lambda(\lambda)^\mu{}_\nu \frac{d}{d\tau}x^\nu$ . Dies sehen wir auch bereits daran, dass  $\frac{d}{d\tau}x^\mu$  die Komponenten der 4er-Geschwindigkeit sind, einem 4er-Vektor.

Also ist

$$L\left(\frac{d}{d\tau}\phi_\lambda x^\mu\right) = \frac{m}{2}\eta_{\mu\nu}\Lambda(\lambda)^\mu{}_\rho \frac{dx^\rho}{d\tau}\Lambda(\lambda)^\nu{}_\sigma \frac{dx^\sigma}{d\tau} = \frac{m}{2}\eta_{\rho\sigma} \frac{dx^\rho}{d\tau} \frac{dx^\sigma}{d\tau} = L\left(\frac{d}{d\tau}x^\mu\right).$$

Das zweite Gleichheitszeichen folgt, weil die quadrierte Länge der 4er-Geschwindigkeit Lorentz-invariant ist. Somit haben wir gezeigt, dass  $\phi_\lambda$  eine Symmetrie ist, und zwar für alle  $\Lambda(\lambda)$ , die  $\phi_\lambda$  zu einem Fluss machen (per Definition muss eine Symmetrie nämlich ein Fluss sein). Insbesondere muss  $\Lambda(0) = \text{id}$  die Identitätsmatrix sein.

- (b) Das erzeugende Vektorfeld des Flusses ist

$$v^\mu = \frac{\partial}{\partial\lambda}\phi_\lambda x^\mu \Big|_{\lambda=0} = \frac{\partial}{\partial\lambda}\Lambda(\lambda)^\mu{}_\nu x^\nu \Big|_{\lambda=0} = A^\mu{}_\nu x^\nu,$$

mit der Notation in der Aufgabenstellung.

*Bemerkung:* der Tensor  $A^\mu{}_\nu$  wird auch der Generator der Einparameter-Gruppe  $\Lambda(\lambda)$  von Lorentz-Transformationen genannt.

Noether's Theorem besagt nun, dass

$$M = \frac{\partial L}{\partial\left(\frac{dx^\mu}{d\tau}\right)}v^\mu = m\eta_{\mu\nu} \frac{dx^\mu}{d\tau} A^\nu{}_\rho x^\rho$$

erhalten ist. Mit dem 4er-Impuls  $p^\mu = m \frac{dx^\mu}{d\tau}$  folgt

$$M = \eta_{\mu\nu} p^\mu A^\nu{}_\rho x^\rho.$$

- (c) Angenommen,  $\Lambda(\lambda) \in L_+^\uparrow$ . Wenn wir die Bedingung  $\Lambda^T \eta \Lambda = \eta$  nach  $\lambda$  ableiten und bei  $\lambda = 0$  auswerten, erhalten wir

$$\begin{aligned} 0 &= \frac{\partial}{\partial\lambda}\Lambda(\lambda)^T \Big|_{\lambda=0} \eta \Lambda(0) + \Lambda(0)^T \eta \frac{\partial}{\partial\lambda}\Lambda(\lambda) \Big|_{\lambda=0}, \\ 0 &= A^T \eta + \eta A, \\ A &= -\eta A^T \eta, \end{aligned}$$

eine Bedingung an  $A$ . Ausgeschrieben in Indexnotation ist sie

$$A^\mu{}_\nu = -\eta^{\mu\rho} A^\sigma{}_\rho \eta_{\sigma\nu}. \quad (10.1)$$

Weil  $\det(\Lambda(\lambda)) = 1$ , muss insbesondere für  $\lambda \ll 1$  gemäss Hinweis

$$1 = \det(\text{id} + \lambda A + \mathcal{O}(\lambda^2)) = 1 + \lambda \text{tr} A + \mathcal{O}(\lambda^2)$$

gelten, also  $\text{tr} A = 0$ . Die Bedingung  $\Lambda(\lambda)^\mu{}_\nu \geq 1$  bedeutet  $1 + \lambda A^0{}_0 + \mathcal{O}(\lambda^2) \geq 1$ . Also muss  $A^0{}_0 = 0$ , weil  $\lambda$  positiv und negativ sein kann.

Wir betrachten den Diagonaleintrag  $A^\mu{}_\mu$  (für den Rest der Teilaufgabe keine Einsteinsche Summenkonvention) und erhalten mit (10.1)

$$A^\mu{}_\mu = -\eta^{\mu\mu} A^\mu{}_\mu \eta_{\mu\mu} = -A^\mu{}_\mu = 0.$$

Hier haben wir verwendet, dass  $\eta_{\mu\mu} = \eta^{\mu\mu} = \pm 1$ . Somit sind alle Diagonaleinträge von  $A$  null. Dies erfüllt gleich auch  $\text{tr} A = 0$  und  $A^0{}_0 = 0$ . Analog erhalten wir für  $i = 1, 2, 3$

$$A^0{}_i = -\eta^{00} A^i{}_0 \eta_{ii} = -(1)A^i{}_0(-1) = A^i{}_0.$$

Und schliesslich für  $i, j = 1, 2, 3$

$$A^i{}_j = -\eta^{ii} A^j{}_i \eta_{jj} = -(-1)A^j{}_i(-1) = -A^j{}_i.$$

Somit muss  $A$  von folgender Form sein:

$$A = \begin{pmatrix} 0 & a^1 & a^2 & a^3 \\ a^1 & 0 & -b^3 & b^2 \\ a^2 & b^3 & 0 & -b^1 \\ a^3 & -b^2 & b^1 & 0 \end{pmatrix},$$

wobei  $a^i, b^j \in \mathbb{R}$ . Tatsächlich kann für jede Wahl der  $a^i$ 's und  $b^j$ 's ein passendes  $\Lambda(\lambda)$  gefunden werden. Die Vorzeichen der  $b^j$ 's sind nur Konvention; wichtig ist aber, dass der  $3 \times 3$ -Block von  $A$  unten rechts antisymmetrisch ist, während der Rest symmetrisch ist.

(d) Wir folgen dem Tipp und betrachten zuerst den Boost in  $x^1$ -Richtung:

$$\Lambda(\lambda) = \begin{pmatrix} \cosh \lambda & \sinh \lambda & 0 & 0 \\ \sinh \lambda & \cosh \lambda & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \text{also } A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

(wir bemerken, dass  $A$  tatsächlich von der in (c) geforderten Form ist). Die Erhaltungsgrösse ist nun

$$M = \eta_{\mu\nu} p^\mu A^\nu_\rho x^\rho = \eta_{\mu 0} p^\mu A^0_1 x^1 + \eta_{\mu 1} p^\mu A^1_0 x^0 = p^0 x^1 - p^1 x^0.$$

Also ist

$$x^1(x^0) = x^1_0 + \frac{p^1}{p^0} x^0 = x^1_0 + v^1 x^0,$$

wobei  $x^1_0$  eine Konstante und  $v^1 = dx^1/dx^0$  die  $x^1$ -Geschwindigkeitskomponente ist. Wenn wir berücksichtigen, dass für das freie Teilchen der  $x^1$ -Impuls  $p^1$  und die Energie  $p^0$  auch erhalten sind, dann besagt die letzte Gleichung gerade, dass sich das Teilchen entlang der  $x^1$ -Achse mit konstanter Geschwindigkeit bewegt. Dies entspricht der  $x^1$ -Komponente des *Schwerpunktsintegrals*, welches wir aus der klassischen Physik als Folge von Symmetrie unter Galilei-Boosts kennen. Statt von einem Galilei-Boost kommt die Erhaltungsgrösse in der speziellen Relativitätstheorie von einem relativistischen Boost.

Betrachten wir nun die Rotation um die  $x^1$ -Achse,

$$\Lambda(\lambda) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos \lambda & -\sin \lambda \\ 0 & 0 & \sin \lambda & \cos \lambda \end{pmatrix}, \quad \text{also } A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

zur Erhaltungsgrösse

$$M = \eta_{\mu 2} p^\mu A^2_3 x^3 + \eta_{\mu 3} p^\mu A^3_2 x^2 = -p^3 x^2 + p^2 x^3 = -(x^2 p^3 - x^3 p^2).$$

führt. Das ist die  $x^1$ -Komponente des Vektors  $-\vec{x} \times \vec{p}$ , wobei  $\vec{x} = (x^1, x^2, z^2)^T$  und  $\vec{p} = (p^1, p^2, p^3)^T$ . Es macht also durchaus Sinn, dies (bis auf das Vorzeichen) die  $x^1$ -Komponente des relativistischen Drehimpuls zu nennen.

Wenn wir statt in  $x^1$ -Richtung in die anderen beiden Richtung boosten, erhalten wir analog  $p^0 x^2 - p^2 x^0 = \text{konst.}$  und  $p^0 x^3 - p^3 x^0 = \text{konst.}$  Wenn wir um die anderen beiden Achsen drehen, erhalten wir analog  $x^1 p^2 - x^2 p^1 = \text{konst.}$  und  $x^3 p^1 - x^1 p^3 = \text{konst.}$  Wir haben so gezeigt, dass die verschiedenen Komponenten des antisymmetrischen Tensors

$$M^{\mu\nu} = p^\mu x^\nu - p^\nu x^\mu$$

alle erhalten sind.

Die Komponenten  $M^{32} = -M^{23}$ ,  $M^{13} = -M^{31}$  und  $M^{21} = -M^{12}$  sind Verallgemeinerungen der klassischen Drehimpulskomponenten  $L_1$ ,  $L_2$  und  $L_3$ . Dennoch nennen wir den ganzen Tensor  $M^{\mu\nu}$ , inklusive der drei Schwerpunktsintegrale  $M^{01} = -M^{10}$ ,  $M^{02} = -M^{20}$ ,  $M^{03} = -M^{30}$ , den *relativistischen Drehimpuls* des Teilchens.

- (e) Wenn wir verschiedene freie Teilchen mit Massen  $m_i$  betrachten, dann sind die Drehimpulse  $M_i^{\mu\nu}$  der Teilchen einzeln erhalten, also auch der gesamte Drehimpuls. Diese Situation ist aber nicht sonderlich spannend. Erst wenn wir eine Wechselwirkung zwischen den Teilchen betrachten, würde eine Gesamtdrehimpulserhaltung potentiell interessant. Dazu müssen wir alle Teilchen aber durch eine einzige Lagrange-Funktion beschreiben. Das ist schwierig, weil verschiedene Teilchen im Allgemeinen verschiedene Eigenzeiten haben. Wir müssten statt  $\tau$  einen anderen Lorentz-invarianten Parameter  $s$  verwenden. Es bietet sich unmittelbar kein Parameter an, der auch eine physikalische Bedeutung besitzt; dies erschwert die Interpretation der Erhaltungsgrösse.

## Weiterführende Bemerkungen und Literatur

**Lie-Algebra der Lorentz-Gruppe** Alle möglichen  $A$ 's in (c) bilden einen Vektorraum der Dimension 6. Intuitiv ist dies der Vektorraum der „Geschwindigkeiten“, mit denen wir uns von der Identität  $\text{id} \in L_+^\uparrow$  innerhalb von  $L_+^\uparrow$  entfernen können.

Genauer gesagt ist  $L_+^\uparrow$  eine *Lie-Gruppe* der Dimension 6; das heisst eine Gruppe, die auch eine Mannigfaltigkeit der Dimension 6 mit stetiger Gruppenoperation sowie stetiger Inverse-Abbildung ist. Dann ist der Vektorraum der  $A$ 's der *Tangententialraum*  $T_{\text{id}}L_+^\uparrow$  der Mannigfaltigkeit  $L_+^\uparrow$  an die Identität  $\text{id} \in L_+^\uparrow$ . Dieser Tangentialraum wird zusammen mit dem *Kommutator*  $[A, A'] := AA' - A'A$  auch die *Lie-Algebra* von  $L_+^\uparrow$  genannt.

Die Lie-Algebra von  $L_+^\uparrow$  enthält übrigens die Lie-Algebra von  $SO(3)$  (bis auf Darstellung):  $T_{\text{id}}SO(3)$  sind nämlich die antisymmetrischen  $3 \times 3$ -Matrizen. Dies macht intuitiv Sinn, da auch  $SO(3)$  gewissermassen in  $L_+^\uparrow$  enthalten ist.

Lie-Gruppen und Lie-Algebren sind ein Teilgebiet der Differentialgeometrie; für eine detaillierte Behandlung, siehe z.B. [7].

**Klassische Feldtheorien** Das Problem von (e) kann mittels *klassischen Feldtheorien* gelöst werden („klassisch“ bedeutet hier soviel wie „nicht quantenmechanisch“, die Feldtheorien können sehr wohl relativistisch sein).

Wenn wir instantane Fernwirkungen zwischen Teilchen verhindern wollen (da sich Information gemäss spezieller Relativitätstheorie höchstens mit Lichtgeschwindigkeit ausbreitet), ist die Modellierung von Interaktionen zwischen Teilchen gar nicht so einfach. Wir müssen ein Feld einführen, welches die Wirkungen der Teilchen aufeinander mit maximal Lichtgeschwindigkeit vermittelt. Wenn wir dies tun, müssen wir das Feld unweigerlich auch als Teil unseres physikalischen Systems sehen und in die Lagrange-Funktion aufnehmen. Die Mathematik der Variationsrechnung wird dann etwas allgemeiner und die Euler-Lagrange-Gleichungen werden nicht nur die Bewegungsgleichungen der Teilchen beinhalten, sondern auch die *Feldgleichungen* (z.B. die Maxwell-Gleichungen für das elektromagnetische Feld).

Das Noether-Theorem kann auch auf Situationen mit Feldern verallgemeinert werden. Besitzt das System Translationssymmetrie in Raum und Zeit, dann ist der sogenannte *Energie-Impuls-Tensor*  $T^{\mu\nu}$ , eine Verallgemeinerung des gesamten 4er-Impulses eines Systems aus einzelnen Teilchen auf den Fall von kontinuierlichen Feldern, erhalten. Besitzt das System zudem eine Symmetrie unter allen eigentlich orthochronen Lorentz-Transformationen ( $L_+^\uparrow$ ), dann folgt die Erhaltung eines Drehimpulstensors  $J^{\mu\nu\rho}$  dritten Rangs, der sich aus  $x^\mu$  und  $T^{\mu\nu}$  konstruieren lässt.

Für eine Einführung in klassische Feldtheorien, sie z.B. Kapitel 13 von [3]. Für eine rigorose Behandlung der wichtigsten beiden Feldtheorien in der Physik, der Elektrodynamik und der allgemeinen Relativitätstheorie, siehe [10].

# 11 \* Noether-Theorem im Hamilton-Formalismus

Seien  $z = (q^1, \dots, p_1, \dots)$  die Phasenraumkoordinaten eines Systems mit zeitunabhängiger Hamilton-Funktion  $H$ . Sei  $\phi^t$  die *Lösungsabbildung*, die dem Anfangszustand  $z_0 = z(0)$  den Zustand  $z(t)$  des Systems zur Zeit  $t$  zuweist:  $\phi^t(z_0) := z(t)$ .  $\phi^t$  ist ein kanonischer Fluss und wird durch die Hamilton-Funktion erzeugt:

$$\{z, H(z)\} = \frac{d}{dt}z = \left. \frac{d}{dt}\phi^t(z) \right|_{t=0}.$$

Statt  $H$  können wir aber auch irgendeine zeitunabhängige Phasenraumfunktion  $F : z \mapsto F(z)$  betrachten.  $F$  kann als Erzeugende eines Flusses  $\psi^\lambda$  gesehen werden:

$$\left. \frac{d}{d\lambda}\psi^\lambda(z) \right|_{\lambda=0} := \{z, F(z)\}.$$

- (a) Zeigen Sie, dass  $\psi^\lambda$  ein Fluss ist. Es lässt sich zudem zeigen (Bonus: zeigen Sie dies), dass der Fluss kanonisch ist. Erklären Sie die Parallelen zwischen  $\psi^\lambda$  und  $\phi^t$ .

*Hinweis: Denken Sie ans Anfangswertproblem einer gewöhnlichen Differentialgleichung erster Ordnung.*

- (b) Zeigen Sie, dass

$$\left. \frac{d}{dt}F(\phi^t(z)) \right|_{t=0} = - \left. \frac{d}{d\lambda}H(\psi^\lambda(z)) \right|_{\lambda=0}. \quad (11.1)$$

*Hinweis: Verwenden Sie, dass  $dF(z(t))/dt = \{F, H\}(z(t))$  und die Antisymmetrie der Poisson-Klammer.*

- (c) Erklären Sie, dass die linke Seite von (11.1) null ist, genau dann, wenn  $F$  eine Erhaltungsgröße ist. Erklären Sie zudem, wieso die rechte Seite dann als Symmetrieeigenschaft gesehen werden kann. Begründen Sie also, wieso (11.1) das Noether-Theorem im Hamilton-Formalismus ist.

*Hinweis: Dass  $\psi^\lambda$  kanonisch ist, wird für die Symmetrieeigenschaft benötigt.*

- (d) Wenden Sie das Theorem auf ein Beispiel an:  $H = \vec{p}^2/2m$  (freies Teilchen),  $F = \vec{p} \cdot \vec{e}$  für  $|\vec{e}| = 1$ . Bestimmen Sie die Symmetrie (Fluss), unter der  $F$  erhalten ist. Vergleichen Sie das Resultat mit dem Wissen aus der Lagrange-Mechanik.

Mit dieser Version des Noether-Theorems können wir keine Erhaltungsgrößen behandeln, die explizit von der Zeit abhängen (z.B. Schwerpunktsintegrale). Wir nehmen deshalb an, dass  $F$  und  $H$  jetzt auch explizit von der Zeit abhängen können.

- (e) Erklären Sie, dass das Noether-Theorem dann folgendermassen aussieht:

$$\left. \frac{d}{ds}F(\phi^s(z), s+t) \right|_{s=0} = - \left. \frac{d}{d\lambda}H(\psi^\lambda(z), t) \right|_{\lambda=0} + \frac{\partial}{\partial t}F(z, t), \quad (11.2)$$

Was sind die physikalischen Bedeutungen der verschiedenen Terme? Bonus: Erklären Sie, wieso die rechte Seite eine sinnvolle Symmetrieeigenschaft ist.

- (f) Wenden Sie das allgemeine Noether-Theorem (11.2) an:  $H = \sum_{k=1}^N \vec{p}_k^2/2m_k$  ( $N$  freie Teilchen). Zeigen Sie, dass die drei Schwerpunktsintegrale

$$\frac{\vec{P}}{M}t - \vec{X}, \quad \text{wobei} \quad \vec{P} := \sum_{k=1}^N \vec{p}_k, \quad M := \sum_{k=1}^N m_k, \quad \vec{X} := \sum_{k=1}^N m_k \vec{x}_k,$$

erhalten sind und die entsprechenden Symmetrien die Galilei-Boosts sind.

## Lösung

Die Lösung wird möglicherweise später noch hinzugefügt. Bis aufs weitere ist sie der Leserin und dem Leser als Aufgabe überlassen.

## Weiterführende Bemerkungen und Literatur

Diese Aktivität wurde in der Übungsstunde nicht behandelt.

Die Version des Noether-Theorems, die wir in der Lagrange-Mechanik verwenden, wird in einführenden Büchern und Vorlesungen über klassische Mechanik meistens der hier behandelten Hamilton-Version vorgezogen, obwohl die Hamilton-Version vergleichbar mächtig ist. Dies kann verschiedene Gründe haben:

- (i) Lagrange-Mechanik ist zumindest anfangs einfacher zu verstehen als Hamilton-Mechanik, da die Hamilton-Mechanik schon früh fortgeschrittene Konzepte aus der Differentialgeometrie benötigt, wie z.B. *Differentialformen* und später *symplektische Mannigfaltigkeiten*.
- (ii) Die Lagrange-Version lässt sich besser auf relativistische Feldtheorien erweitern, weil das Wirkungsintegral eines Feldes bereits Lorentz-invariant ist. Hingegen können in der Hamiltonschen Feldtheorie nicht so einfach Lorentz-invariante Skalare gefunden werden.
- (iii) Die Lagrange-Version ist ähnlich mächtig wie die Hamilton-Version. Mit dem allgemeinen Lagrange-Noether-Theorem sind also die allermeisten Fälle abgedeckt.

Aber natürlich gehört die Hamiltonsche Mechanik zu einem guten Verständnis der klassischen Mechanik dazu. Fortgeschrittene und mathematisch rigorose Lehrbücher zur klassischen Mechanik sind z.B. [16] und [13].

# Literaturverzeichnis

- [1] C. W. Misner, K. S. Thorne, J. A. Wheeler. *Gravitation*. W. H. Freeman and Company, San Francisco, 1973.
- [2] E. F. Taylor, J. A. Wheeler. *Spacetime Physics: Introduction to Special Relativity*. 2nd Edition. W. H. Freeman and Company, New York, 1991.
- [3] H. Goldstein, C. P. Poole, J. L. Safko. *Classical Mechanics*. 3rd Edition. Addison Wesley, 2000.
- [4] J. D. Bjorken, S. D. Drell. *Relativistic Quantum Fields*. McGraw-Hill, 1965.
- [5] H. Amann und J. Escher. *Analysis III*. 2. Auflage. Birkhäuser, Basel, 2008.
- [6] J. M. Lee. *Introduction to Riemannian Manifolds*. 2nd Edition. Springer, New York, 2018.
- [7] J. M. Lee. *Introduction to Smooth Manifolds*. 2nd Edition. Springer, New York, 2013.
- [8] K. Jänich. *Topologie*. 7. Auflage. Springer, Heidelberg, 2001.
- [9] L. D. Landau, E. M. Lifschitz. *Lehrbuch der Theoretischen Physik, Band 1: Mechanik*. 14. Ausgabe. Europa Lehrmittel, Haan-Gruiten, 1997.
- [10] L. D. Landau, E. M. Lifschitz. *Lehrbuch der Theoretischen Physik, Band 2: Klassische Feldtheorie*. 12. Ausgabe. Europa Lehrmittel, Haan-Gruiten, 1992.
- [11] L. D. Landau, E. M. Lifschitz. *Lehrbuch der Theoretischen Physik, Band 4: Quantenelektrodynamik*. 7. Ausgabe. Europa Lehrmittel, Haan-Gruiten, 1991.
- [12] M. Tenenbaum. *Ordinary Differential Equations: An Elementary Textbook for Students of Mathematics, Engineering, and the Sciences*. Dover Publications, New York, 1985.
- [13] R. Abraham, J. E. Marsden. *Foundations of Mechanics*. 2nd edition. Addison-Wesley, 1987.
- [14] R. Feynman, R. Leighton, M. Sands. *The Feynman Lectures on Physics, Volume I: Mainly Mechanics, Radiation and Heat*. New Millennium Edition. Basic Books, 2011.
- [15] R. Feynman, R. Leighton, M. Sands. *The Feynman Lectures on Physics, Volume II: Mainly Electromagnetism and Matter*. New Millennium Edition. Basic Books, 2011.
- [16] V. I. Arnold. *Mathematical Methods of Classical Mechanics*. 2nd edition. Graduate Texts in Mathematics. Springer, 1989.
- [17] W. Rindler. *Introduction to Special Relativity*. Clarendon Press, Oxford, 1982.