

ALLGEMEINE MECHANIK

Herbstsemester 2013

Renato Renner

Institut für Theoretische Physik

ETH-Hönggerberg

CH-8093 Zürich

renner@phys.ethz.ch

Version vom 01. Juli 2014

Inhaltsverzeichnis

1	Newton'sche Mechanik	6
1.1	Raum und Zeit	6
1.2	Inertialsysteme	6
1.3	Der Schwerpunktsatz und das Konzept der Masse	7
1.4	Das Newton'sche Prinzip der Determiniertheit	8
1.5	Das Galileische Relativitätsprinzip	9
1.6	Erhaltungssätze	11
1.7	Beschleunigte Bezugssysteme	13
2	Zweikörper-Systeme	17
2.1	Der allgemeine Fall	17
2.1.1	Relativkoordinaten und Erhaltungsgrößen	17
2.1.2	Die verschiedenen Bahntypen	19
2.1.3	Gebundene Bahnen	20
2.1.4	Streubahnen	21
2.2	Das Keplerproblem	23
2.2.1	Ellipsenbahnen	24
2.2.2	Hyperbelbahnen	26
2.2.3	Runge-Lenz Vektor	27
3	Ausgewählte Dreikörperprobleme	29
3.1	Gleichgewichtslagen und ihre Stabilität	29
3.2	Die Bewegung des Mondes	32
3.2.1	Periodische Bahnen	37
3.2.2	Variationen der Bahn	39
4	Schwingungsprobleme	41
4.1	Lösungen der homogenen Gleichung	41
4.2	Erzwungene Schwingungen	46
4.3	Parametrische Resonanz	48
4.4	Stabilisierung linearer Systeme	52
5	Lagrange Formulierung	56
5.1	Das Brachistochronenproblem	57
5.2	Die Euler-Lagrange Gleichungen	60
5.3	Beispiele zur Variationsrechnung	61
5.3.1	Kürzeste Verbindungslinie zweier Punkte	61
5.3.2	Die Kettenlinie und Lagrange Multiplikatoren	62
5.4	Das Hamiltonsche Prinzip	65
5.4.1	Holonome Zwangsbedingungen	66
5.4.2	Äquivalente Lagrangefunktionen	71
5.5	Zyklische Koordinaten und konjugierte Impulse	72

5.6	Noether's Theorem	75
5.7	Erhaltungssätze revisited	79
5.8	Das Prinzip von Maupertuis	81
6	Hamiltonsche Systeme	86
6.1	Die Legendretransformation	86
6.2	Die Hamiltonschen Gleichungen	88
6.3	Phasenraum und Poissonklammern	89
6.4	Kanonische Transformationen	92
6.5	Symplektische Geometrie	95
6.6	Kanonische Flüsse	98
6.6.1	Der Satz von Liouville	100
6.6.2	Der Wiederkehrsatze von Poincaré	100
6.7	Erhaltungsgrößen	102
6.8	Das Hamiltonsche Prinzip im Phasenraum	104
6.8.1	Erzeugung kanonischer Transformationen	104
7	Die Hamilton-Jacobi Gleichung	108
7.1	Der zeitunabhängige Fall	108
7.2	Separable Probleme	109
7.2.1	Das ebene Zentralkraftproblem	109
7.2.2	Das ebene Zweizentren-Problem	110
7.3	Zeit-abhängige Hamilton-Jacobi Gleichung	112
8	Starre Körper	113
8.1	Eulerwinkel	113
8.2	Der Trägheitstensor eines Kreisels	114
8.3	Der freie Kreisel	115
8.3.1	Permanente Rotationen	117
8.3.2	Der symmetrische freie Kreisel	118
8.4	Der schwere symmetrische Kreisel	119
8.4.1	Erhaltungssätze	120
8.4.2	Integration der Bewegungsgleichungen	121
8.4.3	Der schnelle Kreisel	123
8.4.4	Reine Präzession und andere Grenzfälle	126
9	Die spezielle Relativitätstheorie	128
9.1	Das Gesetz der Lichtausbreitung	128
9.2	Die Postulate von Einstein	129
9.3	Lorentzgruppe und Poincarégruppe	130
9.4	Eigenschaften der Lorentztransformationen	132
9.4.1	Kontravariante und kovariante Tensoren	137
9.5	Relativistische Mechanik	139
9.5.1	Zeitdilatation und Längenkontraktion	141

9.6 Lagrange Formulierung	143
A Anhang: Die Legendre-Transformation	147

Dieses Skript basiert grösstenteils auf den Vorlesungsskripten von Gian Michele Graf und Matthias Gaberdiel, denen ich für das Überlassen des \LaTeX Quelltextes danke.

1 Newton'sche Mechanik

Wir beginnen mit der Beschreibung der grundlegenden Prinzipien, die der Formulierung der klassischen Mechanik zugrunde liegen. Diese widerspiegeln die klassische Vorstellung von Raum und Zeit, die über Jahrhunderte als allgemeingültig angesehen worden ist, und wurden erst 1905 von Einstein umformuliert. (Die Relativitätstheorie, die aus dieser Umformulierung resultiert, werden wir später diskutieren.) Trotz dieser Entwicklung ist die klassische Mechanik auch heute noch von grosser Bedeutung, z.B. als Spezialfall der Relativitätstheorie im Grenzfall kleiner Geschwindigkeiten.

1.1 Raum und Zeit

Nach Festlegung von Längen- und Zeiteinheiten sowie eines Bezugssystems charakterisieren wir Ereignisse durch $(t, \mathbf{x}) \in \mathbb{R}^{1+3}$, wobei t die Zeitkoordinate, und $\mathbf{x} = (x^1, x^2, x^3)$ kartesische Koordinaten eines drei-dimensionalen euklidischen Raumes sind. Eine absolute, vom Bezugssystem unabhängige Bedeutung haben die Grössen

- $|t_1 - t_2|$: Zeitabstand von zwei beliebigen Ereignissen $(t_1, \mathbf{x}_1), (t_2, \mathbf{x}_2)$ (1.1.1)
(\implies Gleichzeitigkeit ist absolut);

- falls $t_1 = t_2$:
 $|\mathbf{x}_1 - \mathbf{x}_2|$: Raumabstand von zwei gleichzeitigen Ereignissen. (1.1.2)

Die Koordinatentransformationen, die diese Grössen invariant lassen, sind

$$\begin{aligned} t' &= \lambda t + a, & (\lambda = \pm 1, a \in \mathbb{R}), \\ \mathbf{x}' &= R(t)\mathbf{x} + \mathbf{b}(t), & (R(t) \in O(3), \mathbf{b}(t) \in \mathbb{R}^3). \end{aligned} \quad (1.1.3)$$

Hier ist $R(t) \in O(3)$ eine von t abhängige orthogonale 3×3 Matrix. Diese Transformationen erlauben es uns insbesondere, (i) die Zeitrichtung umzukehren ($\lambda = -1$); (ii) den Zeitnullpunkt zu verschieben ($a \neq 0$); und (iii) das räumliche Bezugssystem (in zeitabhängiger Weise) zu rotieren und verschieben ($R(t), \mathbf{b}(t)$).

Man kann leicht nachprüfen, dass die Transformationen (1.1.3) eine Gruppe bilden. Dies ist im wesentlichen eine Folge davon, dass sie durch die Invarianzen (1.1.1) und (1.1.2) charakterisiert sind.

1.2 Inertialsysteme

Vom Standpunkt der Mechanik sind jedoch nicht alle diese Koordinatensysteme äquivalent. Die besondere Klasse der *Inertialsysteme* ist durch das *Trägheitsgesetz* ausgezeichnet. In einem Inertialsystem bewegt sich ein *freies Teilchen*, d.h. ein Teilchen, das keinen (äusseren) Kräften ausgesetzt ist, geradlinig und gleichförmig (Trägheitsgesetz):

$$\ddot{\mathbf{x}} = 0, \quad (\cdot = \frac{d}{dt}). \quad (1.2.1)$$

Bemerkung:

Ein freies Teilchen ist umgekehrt dadurch definiert, dass es sich in einem Inertialsystem geradlinig und gleichförmig bewegt. Zum Beispiel ist ein Teilchen im Gravitationsfeld der Erde frei, wenn man ein frei fallendes Bezugssystem wählt.

Daher gilt

$$\mathbf{x}(t) = \mathbf{x}(0) + \dot{\mathbf{x}}(0) t. \quad (1.2.2)$$

In einem Inertialsystem ist also die Bahn eines freien Teilchens $t \mapsto (t, \mathbf{x}(t))$ eine Gerade im \mathbb{R}^4 . Die Untergruppe der Koordinatentransformationen (1.1.3), die Inertialsysteme in Inertialsysteme abbildet, sind dann die geradentreuen Abbildungen

$$\begin{aligned} t' &= \lambda t + a, & (\lambda = \pm 1, a \in \mathbb{R}), \\ \mathbf{x}' &= R\mathbf{x} + \mathbf{v}t + \mathbf{b}, & (R \in O(3), \mathbf{v}, \mathbf{b} \in \mathbb{R}^3). \end{aligned} \quad (1.2.3)$$

Diese Abbildungen bilden wiederum eine Gruppe, die *Galileigruppe* oder Gruppe der *Galileitransformationen*.

Wir postulieren die Existenz eines Inertialsystems (und somit unendlich vieler, gleichförmig gegeneinander bewegter Inertialsysteme).¹ In der Praxis werden unterschiedliche Bezugssysteme als Inertialsysteme verwendet. So kann man bei Labor-Experimenten auf kleinen Zeit- und Grössenskalen approximativ die Erde als Inertialsystem annehmen. Andererseits ist bei der Betrachtung von Planetenbahnen in unserem Sonnensystem dessen Schwerpunkt mit nach Fixsternen orientierten Achsen (mit hoher Präzision) ein Inertialsystem. Auf noch grösseren Zeit- und Längenskalen müsste man dann beispielsweise zum Schwerpunkt unserer Milchstrasse übergehen.

Soweit nichts anderes gesagt wird, bezeichnen im folgenden (t, \mathbf{x}) die Koordinaten eines Ereignisses in einem beliebigen Inertialsystem.

Das Galileische Relativitätsprinzip besagt, dass *alle Naturgesetze invariant sind unter Transformationen zwischen Inertialsystemen*. Die Naturgesetze sind also in allen Inertialsystemen identisch, und damit sind alle Inertialsysteme gleichberechtigt. Dieses fundamentale Prinzip wird natürlich später durch das Relativitätsprinzip der speziellen Relativitätstheorie ersetzt werden; für den Moment wollen wir jedoch damit arbeiten.

1.3 Der Schwerpunktsatz und das Konzept der Masse

Jedes Punktteilchen besitzt eine unveränderliche Masse $m > 0$, die (nach Festlegung einer Masseneinheit) durch den *Schwerpunktsatz* charakterisiert ist: Für ein isoliertes System von N Teilchen mit Koordinaten $\mathbf{x}_1, \dots, \mathbf{x}_N$ gilt

$$\sum_{i=1}^N m_i \ddot{\mathbf{x}}_i = \frac{d}{dt} \sum_{i=1}^N \mathbf{p}_i = 0. \quad (1.3.1)$$

¹Damit sind freie Teilchen definiert als diejenigen, die sich in einem Inertialsystem geradlinig und gleichförmig bewegen.

Die Definition der Masse m_i ist von der Wahl des Inertialsystems unabhängig, da $\ddot{\mathbf{x}}'_i = R\ddot{\mathbf{x}}_i$.

Die Grösse $\mathbf{p}_i = m_i\dot{\mathbf{x}}_i$ ist der Impuls des i -ten Teilchens. Die obige Gleichung ist dann der Erhaltungssatz für den Gesamtimpuls $\mathbf{P} = \sum_{i=1}^N \mathbf{p}_i$. Definieren wir den Schwerpunkt durch

$$\mathbf{X} = \frac{1}{M} \sum_{i=1}^N m_i \mathbf{x}_i, \quad M = \sum_{i=1}^N m_i, \quad (1.3.2)$$

dann lässt er sich auch als $\ddot{\mathbf{X}} = 0$ schreiben. Der Schwerpunkt eines isolierten Systems bewegt sich daher auf einer Geraden

$$\mathbf{X}(t) = \mathbf{X}(0) + \frac{\mathbf{P}}{M}t. \quad (1.3.3)$$

Bei Streuprozessen können Teilchen erzeugt oder vernichtet werden. Jedoch muss der Gesamtimpuls \mathbf{P} erhalten bleiben. Das Galileische Relativitätsprinzip impliziert dann, dass die Gesamtmasse konstant sein muss. Dazu betrachten wir ein zweites Inertialsystem \mathbf{x}' , das sich mit Geschwindigkeit \mathbf{v} relativ zu dem ersten Inertialsystem bewegt, $\mathbf{x}' = \mathbf{x} + \mathbf{v}t$. Der Impuls des i -ten Teilchens im gestrichelten Inertialsystem ist dann $\mathbf{p}'_i = \mathbf{p}_i + m_i\mathbf{v}$. Da sowohl \mathbf{P} als auch \mathbf{P}' konstant sind, folgt dann, dass auch M konstant sein muss. Die Masse ist daher additiv, falls sich mehrere Teilchen zu einem verbinden.

Die hier definierte Masse ist die *träge Masse* m_T . Im nächsten Kapitel wird die *schwere Masse* m_S als Proportionalitätsfaktor im Gravitationsgesetz eingeführt. Experimentell findet man, dass die beiden gleich sind

$$m_S = m_T, \quad \text{Experiment: } m_S/m_T = 1 \pm 10^{-12}. \quad (1.3.4)$$

1.4 Das Newton'sche Prinzip der Determiniertheit

Wir betrachten ein isoliertes System von N Teilchen, deren Ortskoordinaten durch $\mathbf{x}_i(t)$ beschrieben sind. Das Newton'sche Prinzip der Determiniertheit sagt dann, dass die Bahnkurven $\mathbf{x}_i(t)$ für alle t eindeutig festgelegt sind, falls die Orte $\mathbf{x}_i(t_0)$ und die Geschwindigkeiten $\dot{\mathbf{x}}_i(t_0)$ zu einem Zeitpunkt t_0 gegeben sind. Insbesondere sind dann also die Beschleunigungen $\ddot{\mathbf{x}}_i(t_0)$ bestimmt, d.h. es gibt Funktionen \mathbf{F}_i , so dass

$$m_i \ddot{\mathbf{x}}_i(t) = \underbrace{\mathbf{F}_i(\mathbf{x}_1(t), \dots, \mathbf{x}_N(t), \dot{\mathbf{x}}_1(t), \dots, \dot{\mathbf{x}}_N(t))}_{\text{Kraftgesetz des Systems}}. \quad (1.4.1)$$

Das sind die sogenannten Newtonsche Bewegungsgleichungen. Umgekehrt folgt aus dem Existenz- und Eindeutigkeitsatz für gewöhnliche Differentialgleichungen, dass (unter schwachen Voraussetzungen an die \mathbf{F}_i) die Bahnen $\mathbf{x}_i(t)$ durch (1.4.1) und die obigen Anfangsdaten zu $t = t_0$ zumindest für ein kleines Zeitintervall um t_0 eindeutig bestimmt sind.

Allgemeiner nennen wir ein System von N Teilchen ein *mechanisches System*, falls es einem Kraftgesetz der Form (1.4.1) genügt. Beispiele mechanischer Systeme sind das

Sonnensystem, dessen Kraftgesetz durch

$$m_i \ddot{\mathbf{x}}_i = -G \sum_{k \neq i} m_i m_k \frac{\mathbf{x}_i - \mathbf{x}_k}{|\mathbf{x}_i - \mathbf{x}_k|^3} \quad (1.4.2)$$

gegeben ist, oder ein System geladener Teilchen, für das das Kraftgesetz

$$m_i \ddot{\mathbf{x}}_i = \sum_{k \neq i} e_i e_k \frac{\mathbf{x}_i - \mathbf{x}_k}{|\mathbf{x}_i - \mathbf{x}_k|^3} \quad (1.4.3)$$

ist; beide dieser Systeme sind tatsächlich isolierte Systeme, und die Kraft \mathbf{F}_i auf das i -te Teilchen hängt nur von den Positionen der Teilchen ab,

$$\mathbf{F}_i \equiv \mathbf{F}_i(\mathbf{x}_1, \dots, \mathbf{x}_N). \quad (1.4.4)$$

Zudem ist \mathbf{F}_i einfach eine Superposition von 2-Körperkräften

$$\mathbf{F}_i = \sum_{k \neq i} \mathbf{F}_{ik}, \quad (1.4.5)$$

wobei $\mathbf{F}_{ik} = \mathbf{F}_{ik}(\mathbf{x}_i, \mathbf{x}_k)$ die vom Teilchen k auf das Teilchen i wirkende Kraft ist. Für das (isolierte) 2-Teilchensystem gilt dann wegen (1.3.1) einfach

$$\mathbf{F}_{ik} + \mathbf{F}_{ki} = 0 \quad (1.4.6)$$

(actio = reactio).

Beispiele nicht-isolierter mechanischer Systeme sind geladene Teilchen in einem externen elektromagnetischen Feld $\mathbf{E}(\mathbf{x}, t)$ und $\mathbf{B}(\mathbf{x}, t)$:

$$m \ddot{\mathbf{x}} = e \mathbf{E}(\mathbf{x}, t) + \frac{e}{c} \dot{\mathbf{x}} \wedge \mathbf{B}(\mathbf{x}, t), \quad (1.4.7)$$

oder ein 1-dimensionaler Oszillator mit Reibungskraft und Anregung

$$m \ddot{x} = -fx - r \dot{x} + k(t). \quad (1.4.8)$$

Die Beschreibung dieser Systeme als mechanische Systeme setzt häufig Näherungen voraus (selbst im Rahmen der klassischen Mechanik). Zum Beispiel sind in (1.4.7) $\mathbf{E}(\mathbf{x}, t)$, und $\mathbf{B}(\mathbf{x}, t)$ vorgegebene äussere elektromagnetische Felder, und der Einfluss des geladenen Teilchens auf sie wurde ignoriert. In (1.4.8) ist $-r \dot{x}$ eine summarische Beschreibung der Reibung, ohne Berücksichtigung der Dynamik des dämpfenden Mediums.

1.5 Das Galileische Relativitätsprinzip

Im Folgenden betrachten wir Systeme unter der Bedingung (1.4.4), d.h., die Kraft auf das i -te Teilchen ist unabhängig von den Geschwindigkeiten der Teilchen.

Das klassische (oder Galileische) Relativitätsprinzip verlangt, dass die Bewegungsgleichungen eines isolierten Systems in jedem Inertialsystem gleich lauten, also invariant unter Galilei-Transformationen (1.2.3) sind. Insbesondere muss dies für das Kraftgesetz \mathbf{F}_{ik} gelten. Wir wollen nun zeigen, dass dann \mathbf{F}_{ik} von der Form

$$\mathbf{F}_{ik} = f_{ik}(|\mathbf{x}_i - \mathbf{x}_k|) \frac{\mathbf{x}_i - \mathbf{x}_k}{|\mathbf{x}_i - \mathbf{x}_k|} \quad (1.5.1)$$

sein muss, wobei $f_{ik}(r)$ eine beliebige skalare Funktion einer Variablen ist. [Der Faktor $1/|\mathbf{x}_i - \mathbf{x}_k|$ hätte natürlich auch in die Definition von f_{ik} absorbiert werden können; für das folgende ist jedoch die obige Form bequemer.] Dies bedeutet, dass die Kraft entlang der Verbindungslinie der beiden Teilchen wirkt und in ihrer Stärke nur vom Abstand der beiden Teilchen abhängt.

Zum Beweis betrachten wir ein mechanisches 2-Teilchensystem, dessen Kraftgesetz in einem Inertialsystem von der Form

$$m_1 \frac{d^2 \mathbf{x}_1}{dt^2} = \mathbf{F}_{12}(\mathbf{x}_1, \mathbf{x}_2) \quad (1.5.2)$$

ist. Sei \mathbf{x}' ein anderes Inertialsystem, das durch eine Galileitransformation aus dem ungestrichenen Inertialsystem hervorgeht,

$$\mathbf{x}'_i(t') = R\mathbf{x}_i(\lambda(t' - a)) + \mathbf{v}\lambda(t' - a) + \mathbf{b}. \quad (1.5.3)$$

[Beachte, dass $t = \lambda(t' - a)$.] Das Relativitätsprinzip verlangt dann, dass in dem gestrichenen Inertialsystem dieselbe Bewegungsgleichung gilt, also dass

$$m_1 \frac{d^2 \mathbf{x}'_1}{dt'^2} = \mathbf{F}_{12}(\mathbf{x}'_1, \mathbf{x}'_2), \quad (1.5.4)$$

wobei \mathbf{F}_{12} dieselbe Funktion ist. Die Invarianz unter einer reinen Translation ($\mathbf{b} \neq \mathbf{0}$) impliziert sofort, dass \mathbf{F}_{12} nur von $\mathbf{x}_1 - \mathbf{x}_2$ abhängen kann, $\mathbf{F}_{12}(\mathbf{x}_1, \mathbf{x}_2) = \mathbf{F}_{12}(\mathbf{x}_1 - \mathbf{x}_2)$. Invarianz unter Rotationen $R \in \text{SO}(3)$ führt dann zu der Bedingung, dass

$$R\mathbf{F}_{12}(\mathbf{x}) = \mathbf{F}_{12}(R\mathbf{x}) \quad (1.5.5)$$

gilt. Wählt man speziell R mit Achse in Richtung $\mathbf{x} = \mathbf{x}_1 - \mathbf{x}_2$ der Verbindungsgeraden, d.h. $R\mathbf{x} = \mathbf{x}$, so folgt mit $R\mathbf{F}_{12}(\mathbf{x}) = \mathbf{F}_{12}(\mathbf{x})$ die Behauptung über die Richtung der Kraft:

$$\mathbf{F}_{12}(\mathbf{x}) = f_{12}(\mathbf{x}) \frac{\mathbf{x}}{|\mathbf{x}|}. \quad (1.5.6)$$

Aus (1.5.5) folgt nun $f_{12}(\mathbf{x}) = f_{12}(R\mathbf{x})$ für jede Drehung R , also $f_{12}(\mathbf{x}) = f_{12}(|\mathbf{x}|)$. Dies beweist die Behauptung.

Die Kräfte (1.5.1) besitzen stets ein Potential:

$$\mathbf{F}_{ik} = -\frac{\partial}{\partial \mathbf{x}_i} V_{ik}(|\mathbf{x}_i - \mathbf{x}_k|), \quad V'_{ik}(r) = -f_{ik}(r). \quad (1.5.7)$$

Dementsprechend gilt für die Superposition von 2-Körperkräften

$$\mathbf{F}_i = -\frac{\partial}{\partial \mathbf{x}_i} V(\mathbf{x}_1, \dots, \mathbf{x}_N), \quad V = \sum_{j < k} V_{jk}(|\mathbf{x}_j - \mathbf{x}_k|). \quad (1.5.8)$$

Etwas allgemeiner betrachten wir auch Kräfte der Form

$$\mathbf{F}_i = -\frac{\partial}{\partial \mathbf{x}_i} V(\mathbf{x}_1, \dots, \mathbf{x}_N), \quad (1.5.9)$$

d.h. wir verlangen nicht, dass sich die Kräfte als Summe von 2-Körperkräften schreiben lassen. Das Relativitätsprinzip verlangt dann, dass die beiden Bewegungsgleichungen

$$m_i \frac{d^2 \mathbf{x}_i}{dt^2} = -\frac{\partial}{\partial \mathbf{x}_i} V(\mathbf{x}_1, \dots, \mathbf{x}_N) \quad \text{und} \quad m_i \frac{d^2 \mathbf{x}'_i}{dt'^2} = -\frac{\partial}{\partial \mathbf{x}'_i} V(\mathbf{x}'_1, \dots, \mathbf{x}'_N) \quad (1.5.10)$$

zueinander äquivalent sind. Dies ist der Fall, falls das Potential unter den euklidischen Transformationen invariant ist,

$$V(R\mathbf{x}_1 + \mathbf{a}, \dots, R\mathbf{x}_N + \mathbf{a}) = V(\mathbf{x}_1, \dots, \mathbf{x}_N), \quad (R \in O(3), \mathbf{a} \in \mathbb{R}^3). \quad (1.5.11)$$

Umgekehrt folgt aus einer analogen Argumentation zu vorher, dass das Relativitätsprinzip impliziert, dass das Potential unter euklidischen Transformationen invariant ist.

Bis jetzt haben wir die ‘passive Interpretation’ des Relativitätsprinzips betrachtet: falls $\mathbf{x}(t)$ und $\mathbf{x}'(t')$ die *gleiche* Bahn in unterschiedlichen Inertialsystemen beschreiben, dann müssen sie in beiden Inertialsystemen die Bewegungsgleichungen erfüllen (die die gleiche Form haben müssen). Man kann jedoch auch die ‘aktive Interpretation’ benutzen: sei $\mathbf{x}(t)$ eine Lösung der Bewegungsgleichung in einem Inertialsystem, dann ist $\mathbf{x}(t)$ auch eine Lösung bezüglich jedes anderen Inertialsystems. [$\mathbf{x}(t)$ beschreibt dann im allgemeinen eine andere Bahn, da diese Funktion jetzt die Koordinaten bezüglich eines anderen Koordinatensystems bezeichnet!] Betrachte zum Beispiel den Fall der Zeitumkehr ($\lambda = -1$): falls $\mathbf{x}(t)$ eine Lösung ist, dann ist auch $\mathbf{x}(-t)$ eine Lösung. Die zeitumgekehrte Bewegung ist also immer auch eine Lösung der Bewegungsgleichungen!

1.6 Erhaltungssätze

Wir wollen nun einige allgemeine Aussagen über mechanische Systeme machen. Mechanische Systeme sind dadurch charakterisiert, dass sie ein Kraftgesetz von der Form (1.4.1) erfüllen.

(a) Impulssatz: Es folgt direkt aus (1.4.1), dass

$$\frac{d}{dt} \underbrace{\sum_i \mathbf{p}_i}_{\mathbf{P}} = \underbrace{\sum_i \mathbf{F}_i}_{\mathbf{F}: \text{resultierende Kraft}}. \quad (1.6.1)$$

(b) **Drehimpulssatz:** Da die Zeitableitung von \mathbf{x}_i proportional zu \mathbf{p}_i ist, gilt weiterhin

$$\frac{d}{dt} \underbrace{\sum_i \mathbf{x}_i \wedge \mathbf{p}_i}_{\mathbf{L}} = \underbrace{\sum_i \mathbf{x}_i \wedge \mathbf{F}_i}_{\mathbf{M}: \text{resultierendes Drehmoment bzgl. } \mathbf{x} = 0}. \quad (1.6.2)$$

Der Drehimpuls \mathbf{L} lässt sich zerlegen in Anteile der Schwerpunktsbewegung und der inneren Bewegung relativ zum Schwerpunkt,

$$\mathbf{L} = \mathbf{X} \wedge \mathbf{P} + \mathbf{L}_S, \quad \text{mit} \quad \mathbf{L}_S = \sum_i (\mathbf{x}_i - \mathbf{X}) \wedge (\mathbf{p}_i - m_i \dot{\mathbf{X}}), \quad (1.6.3)$$

wobei wir benützt haben, dass

$$\mathbf{X} \wedge \mathbf{P} = \mathbf{X} \wedge \sum_i \mathbf{p}_i = \sum_i m_i \mathbf{x}_i \wedge \dot{\mathbf{X}} = \mathbf{X} \wedge \sum_i m_i \dot{\mathbf{X}}, \quad (1.6.4)$$

da $\sum_i m_i \mathbf{x}_i = M \mathbf{X}$ und $M \dot{\mathbf{X}} = \mathbf{P}$. [Hier bezeichnet M die Gesamtmasse, $M = \sum_i m_i$.]

(c) **Energiesatz:** Multiplikation von (1.4.1) mit $\dot{\mathbf{x}}_i$ und Summation über i führt nun zu

$$\frac{d}{dt} \underbrace{\sum_i \frac{1}{2} m_i \dot{\mathbf{x}}_i^2}_T = \underbrace{\sum_i \mathbf{F}_i \cdot \dot{\mathbf{x}}_i}_{\text{Leistung der Kräfte}}. \quad (1.6.5)$$

Die *kinetische Energie* lässt sich wiederum zerlegen in

$$T = \frac{1}{2} M \dot{\mathbf{X}}^2 + T_S, \quad \text{wobei} \quad T_S = \sum_i \frac{1}{2} m_i (\dot{\mathbf{x}}_i - \dot{\mathbf{X}})^2. \quad (1.6.6)$$

Für Systeme, deren Kräfte aus einem Potential (1.5.9) herkommen, das unter euklidischen Transformationen invariant ist (siehe (1.5.11)), gilt weiterhin folgendes:

- (i) die resultierende Kraft $\mathbf{F} = 0$ verschwindet
- (ii) das resultierende Drehmoment $\mathbf{M} = 0$ verschwindet,
- (iii) die Leistung verringert das Potential,

$$\sum_i \mathbf{F}_i \cdot \dot{\mathbf{x}}_i = -\frac{dV}{dt}. \quad (1.6.7)$$

Um dies zu beweisen betrachten wir einen beliebigen Einheitsvektor \mathbf{e} . Dann gilt

$$\mathbf{e} \cdot \mathbf{F} = -\frac{d}{d\lambda} V(\mathbf{x}_1 + \lambda \mathbf{e}, \dots, \mathbf{x}_N + \lambda \mathbf{e}) \Big|_{\lambda=0} = 0, \quad (1.6.8)$$

da das Potential unter Translationen invariant ist. Da dies für beliebige Einheitsvektoren \mathbf{e} gilt, folgt damit (i). Um (ii) zu zeigen betrachten wir eine Drehung $R(\varphi)$ vom Winkel φ um \mathbf{e} . Dann gilt

$$\left. \frac{d}{d\varphi} R(\varphi) \mathbf{x} \right|_{\varphi=0} = \mathbf{e} \wedge \mathbf{x}, \quad (1.6.9)$$

und daher also

$$\mathbf{e} \cdot \mathbf{M} = \sum_i \mathbf{e} \cdot (\mathbf{x}_i \wedge \mathbf{F}_i) = \sum_i (\mathbf{e} \wedge \mathbf{x}_i) \cdot \mathbf{F}_i = - \left. \frac{d}{d\varphi} V(R(\varphi) \mathbf{x}_1, \dots, R(\varphi) \mathbf{x}_N) \right|_{\varphi=0} = 0. \quad (1.6.10)$$

Da \mathbf{e} beliebig ist, folgt daraus (ii). Der Beweis von (iii) folgt direkt aus der Kettenregel.

Ein solches System besitzt daher also die zehn klassischen Erhaltungsgrößen (Integrale der Bewegung)

$$\begin{array}{lll} \mathbf{P} \quad \text{und} \quad M \mathbf{X} - \mathbf{P}t & 6 \text{ Schwerpunktsintegrale} \\ \mathbf{L} \quad (\text{bzw } \mathbf{L}_S) & 3 \text{ Drehimpulsintegrale} \\ T + V \quad (\text{bzw } T_S + V) & \text{Energieintegral.} \end{array} \quad (1.6.11)$$

Diese 10 Erhaltungsgrößen stehen in Beziehung zu den 10 kontinuierlichen Parametern \mathbf{b} , \mathbf{v} , R und a (3+3+3+1) der Galilei-Gruppe (siehe Kapitel 5.6).

1.7 Beschleunigte Bezugssysteme

Die Bewegungsgleichungen haben, wie wir gesehen haben, eine einfache Form in allen Inertialsystemen, nämlich

$$m \ddot{\mathbf{x}} = \mathbf{F}. \quad (1.7.1)$$

Manchmal ist es jedoch bequem in Koordinatensystemen zu arbeiten, die keine Inertialsysteme sind. In solchen beschleunigten Bezugssystemen treten dann sogenannte ‘Scheinkräfte’ auf.

Bezeichne \mathbf{y} die Koordinaten eines beliebigen Bezugssystems, das durch

$$\mathbf{x} = R(t) \mathbf{y} + \mathbf{b}(t), \quad (R(t) \in SO(3) \quad \mathbf{b}(t) \in \mathbb{R}^3) \quad (1.7.2)$$

mit dem Inertialsystem \mathbf{x} in Beziehung steht. Dann gilt

$$\dot{\mathbf{x}} = \dot{R} \mathbf{y} + R \dot{\mathbf{y}} + \dot{\mathbf{b}}, \quad \ddot{\mathbf{x}} = \ddot{R} \mathbf{y} + 2\dot{R} \dot{\mathbf{y}} + R \ddot{\mathbf{y}} + \ddot{\mathbf{b}}, \quad (1.7.3)$$

also wegen $R^T R = 1$ (R^T ist die transponierte Matrix zu R)

$$m \ddot{\mathbf{y}} = R^T \mathbf{F} - 2m R^T \dot{R} \dot{\mathbf{y}} - m R^T \ddot{R} \mathbf{y} - m R^T \ddot{\mathbf{b}}. \quad (1.7.4)$$

Der erste Term auf der rechten Seite

$$\mathbf{K} \equiv R^T \mathbf{F} \quad (1.7.5)$$

ist der Kraftvektor in \mathbf{y} -Komponenten, und

$$\mathbf{a} \equiv R^T \ddot{\mathbf{b}} \quad (1.7.6)$$

ist die Beschleunigung des Punktes $\mathbf{y} = 0$ in \mathbf{y} -Komponenten. Weiterhin sieht man leicht, dass die Abbildung $\Omega = R^T \dot{R}$ antisymmetrisch ist, denn aus $R^T R = 1$ folgt

$$R^T \dot{R} + \dot{R}^T R = \Omega + \Omega^T = 0. \quad (1.7.7)$$

Somit hat Ω die Form

$$\Omega = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix} : \quad \Omega \mathbf{y} = \boldsymbol{\omega} \wedge \mathbf{y}, \quad (\boldsymbol{\omega} = (\omega_1, \omega_2, \omega_3)). \quad (1.7.8)$$

Ein im \mathbf{y} -System ruhender Punkt hat im \mathbf{x} -System die Geschwindigkeit $\dot{\mathbf{x}} = \dot{R} \mathbf{y} + \dot{\mathbf{b}}$. Die Komponenten dieser Geschwindigkeit im \mathbf{y} -System sind daher $R^T \dot{\mathbf{x}} = \boldsymbol{\omega} \wedge \mathbf{y} + R^T \dot{\mathbf{b}}$. Insbesondere sind daher $\boldsymbol{\omega}$ die \mathbf{y} -Komponenten der Winkelgeschwindigkeit des \mathbf{y} -Systems relativ zum \mathbf{x} -System.²

Weiter gilt

$$R^T \ddot{R} = \dot{\Omega} - \dot{R}^T \dot{R} = \dot{\Omega} - \dot{R}^T R R^T \dot{R} = \dot{\Omega} + \Omega^2, \quad (1.7.9)$$

$$R^T \ddot{R} \mathbf{y} = \dot{\boldsymbol{\omega}} \wedge \mathbf{y} + \boldsymbol{\omega} \wedge (\boldsymbol{\omega} \wedge \mathbf{y}). \quad (1.7.10)$$

Insgesamt lautet (1.7.4) nun also

$$m \ddot{\mathbf{y}} = \mathbf{K} - 2m (\boldsymbol{\omega} \wedge \dot{\mathbf{y}}) - m (\dot{\boldsymbol{\omega}} \wedge \mathbf{y}) - m \boldsymbol{\omega} \wedge (\boldsymbol{\omega} \wedge \mathbf{y}) - m \mathbf{a}. \quad (1.7.11)$$

Die rechts neben \mathbf{K} auftretenden Terme heissen Scheinkräfte, insbesondere ist

$$\begin{array}{ll} -2m \boldsymbol{\omega} \wedge \dot{\mathbf{y}} & \text{Coriolis-Kraft} \\ -m \boldsymbol{\omega} \wedge (\boldsymbol{\omega} \wedge \mathbf{y}) & \text{Zentrifugalkraft.} \end{array} \quad (1.7.12)$$

Im Gegensatz zur Zentrifugalkraft, welche auch ein im beschleunigten System ruhender Körper spührt, wirkt die Corioliskraft nur auf bewegte Körper. Letztere ist orthogonal zur Bewegungsrichtung und führt zum Beispiel zu Wirbeln in der Atmosphäre. Die Zentrifugal- und die Corioliskraft sind die einzigen Scheinkräfte bei gleichförmiger Rotation des \mathbf{y} -Systems ($\boldsymbol{\omega} = \text{konstant}$, $\mathbf{a} = 0$). Der Term $-m\mathbf{a}$, der für $\mathbf{a} \neq 0$ auftritt, wird Führungskraft genannt.

Es sei darauf hingewiesen, dass die obigen Scheinkräfte proportional zur *trägen Masse* m_T des Teilchens sind, wohingegen die Gravitationskraft $m_S \mathbf{g}$ zur *schweren Masse* m_S proportional ist. Das *Eötvös-Experiment* war einer der ersten Versuche, in dem gezeigt wurde, dass die beiden gleich sind. Dabei wurden zwei Testmassen aus unterschiedlichen Materialien und gleicher schwerer Masse $m_S^1 = m_S^2$ an einem Stab aufgehängt. Die auf

²Dies lässt sich sehen, indem man zwei Punkte \mathbf{y}_1 und \mathbf{y}_2 betrachtet.

die beiden Massen wirkende Zentrifugalkraft ist proportional zu ihren trägen Massen m_T^1 und m_T^2 . Falls letztere nicht gleich wären, würde dies zu einer beobachtbaren Drehung des Stabes führen. Dies lässt sich aber mit einer Genauigkeit

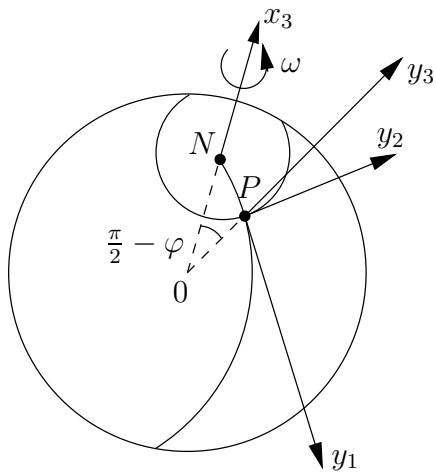
$$\eta = 2 \frac{|\mathbf{a}_1 - \mathbf{a}_2|}{|\mathbf{a}_1 + \mathbf{a}_2|} \approx 10^{-12}$$

widerlegen, woraus geschlossen werden kann, dass auch $m_T^1 = m_T^2$ ist. Damit folgt eigentlich erst

$$\frac{m_S^1}{m_T^1} = \frac{m_S^2}{m_T^2}.$$

Das Verhältnis kann aber gleich eins gesetzt werden, da ein eventueller Proportionalitätsfaktor in die Gravitationskonstante \mathbf{g} absorbiert werden kann.

Als Anwendung betrachten wir den *freien Fall auf der Erdoberfläche*. Sei P ein Punkt auf der Erdoberfläche, dessen geographische Breite durch φ beschrieben ist.



$(0, \mathbf{x})$ beschreibt ein ‘raumfestes’ Inertialsystem, dessen x_3 -Achse in Richtung des Nordpols N zeigt. (Wir vernachlässigen hier die Bewegung der Erde um die Sonne, Polschwankungen, usw.) Das rotierende Koordinatensystem wird durch \mathbf{y} bezeichnet. Der Vektor \mathbf{b} ist gleich $\mathbf{b} = R(t)\mathbf{p}$, wobei \mathbf{p} der ortsfeste Vektor vom Erdmittelpunkt zu P (zur Zeit $t = 0$) ist.

Dann lauten die Bewegungsgleichungen für einen Massenpunkt in der Nähe von P :

$$\ddot{\mathbf{y}} = \mathbf{g}' - 2\boldsymbol{\omega} \wedge \dot{\mathbf{y}} \quad (1.7.13)$$

wobei

$$\boldsymbol{\omega} = (-\omega \cos \varphi, 0, \omega \sin \varphi), \quad \omega = 2\pi/\text{Tag}.$$

und

$$\begin{aligned} \mathbf{g}' &= \mathbf{g} - \boldsymbol{\omega} \wedge (\boldsymbol{\omega} \wedge \mathbf{y}) - \mathbf{a} = \mathbf{g} - \boldsymbol{\omega} \wedge (\boldsymbol{\omega} \wedge (\mathbf{p} + \mathbf{y})) \\ &\approx (0, 0, -g), \quad (g \cong 9,81 \text{ m/s}^2) \end{aligned} \quad (1.7.14)$$

Dabei haben wir im ersten Schritt verwendet, dass $\mathbf{a} = R^T \ddot{\mathbf{b}}$ und $\mathbf{b} = R(t)\mathbf{p}$ gilt. Die Umformung ist dann analog zu (1.7.10). In der letzten Zeile haben wir die y^3 -Richtung als vertikale Richtung definiert, entlang derer die Schwerebeschleunigung bei P wirkt. Dabei haben wir angenommen, dass die Schwerebeschleunigung annäherungsweise überall auf der Erde gleich gross ist; der Effekt der Zentrifugalkraft kann daher vernachlässigt

werden. Wir haben ferner verwendet, dass $|\mathbf{y}| \ll |\mathbf{p}|$, so dass die Schwerebeschleunigung unabhängig von \mathbf{y} ist.

In Komponenten lautet jetzt Gleichung (1.7.13)

$$\begin{aligned}\ddot{y}_1 &= 2\omega \sin \varphi \dot{y}_2, \\ \ddot{y}_2 &= -2\omega \sin \varphi \dot{y}_1 - 2\omega \cos \varphi \dot{y}_3, \\ \ddot{y}_3 &= -g + 2\omega \cos \varphi \dot{y}_2.\end{aligned}\tag{1.7.15}$$

Die uns interessierenden Anfangsbedingungen zu $t = 0$ sind $y_1 = y_2 = 0$ und $y_3 = h > 0$, sowie $\dot{\mathbf{y}} = 0$. Die erste, und dritte Gleichung liefern dann

$$\dot{y}_1 = 2\omega \sin \varphi y_2, \quad \text{und} \quad \dot{y}_3 = -gt + 2\omega \cos \varphi y_2.\tag{1.7.16}$$

Einsetzen in die zweite Gleichung führt dann zu der Differentialgleichung für y_2 :

$$\ddot{y}_2 + 4\omega^2 y_2 = 2g t \omega \cos \varphi.\tag{1.7.17}$$

Die allgemeine Lösung davon ist

$$y_2 = \frac{g \cos \varphi}{2\omega} t + A \sin 2\omega t + B \cos 2\omega t.\tag{1.7.18}$$

Die Anfangsbedingung $y_2(0) = 0$ liefert $B = 0$, und $\dot{y}_2(0) = 0$ gibt $A = -(2\omega)^{-2} g \cos \varphi$, und daher finden wir

$$y_2 = \frac{g \cos \varphi}{2\omega} \left(t - \frac{1}{2\omega} \sin 2\omega t \right).\tag{1.7.19}$$

Für $t > 0$ ist $y_2 > 0$, d.h. es findet eine *Ostablenkung* statt. Da $\omega t \sim \text{Fallzeit/Tag} \ll 1$ entwickeln wir $\sin 2\omega t$ nach Potenzen von ωt

$$\sin 2\omega t \approx 2\omega t - \frac{1}{6}(2\omega t)^3$$

und finden

$$y_2 = \frac{gt^2}{3} \cos \varphi \cdot \omega t.\tag{1.7.20}$$

Mit Hilfe von (1.7.16) können wir nun natürlich auch y_1, y_3 bestimmen — der Effekt ist aber nur von der Ordnung $(\omega t)^2$. Der Aufprall findet zur Zeit $gt^2/2 = h$ statt, und daher erhalten wir

$$y_2 = \frac{1}{3}\omega \left(\frac{8h^3}{g} \right)^{1/2} \cos \varphi.\tag{1.7.21}$$

Zum Beispiel ist für $\varphi = 45^\circ, h = 100$ m die Ostablenkung gerade $y_2 \cong 1,6$ cm.

2 Zweikörper-Systeme

Bevor wir die Struktur der klassischen Mechanik weiter untersuchen wollen, ist es instruktiv zunächst ein paar Beispiele zu studieren. Wir beginnen mit dem einfachsten Fall, einem mechanischen System von zwei Teilchen.

2.1 Der allgemeine Fall

Wie wir in Kapitel 1.5 erklärt haben wirkt die Kraft entlang der Verbindungslinie der beiden Teilchen und hängt in ihrer Stärke nur vom Abstand der beiden Teilchen ab.³ Die Bewegungsgleichungen sind daher von der Form

$$\ddot{\mathbf{x}}_1 = -\frac{1}{m_1} \frac{\partial}{\partial \mathbf{x}_1} V(|\mathbf{x}_1 - \mathbf{x}_2|), \quad \ddot{\mathbf{x}}_2 = \frac{1}{m_2} \frac{\partial}{\partial \mathbf{x}_1} V(|\mathbf{x}_1 - \mathbf{x}_2|), \quad (2.1.1)$$

wobei $V(r)$ das Potential ist. Zum Beispiel ist für zwei Teilchen, die nur durch die Gravitationskraft wechselwirken $V(r) = -Gm_1m_2/r$. Dann gelten (2.1.1) auch für ausgedehnte, sphärisch symmetrische Körper (Himmelskörper, in guter Näherung), da ausserhalb eines solchen Körpers die Schwerkraft so ist, als ob sich die gesamte Masse im Mittelpunkt befände (Newtonscher Satz).⁴

2.1.1 Relativkoordinaten und Erhaltungsgrössen

Um das obige Problem zu lösen, ist es hilfreich, die Schwerpunktsbewegung zu separieren. Dazu definiert man, wie schon zuvor⁵

$$\mathbf{X} = \frac{1}{M}(m_1\mathbf{x}_1 + m_2\mathbf{x}_2), \quad (M = m_1 + m_2), \quad (2.1.2)$$

$$\mathbf{x} = \mathbf{x}_1 - \mathbf{x}_2. \quad (2.1.3)$$

Bezüglich dieser Variablen lauten dann die Bewegungsgleichungen

$$M\ddot{\mathbf{X}} = 0 \quad (2.1.4)$$

$$\mu\ddot{\mathbf{x}} = -\frac{\partial}{\partial \mathbf{x}} V(|\mathbf{x}|), \quad \left(\frac{1}{\mu} = \frac{1}{m_1} + \frac{1}{m_2}\right). \quad (2.1.5)$$

Die Schwerpunktsbewegung \mathbf{X} ist dann einfach eine Trägheitsbahn, und die Gleichung für die Relativbewegung $\mathbf{x}(t)$ ist die eines Teilchens mit reduzierter Masse μ unter dem Einfluss eines äusseren Potentials $V(|\mathbf{x}|)$. Dieses System besitzt nun zwei Erhaltungsgrössen:

³Wie im vorherigen Kapitel formal gezeigt wurde, ist dies eine Folge des Relativitätsprinzips. Die einzige Möglichkeit, dass das Kraftgesetz invariant unter Rotationen entlang der Verbindungslinie bleibt, ist, dass sie in Richtung dieser Achse zeigt.

⁴Auf ein Teilchen am Punkt r_0 innerhalb einer sphärisch symmetrischen Massenverteilung mit Ausdehnung R wirkt der Anteil der Masse mit $r < r_0$ so, als ob er im Mittelpunkt konzentriert wäre. Der Massenanteil mit $R > r > r_0$ übt keine resultierende Kraft auf das Teilchen aus.

⁵Gleichung (2.1.2) ist ein Konsistenz-Test: Dass die Bewegungsgleichungen in der Form (2.1.1) geschrieben werden können, wurde unter der Annahme hergeleitet, dass der Schwerpunktsatz gilt.

Drehimpuls: Der (relative) Drehimpuls

$$\mathbf{L} = \mu \mathbf{x} \wedge \dot{\mathbf{x}} \quad (2.1.6)$$

ist erhalten, da $\dot{\mathbf{L}} = -\mathbf{x} \wedge \partial_{\mathbf{x}} V(|\mathbf{x}|) = 0$.⁶ Somit liegt die Bahn $\mathbf{x}(t)$ in der Ebene $\perp \vec{L}$, der *Bahnebene*. In dieser benützen wir Polarkoordinaten (r, φ) mit zugehörigen Einheitsvektoren $\mathbf{e}_r = (\cos \varphi, \sin \varphi, 0)$ und $\mathbf{e}_\varphi = (-\sin \varphi, \cos \varphi, 0)$ in radialer bzw. azimuthaler Richtung. Mit $\dot{\mathbf{e}}_r = \dot{\varphi} \mathbf{e}_\varphi$, finden wir für

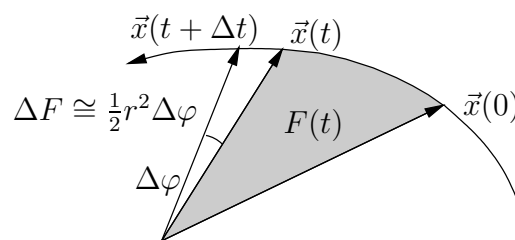
$$\mathbf{x} = r \mathbf{e}_r, \quad \dot{\mathbf{x}} = \dot{r} \mathbf{e}_r + r \dot{\varphi} \mathbf{e}_\varphi, \quad (2.1.7)$$

und daher ist der Betrag des Drehimpulses

$$l \equiv |\mathbf{L}| = \mu r^2 \dot{\varphi} = \text{konstant}. \quad (2.1.8)$$

Dies ist der Flächensatz (*2. Kepler-Gesetz*): Wenn $F(t)$ die vom Vektor $\mathbf{x}(t)$ in der Bahnebene überstrichene Fläche bezeichnet, so ist

$$\dot{F}(t) = \frac{1}{2} r^2 \dot{\varphi} = \frac{l}{2\mu} = \text{konstant}. \quad (2.1.9)$$



Der Flächensatz wurde von Kepler ursprünglich für Planetenbahnen formuliert. Er gilt aber allgemein für Systeme deren Drehimpuls erhalten ist und nicht nur für das Gravitationsgesetz.

Energie: Die andere Erhaltungsgröße ist die Relativenergie

$$T + V = \frac{\mu}{2} \dot{\mathbf{x}}^2 + V(r) = \frac{\mu}{2} (\dot{r}^2 + r^2 \dot{\varphi}^2) + V(r) = E = \text{konstant}. \quad (2.1.10)$$

[Die Konstanz von E folgt wie zuvor: $\dot{E} = \dot{\mathbf{x}} \cdot \mu \ddot{\mathbf{x}} + \dot{\mathbf{x}} \cdot \partial_{\mathbf{x}} V(|\mathbf{x}|) = 0$.] Einsetzen von $\dot{\varphi}$ aus (2.1.8) liefert

$$\frac{1}{2} \mu \dot{r}^2 + U(r) = E, \quad U(r) = \frac{l^2}{2\mu r^2} + V(r). \quad (2.1.11)$$

Die radiale Bewegung $r(t)$ ist die eines Massenpunktes μ auf $0 < r < \infty$ unter dem Einfluss des effektiven Potentials $U(r)$. Auflösen nach $\dot{r} = \frac{dr}{dt}$ führt zu

$$t(r) - t(r_0) = \pm \int_{r_0}^r \frac{dx}{\sqrt{\frac{2}{\mu}(E - U(x))}}. \quad (2.1.12)$$

⁶ $\partial_{\mathbf{x}} |\mathbf{x}| = \frac{\mathbf{x}}{|\mathbf{x}|}$

Weiterhin gilt

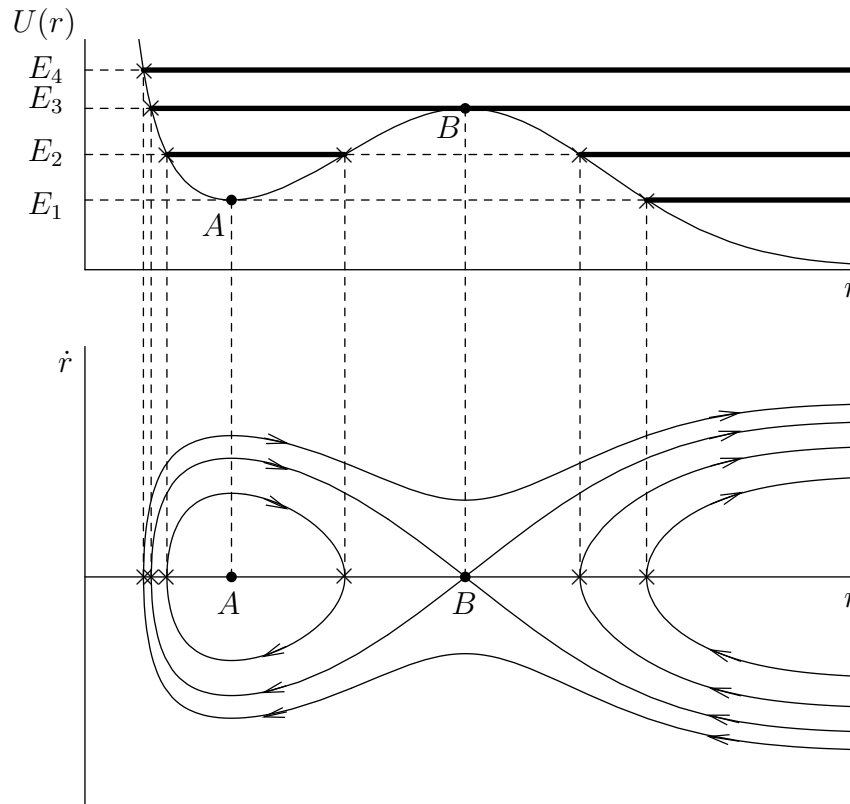
$$\frac{d\varphi}{dr} = \frac{\dot{\varphi}}{\dot{r}} = \frac{\pm l}{\mu r^2 \sqrt{\frac{2}{\mu}(E - U(r))}}, \quad (2.1.13)$$

und daher erhält man

$$\varphi(r) - \varphi(r_0) = \pm \int_{r_0}^r \frac{l dx}{x^2 \sqrt{2\mu(E - U(x))}}. \quad (2.1.14)$$

Dies beschreibt die Bahnkurve in der Ebene, d.h. die Bahn $(r(t), \varphi(t))$ unbeachtet ihres zeitlichen Ablaufs. Durch Auflösen von (2.1.12) findet man $r(t)$ und damit $\varphi(t)$. Durch (2.1.12) und (2.1.14) ist die Bestimmung der Bahn auf die Berechnung von Integralen reduziert: das Problem ist integrabel. Dies ist im Wesentlichen eine Folge davon, dass das Problem separierbar ist (d.h. unabhängige Bewegungsgleichungen können für r und φ hergeleitet werden. In der Praxis sind Probleme nur selten integrabel und müssen im Allgemeinen numerisch gelöst werden.

2.1.2 Die verschiedenen Bahntypen



Anhand eines Beispiels illustrieren wir mögliche Bewegungstypen in einer Dimension (die für die Relativbewegung wie oben erklärt relevant ist). Im Phasenraum mit den

Koordinaten (r, \dot{r}) ist die Bewegung für vier verschiedene Energien dargestellt (Phasenportrait). Sie verlaufen stets im Gebiet

$$E - U(r) = \frac{\mu}{2} \dot{r}^2 \geq 0. \quad (2.1.15)$$

An den Grenzen dieses Gebiets ist $\dot{r} = 0$; das sind die Umkehrpunkte der Bahn (im Diagramm mit \times bezeichnet). Damit $\dot{r} = 0$ ein Umkehrpunkt ist, muss weiterhin gelten

$$\mu \ddot{r} = -\frac{dU}{dr} \neq 0. \quad (2.1.16)$$

Falls $dU/dr = 0$ für $r = r_0$ (das sind die Punkte A und B im Diagramm) gibt es eine Gleichgewichtslösung $r = r_0$. Die Lösung bei A ist stabil (d.h. eine Bahn mit Anfangsbedingungen (r, \dot{r}) nahe bei A wird stets nahe bei A bleiben), wohingegen B instabil ist. Allgemein unterscheiden wir *gebundene Bahnen*, die ganz im Endlichen verlaufen, und *Streubahnen* mit $r(t) \rightarrow \infty$ für $t \rightarrow \pm\infty$.

2.1.3 Gebundene Bahnen

Für gebundene Bahnen ist $r(t)$ periodisch mit der Periode

$$T(E) = 2 \int_{r_{\min}}^{r_{\max}} \frac{dx}{\sqrt{\frac{2}{\mu}(E - U(x))}}, \quad (2.1.17)$$

wobei r_{\min} und r_{\max} die Nullstellen von $E - U(r)$ sind, die das erlaubte r -Intervall begrenzen. [Diese Formel folgt direkt aus (2.1.12).] Falls r_{\min} und r_{\max} Umkehrpunkte sind, so ist $\frac{dU}{dr} \Big|_{r_{\max}, r_{\min}} \neq 0$. Dann gilt, dass

$$T(E) = 2 \int \frac{1}{\sqrt{\frac{2}{\mu}(E - U)}} \underbrace{\frac{dr}{dU}}_{< \infty} dU < \infty. \quad (2.1.18)$$

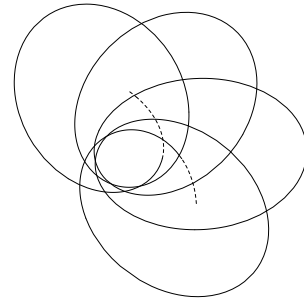
Das Integral konvergiert, da es sich wie $\sqrt{E - U}$ verhält und damit auch für $E \rightarrow U(r)$ endlich bleibt. Falls $\frac{dU}{dr} \Big|_{r_{\max}, r_{\min}} = 0$ ist, kann T divergieren oder endlich bleiben (abhängig davon wie schnell U flach wird).

Für eine periodische Bahn nimmt das Azimut φ während einer Periode T zu um den Winkel

$$\Delta\varphi = 2 \int_{r_{\min}}^{r_{\max}} \frac{l dx}{x^2 \sqrt{2\mu(E - U(x))}}, \quad (2.1.19)$$

was direkt aus (2.1.14) folgt.

So ergibt sich im Allgemeinen eine ‘Rosettenbahn’, im Ring $r_{\min} \leq r \leq r_{\max}$, die sich nur dann schliesst, wenn $\Delta\varphi/2\pi$ rational ist.



2.1.4 Streubahnen

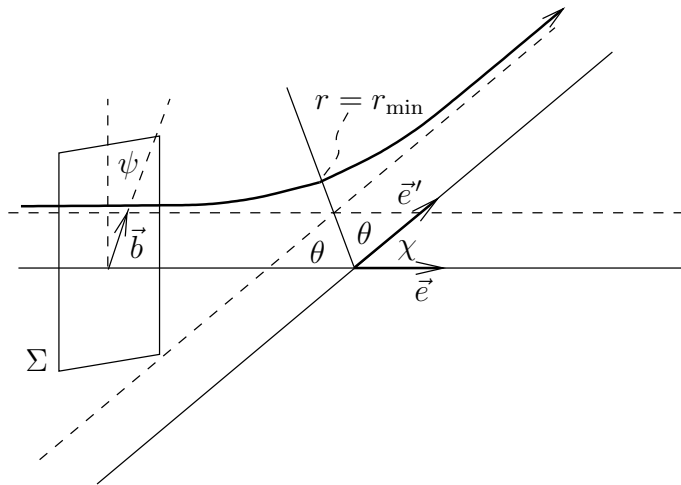
Wir betrachten nun den Fall, dass $V(r) \rightarrow 0$ für $r \rightarrow \infty$. [Im Limes $r \rightarrow \infty$ ist $V(r) - U(r) = 0$.] Dann gibt es Streubahnen ($r(t) \rightarrow \infty$ für $t \rightarrow \pm\infty$) nur für Energien $E \geq 0$. Falls $E > 0$ haben sie gerade Asymptoten für $t \rightarrow \pm\infty$. Die Streubahn ist bestimmt durch die Energie $E > 0$, die Richtung \mathbf{e} ($|\mathbf{e}| = 1$) der einlaufenden Asymptote, sowie den Stossparameter $\mathbf{b} \perp \mathbf{e}$. Für $t \rightarrow -\infty$ gilt dann wegen (2.1.11)

$$\mu|\dot{\mathbf{x}}| \rightarrow \sqrt{2\mu E}, \quad \text{also} \quad l = b\sqrt{2\mu E}, \quad (2.1.20)$$

wobei wir benützt haben, dass $l = \mu|\mathbf{x} \wedge \dot{\mathbf{x}}|$. Daher bestimmen $b = |\mathbf{b}|$ und E den *Streuwinkel* mit (2.1.14)

$$\chi = \pi - 2\theta, \quad \theta = \int_{r_{\min}}^{\infty} \frac{ldx}{x^2 \sqrt{2\mu(E - U(x))}} = \int_{r_{\min}}^{\infty} \frac{b dx}{x^2 \sqrt{1 - V(x)E^{-1} - b^2 x^{-2}}}. \quad (2.1.21)$$

Wobei r_{\min} bestimmt ist durch die Bedingung $U(r_{\min}) = E$.



Bei festen E und \mathbf{e} bestimmt \mathbf{b} den Richtungsvektor \mathbf{e}' ($|\mathbf{e}'| = 1$), der auslaufenden Asymptote. Dies definiert eine Abbildung $\Sigma \ni \mathbf{b} \mapsto \mathbf{e}' \in S^2$. Da \mathbf{e}' den Polarwinkel (χ, ψ) bezüglich \mathbf{e} hat, geht dabei das Flächenelement $d\sigma = b db d\psi$ von Σ über in das

Raumwinkelement $d\Omega = \sin \chi \cdot d\chi d\psi$. Der *differentielle Streuquerschnitt* (mit Dimension einer Fläche) ist dann definiert als

$$\frac{d\sigma}{d\Omega}(\chi) = \left| \frac{b}{\sin \chi} \frac{db}{d\chi} \right| = \frac{b}{\sin \chi} \left| \left(\frac{d\chi}{db} \right)^{-1} \right|, \quad (2.1.22)$$

wobei $\chi = \chi(b)$ durch (2.1.20) und (2.1.21) gegeben ist. Es kann vorkommen, dass verschiedene b_k zum gleichen Streuwinkel χ führen; dann ist (2.1.22) korrekt zu schreiben als

$$\frac{d\sigma}{d\Omega}(\chi) = \sum_k \frac{b_k}{\sin \chi} \left| \left(\frac{d\chi}{db_k} \right)^{-1} \right|. \quad (2.1.23)$$

Dieser Fall tritt zum Beispiel auf, falls $-\infty < V(0) \leq 0$. Dann ist nämlich $\chi(0) = 0$, denn für $b = 0$ verläuft die Bahnkurve gerade durch den Ursprung hindurch. Da auch $\lim_{b \rightarrow \infty} \chi(b) = 0$ gilt, folgt aus dem Zwischenwertsatz, dass jeder (nicht extreme) Wert $\chi(b_k)$ an mindestens zwei Stellen b_k angenommen wird.

In Streuexperimenten liegt eine homogene Stromdichte j (Teilchenzahl pro Flächen- und Zeiteinheit) einfallender Teilchen mit festen E und \mathbf{e} vor. Dann ist $j \cdot (d\sigma/d\Omega) d\Omega$ die Zahl der Stöße pro Zeiteinheit mit auslaufender Asymptotenrichtung $\mathbf{e}' \equiv (\chi, \psi)$ im Raumwinkelement $d\Omega$. Der *totale Streuquerschnitt*

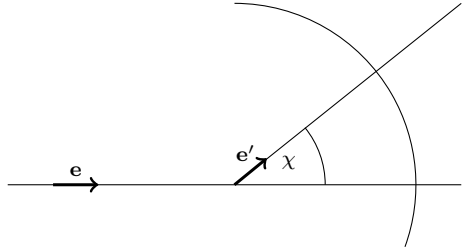
$$\sigma_{\text{tot}} = \int_{S^2 - \mathbf{e}} \frac{d\sigma}{d\Omega} d\Omega \quad (2.1.24)$$

ist gleich der Fläche der Stossparameter $\mathbf{b} \in \Sigma$, die zu einer Streuung führen, $\chi(b) \neq 0$, also

$$\sigma_{\text{tot}} = \pi a^2, \quad (2.1.25)$$

wobei $a = \sup\{r \mid V(r) \neq 0\} (\leq \infty)$ die Reichweite des Potentials ist. Die Integration ist über $S^2 - \mathbf{e}$ statt die ganze Oberfläche, da man diejenigen Teilchen, welche nicht gestreut wurden, nicht mitzählt.

Im Experiment ist $j \cdot (d\sigma/d\Omega) d\Omega$ die messbare Grösse, wobei j und $d\Omega$ (Grösse eines Detektorelements) bekannt sind. Damit lässt sich der differentielle Wirkungsquerschnitt ($d\sigma/d\Omega$) bestimmen. Mittels (2.1.23) und (2.1.21) kann dann eine Hypothese für ein Potential verifiziert werden. Der Stossparameter b lässt sich in der Praxis nicht kontrollieren. Zum Beispiel werden beim Rutherford-Experiment Teilchen durch eine Goldfolie bestreut, wobei jeder einzelne Au-Kern ein Streuzentrum darstellt.



Schematisch lässt sich dies in einer weiteren Figur mit einer anderen Grössenskala darstellen. Da der Stossparameter von der Grössenordnung von Atomabständen ist, hat er keinen direkten geometrischen Einfluss darauf, auf welches Detektorelement ein Teilchen trifft. Die Abhängigkeit ist rein physikalisch durch den Streuwinkel gegeben (welcher natürlich vom Stossparameter abhängt). In einer schematischen Darstellung welche den Detektor beinhaltet, ist die Bahn der aller einfallenden Teilchen identisch.

2.2 Das Keplerproblem

Als wichtigen Spezialfall betrachten wir die Bewegung eines Teilchens mit Masse m im Gravitationsfeld einer Punktmasse M_0

$$V(r) = -\frac{1}{r} G M_0 m, \quad (2.2.1)$$

wobei r der relative Abstand ist. Mit der reduzierten Masse

$$\mu = \frac{M_0 m}{m + M_0}, \quad M_0 m = M\mu, \quad (M = M_0 + m) \quad (2.2.2)$$

findet man dann die Bewegungsgleichung

$$\mu \ddot{\mathbf{x}} = -\frac{\mathbf{x}}{r^3} G M_0 m. \quad (2.2.3)$$

In Polarkoordinaten ist das

$$\dot{r} = \sqrt{\frac{2}{\mu}(E - U(r))}, \quad U(r) = \frac{l^2}{2\mu r^2} - \frac{G M_0 m}{r}, \quad \dot{\varphi} = \frac{l}{\mu r^2}. \quad (2.2.4)$$

Substitution von $r = 1/s$ führt dann zu

$$\begin{aligned} \frac{d\varphi}{ds} &= -\frac{1}{s^2} \frac{d\varphi}{dr} = -\frac{1}{s^2} \frac{\dot{\varphi}}{\dot{r}} \\ &= \pm \frac{l}{\mu} \frac{1}{\sqrt{\frac{2}{\mu}(E - U(r))}} \\ &= \pm (\alpha - s^2 + 2sd^{-1})^{-1/2}, \end{aligned} \quad (2.2.5)$$

wobei $\alpha = 2\mu E l^{-2}$,

$$d = \frac{l^2}{GM\mu^2}. \quad (2.2.6)$$

Dies integriert sich zu

$$\varphi(s) = \arccos\left(\frac{ds - 1}{\sqrt{d^2\alpha + 1}}\right), \quad (2.2.7)$$

wobei wir die freie Konstante gleich Null gesetzt haben. Durch Auflösung nach r ergibt sich die Bahnkurve in Polarkoordinaten

$$r = \frac{d}{1 + \varepsilon \cos \varphi}, \quad (2.2.8)$$

wobei

$$\varepsilon = \sqrt{d^2\alpha + 1}, \quad \text{d.h.} \quad 1 - \varepsilon^2 = -d^2\alpha = \frac{-2El^2}{G^2 M^2 \mu^3}. \quad (2.2.9)$$

Gl. (2.2.8) definiert einen Kegelschnitt mit einem Brennpunkt bei $r = 0$, und zwar je nach dem Wert der Exzentrizität ε eine

Ellipse	:	$\varepsilon < 1$,	d.h.	$E < 0$
Parabel	:	$\varepsilon = 1$,	d.h.	$E = 0$
Hyperbel	:	$\varepsilon > 1$,	d.h.	$E > 0$.

Der Winkel $\varphi = 0$ entspricht dem Perihel (r minimal).

2.2.1 Ellipsenbahnen

Aus der Figur und der Ellipsengleichung $r + \bar{r} = 2a$ ergibt sich

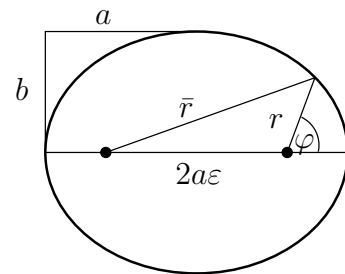
$$\bar{r}^2 = (r \cos \varphi + 2a\varepsilon)^2 + r^2 \sin^2 \varphi = (2a - r)^2$$

und damit, nach Auflösen nach r ,

$$4 a r(1 + \varepsilon \cos \varphi) = 4 a^2 (1 - \varepsilon^2) \quad (2.2.10)$$

$$\Leftrightarrow r = \frac{a(1 - \varepsilon^2)}{(1 + \varepsilon \cos \varphi)}. \quad (2.2.11)$$

Dies ist dann gerade (2.2.8) mit $d = a(1 - \varepsilon^2)$.



Damit haben wir das erste 1. Kepler-Gesetz hergeleitet:⁷

Planetenbahnen beschreiben Ellipsen mit Brennpunkt in der Sonne.

⁷Exakt würde dies nur gelten, falls die Sonne genau im Schwerpunkt wäre.

Johannes Kepler hat die aus den Beobachtungen von Tycho Brahe für den Mars hergeleitet ($\varepsilon = 0.0935$) und dies auf die anderen Planeten erweitert.

Weitere Merkmale der Ellipse sind die kleine Halbachse $b = a\sqrt{1 - \varepsilon^2}$ und die Ellipsenfläche $F = \pi ab = \pi d^{1/2} a^{3/2}$, für die nach dem Flächensatz (2. Kepler-Gesetz) gilt

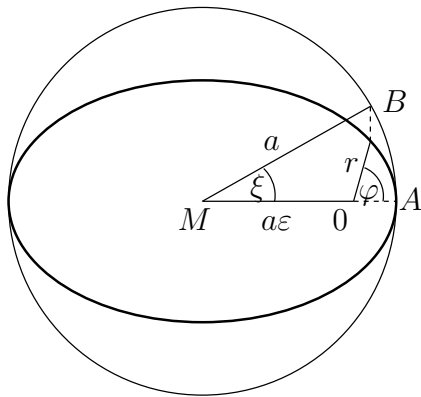
$$F = T\dot{F} = \frac{Tl}{2\mu}, \quad (2.2.12)$$

wobei T die Umlaufdauer ist. Daher ist

$$T = \frac{2\mu}{l} F = \frac{2\mu}{l} \pi d^{1/2} a^{3/2} = \frac{2\pi}{\sqrt{GM_0}} \sqrt{\frac{\mu}{m}} a^{3/2}. \quad (2.2.13)$$

Falls $m \ll M_0$ ist $\mu \cong m$, und das Verhältnis $T^2 : a^3$ ist gleich für alle Planeten (3. Kepler-Gesetz).

Zur Beschreibung des zeitlichen Verlaufs der Bahn parametrisieren wir sie durch die exzentrische Anomalie ξ (d.h. den Winkel relativ zum Mittelpunkt der Ellipse, s. Bild) statt dem Azimut φ . Durch Betrachtung der Projektion von MB auf MA findet man



$$\begin{aligned} \varepsilon \cdot a \cos \xi &= \varepsilon(a\varepsilon + r \cos \varphi) \\ &= a - a(1 - \varepsilon^2) + \varepsilon r \cos \varphi \\ &= a - r(1 + \varepsilon \cos \varphi) + \varepsilon r \cos \varphi \\ &= a - r, \end{aligned}$$

wobei in der dritten Gleichung (2.2.11) benützt wurde. Damit erhält man

$$r = a(1 - \varepsilon \cos \xi). \quad (2.2.14)$$

Mit Hilfe von (2.2.8) kann man (2.2.14) auch als

$$(1 - \varepsilon \cos \xi)(1 + \varepsilon \cos \varphi) = 1 - \varepsilon^2 \quad (2.2.15)$$

schreiben, wobei wir benützt haben, dass $d/a = 1 - \varepsilon^2$. Auflösen nach $\cos \varphi$ führt dann (unter Benützung von $\sin \varphi = \sqrt{1 - \cos^2 \varphi}$) zu

$$\sin \varphi = \frac{\sqrt{1 - \varepsilon^2}}{1 - \varepsilon \cos \xi} \cdot \sin \xi. \quad (2.2.16)$$

Gleichung (2.2.17) kann geschrieben werden als

$$1 + \varepsilon \cos \varphi = \frac{\sqrt{1 - \varepsilon^2}}{1 - \varepsilon \cos \xi} \quad (2.2.17)$$

Ableitung auf beiden Seiten nach ξ liefert

$$\frac{d\varphi}{d\xi} = \frac{\sin \xi}{\sin \varphi} \cdot \frac{1 - \varepsilon^2}{(1 - \varepsilon \cos \xi)^2} = \frac{1 + \varepsilon \cos \varphi}{\sqrt{1 - \varepsilon^2}} = \frac{1}{\sqrt{1 - \varepsilon^2}} \frac{d}{r}, \quad (2.2.18)$$

wobei wir (2.2.16) und (2.2.8) benützt haben. Der Flächensatz (2.1.9) besagt nun, dass

$$\begin{aligned} \frac{dt}{d\xi} &= \frac{d\varphi/d\xi}{d\varphi/dt} = \frac{\mu r^2}{l} \cdot \frac{1}{\sqrt{1 - \varepsilon^2}} \frac{d}{r} \\ &= \frac{1}{\sqrt{GM}} \sqrt{\frac{d}{1 - \varepsilon^2}} r = \frac{a^{3/2}}{\sqrt{GM}} (1 - \varepsilon \cos \xi), \\ t(\xi) - t_0 &= \sqrt{\frac{a^3}{GM}} (\xi - \varepsilon \sin \xi), \end{aligned} \quad (2.2.19)$$

wobei wir in der zweiten Zeile die Formel für d (2.2.6), die Relation $d = a(1 - \varepsilon^2)$ sowie (2.2.14) eingesetzt haben. Diese Gleichung drückt t als Funktion von ξ aus; $t = t_0$ entspricht hier einem Periheldurchgang ($\xi = 0$). [(2.2.14) und (2.2.19) sind die *Kepler Gleichungen*.]

2.2.2 Hyperbelbahnen

Nach (2.2.8) ist $r \rightarrow \infty$ für $\cos \varphi \rightarrow -1/\varepsilon$. Ausgedrückt durch den Winkel θ ist das die Bedingung $\cos \theta = 1/\varepsilon$ und daher $\tan^2 \theta = \varepsilon^2 - 1$. Andererseits ist nach (2.2.9, 2.1.20) aber auch

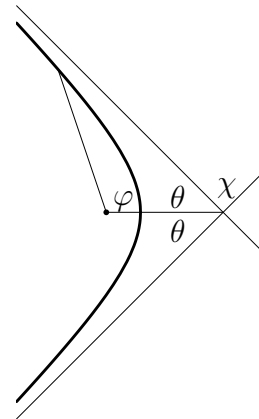
$$\varepsilon^2 - 1 = \frac{2E}{G^2 M_0^2 m^2 \mu} \cdot b^2 \cdot 2\mu E,$$

also

$$\tan \theta = \frac{2E}{GM_0 m} b = \frac{2E}{GM\mu} b,$$

wobei wir verwendet haben, dass

$$\tan^2 \theta = \frac{1 - \cos^2 \theta}{\cos^2 \theta}.$$



Wie im vorherigen Kapitel erwähnt, lässt sich der Stossparameter b in der Praxis nicht kontrollieren. Deshalb wollen wir aus obiger Formel den Wirkungsquerschnitt berechnen, welcher einen Zusammenhang zwischen experimentell überprüfbareren Grössen herstellt

(dem Streuwinkel und der Intensität der gestreuten Teilchen). Für die Berechnung des Wirkungsquerschnitts bemerken wir zuerst, dass aus obiger Formel folgt

$$\frac{d \tan \theta}{db} = \frac{1}{\cos^2 \theta} \frac{d\theta}{db} = \frac{2E}{GM\mu}.$$

Für den Streuwinkel $\chi = \pi - 2\theta$ ergibt sich damit

$$\frac{d\chi}{db} = -2 \frac{d\theta}{db} = \frac{-4E}{GM_0 m} \cos^2 \theta.$$

Weiter finden wir unter Verwendung von

$$\sin \chi = \sin 2\theta = 2 \sin \theta \cos \theta; \quad \tan \theta = \frac{\sin \theta}{\cos \theta},$$

$$\frac{b}{\sin \chi} = \frac{b}{2 \sin \theta \cos \theta} = \frac{GM\mu}{4E} \frac{1}{\cos^2 \theta}.$$

Für den differentiellen Streuquerschnitt erhält man

$$\frac{d\sigma}{d\Omega} = \left| \frac{b}{\sin \chi} \left(\frac{d\chi}{db} \right)^{-1} \right|^2 = \left(\frac{GM\mu}{4E \sin^2 \frac{\chi}{2}} \right)^2, \quad (2.2.20)$$

wobei wir $\cos \theta = \cos(\pi/2 - \chi/2) = \sin(\chi/2)$ geschrieben haben. Für geladene Teilchen erscheint das Produkt der Ladungen $e_1 e_2$ anstelle von $GM\mu$ (*Streuformel von Rutherford*). Überraschenderweise gilt dasselbe Resultat auch in der Quantenmechanik. Dies ist ein Zufall, welcher nur im Fall des Coulomb- und Gravitationspotentials auftritt. Darum war es nicht möglich, mit Hilfe des Rutherford-Experiments einen Unterschied zwischen klassischer und Quanten-Physik festzustellen.

2.2.3 Runge-Lenz Vektor

Das Kepler-Problem weist gegenüber dem allgemeinen Zweikörper-Problem ein zusätzliches Bewegungsintegral auf, nämlich den Runge-Lenz Vektor

$$\mathbf{A} = \mu \dot{\mathbf{x}} \wedge \mathbf{L} - G M_0 m \mu \frac{\mathbf{x}}{r}. \quad (2.2.21)$$

In der Tat rechnet man nämlich nach, dass

$$\begin{aligned} \mu \frac{d}{dt} \frac{\mathbf{x}}{r} &= \mu \left(\frac{\dot{\mathbf{x}}}{r} - \frac{\mathbf{x}}{r^2} \underbrace{\dot{r}}_{\mathbf{x} \cdot \dot{\mathbf{x}}/r} \right) = \frac{\mu}{r^3} (\dot{\mathbf{x}} r^2 - \mathbf{x}(\mathbf{x} \cdot \dot{\mathbf{x}})) \\ &= -\frac{\mu}{r^3} \mathbf{x} \wedge (\mathbf{x} \wedge \dot{\mathbf{x}}) = -\frac{1}{r^3} \mathbf{x} \wedge \mathbf{L}, \end{aligned}$$

wobei die zweitletzte Gleichung aus dem allgemein für drei Vektoren \mathbf{a} , \mathbf{b} und \mathbf{c} geltenden Zusammenhang

$$\mathbf{a} \wedge (\mathbf{b} \wedge \mathbf{c}) = \mathbf{b} \cdot (\mathbf{a} \cdot \mathbf{c}) - \mathbf{c} \cdot (\mathbf{a} \cdot \mathbf{b})$$

folgt. Somit folgt unter Verwendung der Bewegungsgleichung (2.2.3) und der Zeitunabhängigkeit von \mathbf{L}

$$\frac{d\mathbf{A}}{dt} = -\frac{GM_0 m}{r^3}(\mathbf{x} \wedge \mathbf{L} - \mathbf{x} \wedge \mathbf{L}) = 0. \quad (2.2.22)$$

Der Runge-Lenz Vektor ist orthogonal zu \mathbf{L} , $\mathbf{A} \cdot \mathbf{L} = 0$, da beide Terme separat orthogonal zu \mathbf{L} stehen. Seine Länge ist

$$\mathbf{A}^2 = (GM_0 m \mu)^2 + \mu^2 \dot{\mathbf{x}}^2 l^2 - \frac{2GM_0 m \mu}{r} \underbrace{\mu \mathbf{x} \cdot (\dot{\mathbf{x}} \wedge \mathbf{L})}_{\mathbf{L} \cdot (\mathbf{x} \wedge \dot{\mathbf{x}}) = \frac{l^2}{\mu}} = (GM_0 m \mu)^2 + 2\mu E l^2, \quad (2.2.23)$$

wobei wir die Formel für E (2.1.10) verwendet haben. Der Runge-Lenz Vektor liegt in der Bahnebene, und wir können φ so wählen, dass $\mathbf{x} \cdot \mathbf{A} = rA \cos \varphi$. Andererseits berechnen wir direkt aus der Formel für \mathbf{A}

$$\mathbf{x} \cdot \mathbf{A} = \mu(\mathbf{x} \wedge \dot{\mathbf{x}}) \cdot \mathbf{L} - GM_0 m \mu r = l^2 - GM_0 m \mu r, \quad (2.2.24)$$

und daher ist

$$r = \frac{l^2}{GM_0 m \mu + A \cos \varphi} = \frac{d}{1 + \varepsilon \cos \varphi} \quad (2.2.25)$$

mit

$$d = \frac{l^2}{GM_0 m \mu}, \quad 1 - \varepsilon^2 = 1 - \left(\frac{A}{GM_0 m \mu}\right)^2 = \frac{-2El^2}{G^2 M_0^2 m^2 \mu}, \quad (2.2.26)$$

in Übereinstimmung mit (2.2.8). Insbesondere ist die Richtung von \mathbf{A} die vom Ursprung zum Perihel. Aus der Erhaltung von \mathbf{A} folgt schliesslich die Bahnkurve von $\dot{\mathbf{x}}(t)$: wegen

$$\begin{aligned} \mathbf{L} \wedge \mathbf{A} &= \mu \mathbf{L} \wedge (\dot{\mathbf{x}} \wedge \mathbf{L}) - GM_0 m \mu \mathbf{L} \wedge \frac{\mathbf{x}}{r} \\ &= \mu (l^2 \dot{\mathbf{x}} - (\mathbf{L} \cdot \dot{\mathbf{x}}) \mathbf{L}) - GM_0 m \mu \mathbf{L} \wedge \frac{\mathbf{x}}{r} \\ &= l^2 \mu \dot{\mathbf{x}} - GM_0 m \mu \mathbf{L} \wedge \frac{\mathbf{x}}{r} \end{aligned}$$

ist

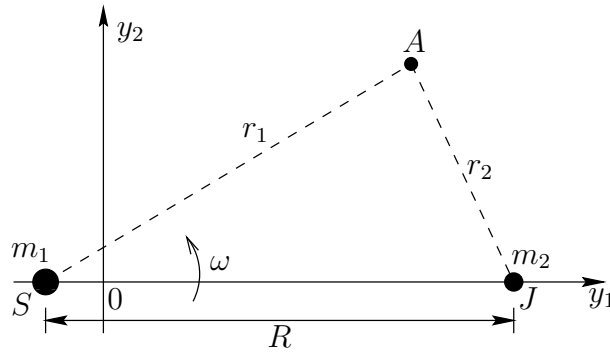
$$\dot{\mathbf{x}} = \frac{1}{\mu l^2} \mathbf{L} \wedge \mathbf{A} + \frac{GM_0 m}{l} \frac{\mathbf{L}}{l} \wedge \frac{\mathbf{x}}{r}, \quad (2.2.27)$$

d.h. $\dot{\mathbf{x}}(t)$ beschreibt einen Kreis um $\frac{1}{\mu l^2} \mathbf{L} \wedge \mathbf{A}$ vom Radius $GM_0 m l^{-1}$.

3 Ausgewählte Dreikörperprobleme

Im Gegensatz zu dem Zweikörperproblem, das wir im vorigen Kapitel behandelt haben, ist das Dreikörperproblem nicht mehr integrabel. Man kann jedoch spezielle Lösungen (zum Teil näherungsweise) finden.

3.1 Gleichgewichtslagen und ihre Stabilität



Sonne S und Jupiter J bewegen sich auf Kreisbahnen mit Winkelgeschwindigkeit ω um den Schwerpunkt 0 . In ihrem Gravitationsfeld bewegt sich ein Asteroid A , der keinen Einfluss auf S und J haben soll. (Das ist eine Näherung!) Als Einheiten wählen wir

$$\omega = 1, \quad R = 1, \quad G = 1. \quad (3.1.1)$$

Da das in der Figur definierte y -System um den in Ursprung gewählten Schwerpunkt rotiert, sind die Bewegungsgleichungen für rotierende Bezugssysteme (1.7.11) zu verwenden. Insbesondere müssen sich die Gravitations- und Zentrifugalkraft auf S und J aufheben; im Schwerpunktsystem hat J den Abstand $r_J = Rm_1/(m_1 + m_2)$ vom Ursprung, während der Abstand von S gerade $r_S = Rm_2/(m_1 + m_2)$ ist. Die Gravitationskraft und Zentrifugalkraft heben sich deshalb gerade auf, falls

$$\frac{G}{R^2} m_1 m_2 = \frac{m_1 m_2}{m_1 + m_2} R \omega^2. \quad (3.1.2)$$

In den obigen Einheiten gilt deshalb auch

$$m_1 + m_2 = 1, \quad (3.1.3)$$

und S , J haben die Koordinaten $(-m_2, 0, 0)$, $(m_1, 0, 0)$. Da alle Kräfte auf A proportional zu seiner Masse sind, fällt diese aus den Bewegungsgleichungen für A heraus; wir können daher formal $m = 1$ setzen. Dann sind die Kräfte, die auf A wirken:

$$\begin{aligned} \text{Gravitationskraft:} \quad \mathbf{G} &= -\frac{m_1}{r_1^3} (y_1 + m_2, y_2, y_3) - \frac{m_2}{r_2^3} (y_1 - m_1, y_2, y_3). \\ \text{Zentrifugalkraft:} \quad \mathbf{Z} &= (y_1, y_2, 0). \\ \text{Corioliskraft:} \quad \mathbf{C} &= 2(\dot{y}_2, -\dot{y}_1, 0). \end{aligned} \quad (3.1.4)$$

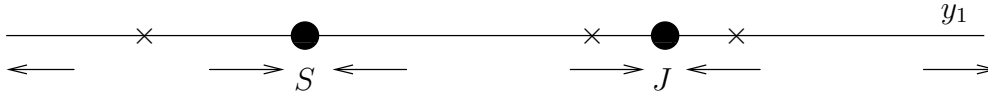
Wir suchen Gleichgewichtslösungen $\mathbf{y} = \text{konstant}$, d.h. Lösungen \mathbf{y} von $\mathbf{G} + \mathbf{Z} = 0$:

$$\begin{aligned} \left(\frac{m_1}{r_1^3} + \frac{m_2}{r_2^3} - 1 \right) y_1 + m_1 m_2 \left(\frac{1}{r_1^3} - \frac{1}{r_2^3} \right) &= 0, \\ \left(\frac{m_1}{r_1^3} + \frac{m_2}{r_2^3} - 1 \right) y_2 &= 0, \quad y_3 = 0. \end{aligned} \quad (3.1.5)$$

Es gibt dafür 2 Möglichkeiten: Entweder ist $y_2 = 0$ (Eulerscher Spezialfall) oder

$$\frac{m_1}{r_1^3} + \frac{m_2}{r_2^3} = 1 \quad (\text{Lagrangescher Spezialfall}). \quad (3.1.6)$$

Im ersten Fall liegt A auf der y_1 -Achse, wo $\mathbf{G} + \mathbf{Z}$ in etwa so gerichtet ist:



Somit gibt es dort 3 Gleichgewichtslagen (× im Diagramm), die wir nicht weiter verfolgen.

Im zweiten Fall folgt aus (3.1.5) und (3.1.6)

$$r_1 = r_2 = 1,$$

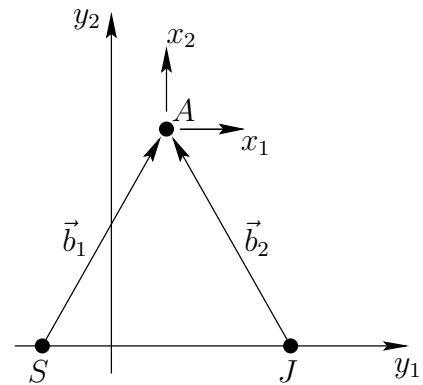
d.h. SJA ist ein gleichseitiges Dreieck in der 12-Ebene. Es gibt also zwei Lagrangesche Gleichgewichtslagen ($y_2 > 0$ oder $y_2 < 0$), deren Stabilität wir untersuchen. Dazu betrachten wir kleine Auslenkungen \mathbf{x} aus der Ruhelage $\mathbf{x} = 0$, d.h. $\mathbf{y} = (\frac{1}{2} - m_2, \frac{\sqrt{3}}{2}, 0) + \mathbf{x}$, in denen wir \mathbf{G} (die einzige nicht-lineare Kraft) linearisieren, das heisst, wir linearisieren $r_i^{-3} = |\mathbf{b}_i + \mathbf{x}|^{-3}$ um $\mathbf{x} = 0$. Da $|\mathbf{b}_i| = 1$ ist, finden wir in erster Ordnung

$$r_i^{-3} = 1 + \frac{\partial}{\partial \mathbf{z}} (|\mathbf{z}|^{-3}) \Big|_{\mathbf{b}_i} \cdot \mathbf{x} = 1 - 3 \mathbf{b}_i \cdot \mathbf{x}$$

und damit

$$\begin{aligned} \mathbf{G} &= -m_1(1 - 3\mathbf{b}_1 \cdot \mathbf{x}) \left(\frac{1}{2} + x_1, \frac{\sqrt{3}}{2} + x_2, x_3 \right) \\ &\quad - m_2(1 - 3\mathbf{b}_2 \cdot \mathbf{x}) \left(-\frac{1}{2} + x_1, \frac{\sqrt{3}}{2} + x_2, x_3 \right) \\ &\simeq \mathbf{G}^0 + \frac{3}{2} m_1 \mathbf{b}_1 \cdot \mathbf{x} (1, \sqrt{3}, 0) + \frac{3}{2} m_2 \mathbf{b}_2 \cdot \mathbf{x} (-1, \sqrt{3}, 0) - (m_1 + m_2)(x_1, x_2, x_3) \\ &\equiv \mathbf{G}^0 + (G_1, G_2, G_3), \end{aligned}$$

wobei \mathbf{G}^0 die Glieder nullter Ordnung darstellt, $\mathbf{G}^0 = (\frac{m_2 - m_1}{2}, -\frac{\sqrt{3}}{2}, 0)$, wir in der zweiten



$$\begin{aligned} \mathbf{b}_1 &= \frac{1}{2}(1, \sqrt{3}, 0), \\ \mathbf{b}_2 &= \frac{1}{2}(-1, \sqrt{3}, 0). \end{aligned}$$

Zeile wiederum Terme höherer Ordnung weggelassen haben, und

$$\begin{aligned} G_1 &= \left(\frac{3}{4} - 1\right) \underbrace{(m_1 + m_2)}_{=1} x_1 + \frac{3\sqrt{3}}{4} (m_1 - m_2) x_2, \\ G_2 &= \frac{3}{4} \sqrt{3} (m_1 - m_2) x_1 + \left(\frac{9}{4} - 1\right) \underbrace{(m_1 + m_2)}_{=1} x_2, \\ G_3 &= -x_3. \end{aligned}$$

Weiterhin ist

$$\mathbf{Z} = \mathbf{Z}^0 + (x_1, x_2, 0), \quad \mathbf{C} = 2(\dot{x}_2, -\dot{x}_1, 0), \quad (3.1.7)$$

wobei $\mathbf{Z}^0 = (\frac{1}{2} - m_2, \frac{\sqrt{3}}{2}, 0)$ wiederum den Beitrag nullter Ordnung darstellt. Nach Konstruktion gilt $\mathbf{G}^0 + \mathbf{Z}^0 = 0$ (Gleichgewichtslage!), und wir erhalten die Bewegungsgleichungen in linearer Näherung:

$$\begin{aligned} \ddot{x}_1 &= \frac{3}{4} x_1 + \frac{3\sqrt{3}}{4} (m_1 - m_2) x_2 + 2\dot{x}_2 \\ \ddot{x}_2 &= \frac{3\sqrt{3}}{4} (m_1 - m_2) x_1 + \frac{9}{4} x_2 - 2\dot{x}_1 \\ \ddot{x}_3 &= -x_3. \end{aligned} \quad (3.1.8)$$

Die x_3 -Bewegung ist entkoppelt (harmonische Schwingung der Frequenz $\omega = 1$, synchron mit der Kreisbewegung von S und J). Zur Lösung der beiden anderen Gleichungen machen wir den Exponentialansatz

$$x_k(t) = a_k e^{i\lambda t}, \quad (k = 1, 2), \quad (3.1.9)$$

denn Real- und Imaginärteile einer komplexen Lösung von (3.1.8) sind wieder Lösungen (da die Koeffizienten der Differentialgleichung reell sind). Einsetzen liefert das homogene Gleichungssystem

$$\begin{aligned} a_1 \left(\lambda^2 + \frac{3}{4}\right) + a_2 \left(\frac{3\sqrt{3}}{4} (m_1 - m_2) + 2i\lambda\right) &= 0 \\ a_1 \left(\frac{3\sqrt{3}}{4} (m_1 - m_2) - 2i\lambda\right) + a_2 \left(\lambda^2 + \frac{9}{4}\right) &= 0. \end{aligned}$$

Die nicht-triviale Lösung verlangt, dass die Determinante des Gleichungssystems verschwindet:

$$\left(\lambda^2 + \frac{3}{4}\right) \left(\lambda^2 + \frac{9}{4}\right) = \frac{27}{16} (m_1 - m_2)^2 + 4\lambda^2. \quad (3.1.10)$$

Wegen $(m_1 - m_2)^2 = (m_1 + m_2)^2 - 4m_1 m_2$ und (3.1.3) lautet diese Gleichung für λ

$$\left(\lambda^2 - \frac{1}{2}\right)^2 = \frac{1}{4} (1 - 27m_1 m_2). \quad (3.1.11)$$

Für $27m_1m_2 < 1$, also in beliebigen Masseneinheiten für

$$\frac{m_1m_2}{(m_1 + m_2)^2} < \frac{1}{27} \quad (3.1.12)$$

d.h. für [schreibe $m_1 = (\frac{1}{2} - x)(m_1 + m_2)$ und $m_2 = (\frac{1}{2} + x)(m_1 + m_2)$; dann ist $m_1m_2 = (\frac{1}{4} - x^2)(m_1 + m_2)^2$]

$$\frac{m_1}{m_1 + m_2} \quad \text{oder} \quad \frac{m_2}{m_1 + m_2} < \frac{1}{2} - \frac{\sqrt{69}}{18} = 0.0385, \quad (3.1.13)$$

gibt es 4 verschiedene reelle Lösungen $\pm\lambda_1, \pm\lambda_2$ zu (3.1.11) und die allgemeine Lösung von (3.1.8) ist eine beliebige Superposition der 4 Eigenschwingungen (3.1.9). Da die λ_i reell sind, verhalten sich alle diese Lösungen oszillatorisch und beschränkt. Zumindest in linearer Näherung ist die Gleichgewichtslage daher stabil, falls (3.1.12) erfüllt ist. Gilt die entgegengesetzte Ungleichung, so treten komplexe Eigenfrequenzen auf (darunter solche mit $\text{Im } \lambda < 0$), d.h. exponentiell wachsende Lösungen. Die Gleichgewichtslage $\vec{x} = 0$ ist dann instabil.

Im Fall von Sonne und Jupiter ist (3.1.13) erfüllt: $m_2 \approx m_1 \cdot 10^{-3}$. Tatsächlich gibt es in der Gegend der Lagrangeschen Gleichgewichtslagen zahlreiche Asteroide.

[Ohne Beweis: Für genügend kleine Werte des Verhältnisses (3.1.12) gilt die Stabilität auch ausserhalb der linearen Approximation. Die Eulerschen Gleichgewichtslagen sind hingegen instabil.]

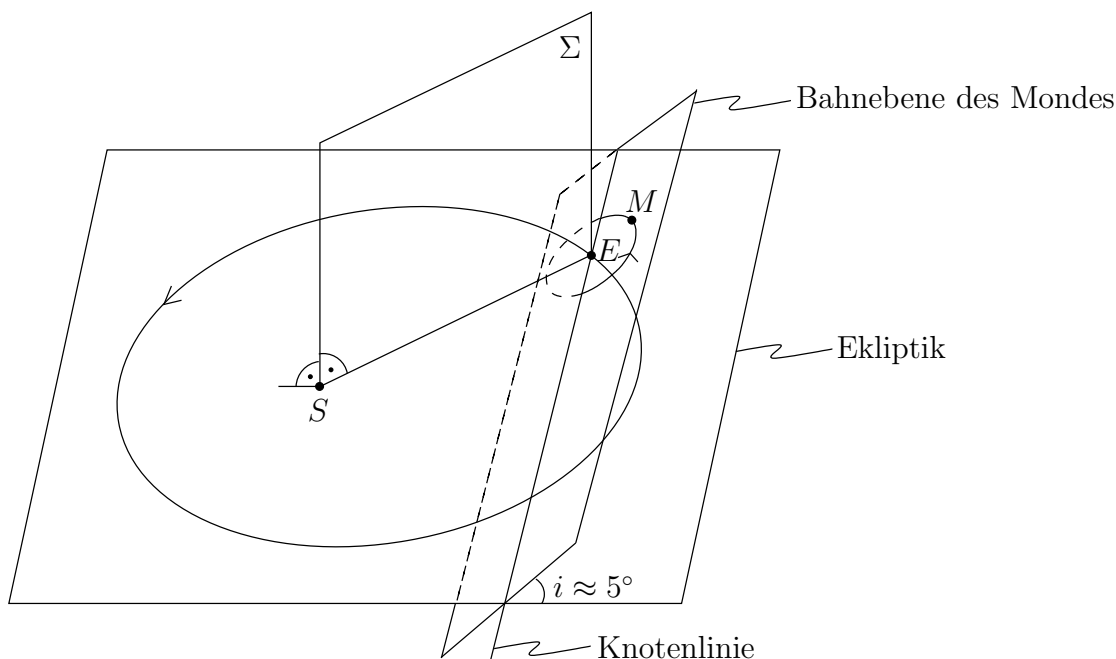
3.2 Die Bewegung des Mondes

Der Mond M kreist um die Erde E in einer Bahnebene die leicht schief zur Ekliptik (Bahnebene der Erde um die Sonne S) steht. Es lassen sich deshalb (mindestens) zwei Perioden unterscheiden:

- $2\pi\mu$ = synodischer Monat
- = Periode zwischen aufeinander folgenden, gleich gerichteten Durchgängen durch Σ (von Neumond zu Neumond)
- = 29.53059 Tage;
- $2\pi\tilde{\mu}$ = drakonitischer Monat
- = Periode zwischen aufeinander folgenden, gleich gerichteten Durchgängen durch die Ekliptik (von aufsteigendem Knoten zu aufsteigendem Knoten)
- = 27.21222 Tage.

Sie stehen im Verhältnis

$$r = \frac{\mu}{\tilde{\mu}} = 1.08520. \quad (3.2.1)$$



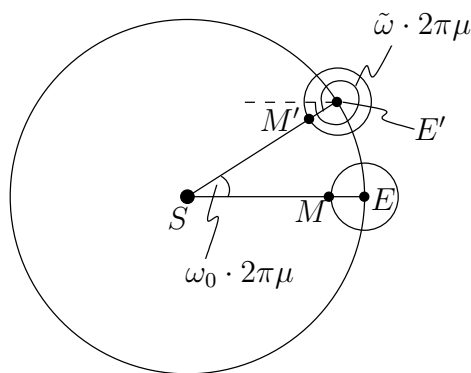
Die Diskrepanz kann zum grössten Teil rein kinematisch erklärt werden. Wäre die Bahnebene des Mondes raumfest orientiert, so wäre (bei Kreisbahnen mit kleiner Neigung i)

$$\tilde{\omega} \cdot 2\pi\mu = 2\pi + \omega_0 \cdot 2\pi\mu,$$

wobei $\tilde{\omega} = \tilde{\mu}^{-1}$, $\omega_0 = J^{-1}$, $2\pi J = 1$ Jahr.
Damit wäre

$$\frac{\mu}{\tilde{\mu}} = 1 + \frac{\mu}{J} = 1.08085 \cong 1 + \frac{1}{12.3}. \quad (3.2.2)$$

Die Abweichung von dem beobachteten Wert (3.2.1) lässt auf eine langsame Drehung der Bahnebene des Mondes schliessen.



Das Verhältnis (3.2.1) bestimmt die Kadenz der Sonnenfinsternisse (bzw. Mondfinsternisse): Ausgehend von einer exakten Konjunktion SME tritt eine weitere Finsternis immer dann auf, wenn

$$p\mu = q\tilde{\mu}, \quad (p, q \text{ ganz})$$

(im selben Knoten) oder

$$p\mu = \left(q + \frac{1}{2}\right)\tilde{\mu}$$

(im entgegengesetzten Knoten). Wegen der endlichen Ausdehnung von E und M müssen diese Bedingungen nur innerhalb einer ‘Toleranz’ Δ erfüllt sein:

$$\text{dist}\left(p \cdot r, \left\{ \begin{array}{l} \mathbb{Z} \\ \mathbb{Z} + \frac{1}{2} \end{array} \right\}\right) < \Delta/2 ; \quad \Delta \cong 0.087 .$$

Wegen $r \cong 1 + 1/12$ tritt nach 12 synodischen Monaten erneut eine Finsternis im gleichen Knoten auf:

$$12r = 13.0224 .$$

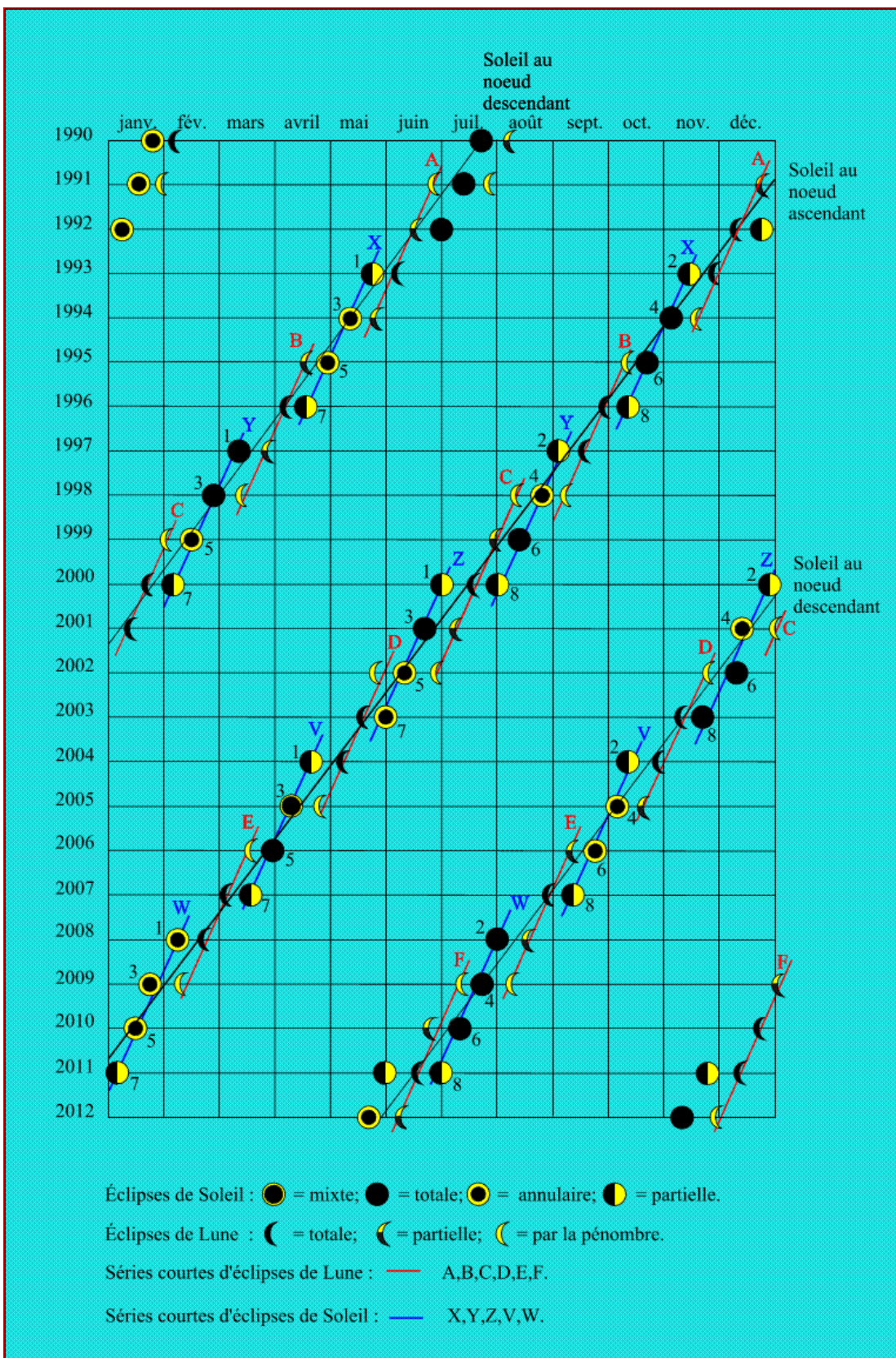
Diese Periodizität ist aber nicht exakt: Bereits nach

$$\frac{\Delta}{12r - 13} \cong 3.88$$

Zyklen, bzw. dem ganzen Teil davon, ist die Toleranz erschöpft und die Periodizität bricht ab. Weitaus genauer ist die Periodizität

$$\begin{aligned} 223r &= 241.999 \cong 242 , \\ 223 \cdot 2\pi\mu &\equiv 1 \text{ Saros} \cong 18 \text{ Jahre} + 11 \text{ Tage} + 8 \text{ Stunden} , \end{aligned}$$

die schon den Chaldäern bekannt war. Sie dauert etwa 1200 Jahre an, bis sie abbricht. [Siehe Kanon der Finsternisse (www.bdl.fr/ephem/eclipses/soleil/Soleil.html).]



Wir wollen nun das wichtige Verhältnis (3.2.1) möglichst genau theoretisch als Funktion von μ/J bestimmen. Dazu behandeln wir das System S, E, M (mit Massen $M \gg m \gg m'$) als restringiertes Dreikörperproblem, d.h. wir betrachten den Mond als Körper in dem externen Gravitationsfeld von Sonne und Erde. Die Erde beschreibt eine Kreisbahn (Radius R) um die Sonne, bzw. umgekehrt beschreibt die Sonne eine Kreisbahn im rotierenden Bezugssystem der Erde mit Ursprung E. Wir wählen die Einheiten:

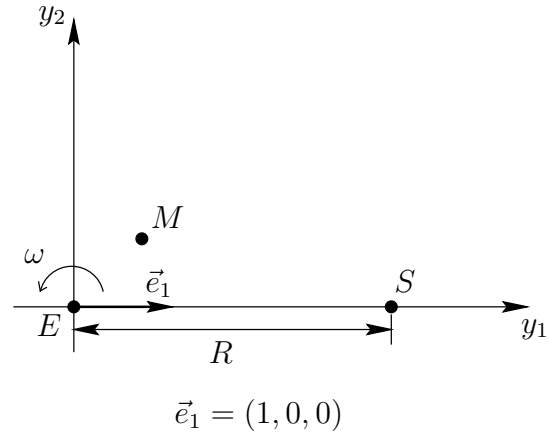
$$\omega = J = 1, \quad m = 1, \quad G = 1,$$

also

$$\mu = 0.08085. \quad (3.2.3)$$

Im 2-Körpersystem Sonne-Erde müssen sich wiederum Gravitations- und Zentrifugalkraft aufheben:

$$G \frac{mM}{R^2} = \frac{mM}{M+m} R\omega^2 \cong m R \omega^2,$$



und daher muss gelten

$$M = R^3$$

in unseren Einheiten. Auf den Mond wirken folgende Kräfte (nach Division durch m'). Zum einen haben wir die Gravitationskräfte von Erde und Sonne:

$$\begin{aligned} \mathbf{G}_E &= \frac{\partial}{\partial \mathbf{y}} \frac{1}{|\mathbf{y}|} = -\frac{1}{|\mathbf{y}|^3} \mathbf{y}, \\ \mathbf{G}_S &= M \frac{\partial}{\partial \mathbf{y}} \frac{1}{|R\mathbf{e}_1 - \mathbf{y}|} \\ &= M \left(\frac{\mathbf{e}_1}{R^2} + \frac{3(\mathbf{e}_1 \cdot \mathbf{y})\mathbf{e}_1 - \mathbf{y}}{R^3} + O(y^2) \right) \\ &= R\mathbf{e}_1 + 3y_1\mathbf{e}_1 - \mathbf{y} + \dots \end{aligned} \quad (3.2.4)$$

Die Linearisierung in \mathbf{y} ist angebracht, da längs der Mondbahn $|\mathbf{y}|/R \cong 1/390$. Der erste Term kompensiert natürlich die Führungsbeschleunigung (siehe (1.7.11))

$$-\mathbf{a} = \boldsymbol{\omega} \wedge (\boldsymbol{\omega} \wedge R\mathbf{e}_1) = -\omega^2 R\mathbf{e}_1 = -R\mathbf{e}_1.$$

Weiterhin wirkt die Zentrifugalbeschleunigung

$$\mathbf{Z} = (y_1, y_2, 0)$$

und kompensiert damit die 1,2-Komponenten von $-\mathbf{y}$ in (3.2.4). Schliesslich ist die Coriolisbeschleunigung

$$\mathbf{C} = 2(\dot{y}_2, -\dot{y}_1, 0).$$

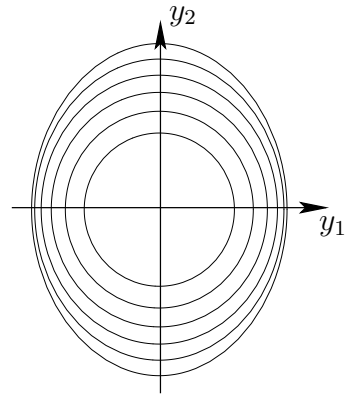
Also lauten die Bewegungsgleichungen des Mondes mit $r = |\mathbf{y}|$

$$\begin{aligned} \ddot{y}_1 - 2\dot{y}_2 &= \left(3 - \frac{1}{r^3}\right) y_1, \\ \ddot{y}_2 + 2\dot{y}_1 &= -\frac{1}{r^3} y_2, \\ \ddot{y}_3 &= -\left(\frac{1}{r^3} + 1\right) y_3. \end{aligned} \quad (3.2.5)$$

Da die Neigung der Mondbahn gegenüber der Ekliptik klein ist ($i \cong 5^\circ$) bestimmen wir nach Hill (1878) zunächst periodische Bahnen in der 12-Ebene und dann, als deren Störung, die eigentliche Mondbahn.

3.2.1 Periodische Bahnen

Sei $y_3 \equiv 0$. Wir konstruieren eine 1-parametrische Schar periodischer Bahnen zu (3.2.5), die um $\mathbf{y} = 0$ im Gegenuhrzeigersinn kreisen, und zwar genau eine Bahn (bis auf Verschiebung der Zeit) durch jeden Punkt der (y_1, y_2) -Ebene in der Nähe von $\mathbf{y} = 0$. Zunächst ist aus (3.2.5) ersichtlich, dass mit jeder Lösung $(y_1(t), y_2(t))$ auch $(y_1(-t), -y_2(-t))$ und $(-y_1(-t), y_2(-t))$ Lösungen sind. Sie entstehen durch Zeitumkehr und Spiegelung an der 1- bzw. 2-Achse. Für einen Anfangspunkt $(y_1(0) > 0, y_2(0) = 0)$ auf der 1-Achse kann man eine vertikale Anfangsgeschwindigkeit $(\dot{y}_1(0) = 0, \dot{y}_2(0) > 0)$ so wählen (und wir werden dies z.T. verifizieren), dass die Bahn die 2-Achse horizontal schneidet: $y_1(\pi\mu/2) = 0, \dot{y}_2(\pi\mu/2) = 0$. Aus den erwähnten Symmetrien und der Eindeutigkeit der Lösung zu gegebenen Anfangsbedingungen folgt



$$\begin{aligned} y_1(t) &= y_1(-t), & y_2(t) &= -y_2(-t), \\ y_1(\pi\mu/2 + t) &= -y_1(\pi\mu/2 - t), & y_2(\pi\mu/2 + t) &= y_2(\pi\mu/2 - t). \end{aligned} \quad (3.2.6)$$

Insbesondere ist die Bahn periodisch, $\mathbf{y}(t+2\pi\mu) = \mathbf{y}(t)$, und oval: spiegelsymmetrisch bzgl. der 1- und 2-Achsen. Als Scharparameter wählen wir die Umlaufszeit $2\pi\mu$ (synodischer Monat) und wählen dann die Zeiteinheit neu:

$$\tau = t/\mu$$

(neu: synodischer Monat = 2π). Dann lautet (3.2.5) mit $' = d/d\tau$

$$\begin{aligned} y_1'' - 2\mu y_2' &= -\frac{\mu^2}{r^3} y_1 + \underline{3\mu^2 y_1}, \\ y_2'' + 2\mu y_1' &= -\frac{\mu^2}{r^3} y_2, \\ y_3'' &= -\mu^2 \left(\frac{1}{r^3} + 1\right) y_3. \end{aligned} \quad (3.2.7)$$

Ohne den unterstrichenen Term (Gezeitenkräfte), den wir als Störung betrachten, lauten die ersten beiden Gleichungen

$$(y_1 + iy_2)'' + 2i\mu(y_1 + iy_2)' = -\frac{\mu^2}{r^3}(y_1 + iy_2)$$

und die gesuchte (nun 2π -)periodische Lösung ist eine Kreisbahn

$$y_1(\tau) + iy_2(\tau) = r_0 e^{i\tau}, \quad y_3(\tau) = 0 \quad (3.2.8)$$

mit

$$r_0 = \frac{\mu^{\frac{2}{3}}}{(1 + 2\mu)^{\frac{1}{3}}}.$$

Nachträglich können wir nun die Gezeitenstörung in (3.2.7) behandeln; wenn wir die Lösung (3.2.8) in den Korrekturterm einsetzen, ist er um μ^2 kleiner als der führende Term. Man kann ihn daher in einer Störungsrechnung berücksichtigen. Statt \mathbf{y} führen wir noch neue Koordinaten $(z, \zeta) \in \mathbb{C} \times \mathbb{R}$ ein durch

$$\begin{aligned} y_1 + iy_2 &= \mu^{2/3}(1 + 2\mu)^{-1/3} e^{i\tau} z, \\ y_3 &= \mu^{2/3}(1 + 2\mu)^{-1/3} \zeta \end{aligned}$$

(‘Variation der Konstanten’). Sie sind der obigen Näherungslösung (3.2.8) angepasst ($z(\tau) = 1, \zeta(\tau) = 0$) und entsprechen einem Bezugssystem, das sich zusätzlich mit einer Umdrehung pro Monat um die y_3 -Achse dreht. Gl. (3.2.7) lautet nun

$$z'' + 2i(1 + \mu)z' = (1 + 2\mu)\left(1 - \frac{1}{\rho^3}\right)z + \frac{3}{2}\mu^2(z + e^{-2i\tau}\bar{z}), \quad (3.2.9)$$

$$\zeta'' = -\left(\frac{1 + 2\mu}{\rho^3} + \mu^2\right)\zeta, \quad (3.2.10)$$

wobei $\rho^2 = |z|^2 + \zeta^2$. Natürlich ist ohne den Störterm die vorige Näherungslösung (also $z(\tau) = 1, \zeta(\tau) = 0$) immer noch gültig. Die gesuchte 2π -periodische Lösung setzen wir nun als Potenzreihe in μ an:

$$z(\tau) = 1 + \sum_{k=2}^{\infty} \mu^k z_k(\tau) \equiv 1 + \tilde{z}(\tau) \quad (3.2.11)$$

mit

$$\begin{aligned} z_k(\tau + 2\pi) &= z_k(\tau) \\ \overline{z_k(\tau)} &= z_k(-\tau) \\ \overline{z_k\left(\frac{\pi}{2} + \tau\right)} &= z_k\left(\frac{\pi}{2} - \tau\right). \end{aligned} \quad (3.2.12)$$

[Der Störterm ist $O(\mu^2)$, also beginnt die Entwicklung (3.2.11) mit $k = 2$; (3.2.12) folgt aus (3.2.6).]

Zur Berechnung von $z_2(\tau)$, entwickeln wir ρ zu führender Ordnung. Für $\zeta \equiv 0$ gilt

$$\begin{aligned}\rho^2 &= |z|^2 = (1 + \tilde{z})(1 + \bar{\tilde{z}}) = 1 + (z_2 + \bar{z}_2)\mu^2 + O(\mu^4) \\ \rho^{-3} &= 1 - \frac{3}{2}(z_2 + \bar{z}_2)\mu^2 + O(\mu^4) \\ z\rho^{-3} &= 1 + \tilde{z} - \frac{3}{2}(\tilde{z} + \bar{\tilde{z}}) + O(\mu^4) = 1 - \left(\frac{1}{2}z_2 + \frac{3}{2}\bar{z}_2\right)\mu^2 + O(\mu^4).\end{aligned}\quad (3.2.13)$$

Vergleich der Terme $\propto \mu^2$ in (3.2.9) ergibt

$$z_2'' + 2i z_2' = \frac{3}{2}(z_2 + \bar{z}_2) + \frac{3}{2}(1 + e^{-2i\tau}). \quad (3.2.14)$$

Der Ansatz $z_2(\tau) = a_0 + a_-e^{-2i\tau} + a_+e^{2i\tau}$ (mit $a_{0,\pm} \in \mathbb{R}$ wegen (3.2.12)) liefert

$$z_2(\tau) = -\frac{1}{2} - \frac{19}{16}e^{-2i\tau} + \frac{3}{16}e^{2i\tau}. \quad (3.2.15)$$

[Die hinzukommende Lösung der homogenen Gleichung (3.2.14), $z_2^0(\tau) = ae^{i\tau} - 3\bar{a}e^{-i\tau}$, genügt nicht (3.2.12), und trägt daher nicht bei.] Gl. (3.2.15) beschreibt in führender Ordnung die ovale Gestalt der periodischen Bahn. Der Unterschied der ‘Halbachsen’ beträgt

$$\Delta = \mu^2 \left(z_2 \left(\frac{\pi}{2} \right) - z_2(0) \right) = 2\mu^2. \quad (3.2.16)$$

Für $\mu = 0.08085$ ist das ungefähr $\Delta = 1.3 \cdot 10^{-2}$; mit unseren Einheiten ist der Abstand Erde-Mond auf 1 normiert.

Die Bahn des Mondes ist tatsächlich jedoch nicht periodisch, und die Änderungen des Abstandes Erde-Mond sind in Wahrheit grösser.

3.2.2 Variationen der Bahn

Der Ansatz für Bahnen, die nahe bei der periodischen Lösung ($z(\tau), \zeta(\tau) \equiv 0$), liegen, lautet

$$(z(\tau) + \hat{z}(\tau), \hat{\zeta}(\tau)), \quad (3.2.17)$$

wobei \hat{z} und $\hat{\zeta}$ klein sind. Wir setzen diesen Ansatz in die Bewegungsgleichungen (3.2.9) und (3.2.10) ein, und linearisieren in $\hat{z}, \hat{\zeta}$, d.h. wir ‘variieren’ die periodische Lösung.

Da $\frac{\partial \rho}{\partial \zeta}|_{\zeta=0} = 0$ enthält die Gleichung, die man von (3.2.9) erhält, $\hat{\zeta}$ nicht, d.h. es kommt nur \hat{z} vor. Insbesondere wird daher diese Gleichung durch $\hat{z} \equiv 0$ gelöst. Die Variation von (3.2.10) lautet dann

$$\hat{\zeta}'' = -\left(\frac{1 + 2\mu}{\rho^3} + \mu^2\right)\hat{\zeta}, \quad (3.2.18)$$

wobei sich ρ nun auf die periodische Bahn bezieht. Einsetzen von (3.2.15) in (3.2.13) führt zu

$$\rho^{-3} = 1 + \frac{3}{2}\mu^2(1 + e^{2i\tau} + e^{-2i\tau}) + O(\mu^4). \quad (3.2.19)$$

Damit wird (3.2.18)

$$\begin{aligned}\hat{\zeta}'' &= -\left(1 + 2\mu + \frac{3}{2}\mu^2(1 + 2\cos 2\tau) + \mu^2 + \dots\right)\hat{\zeta} \\ &= -\left(1 + 2\mu + \frac{5}{2}\mu^2 + 3\mu^2\cos 2\tau + \dots\right)\hat{\zeta}.\end{aligned}\quad (3.2.20)$$

Lässt man den Term

$$3\mu^2\cos 2\tau \quad (3.2.21)$$

weg, so resultiert ein harmonischer Oszillator in $\hat{\zeta}$ der Frequenz

$$r = \left(1 + 2\mu + \frac{5}{2}\mu^2\right)^{1/2} = 1 + \mu + \frac{3}{4}\mu^2 + \dots \quad (3.2.22)$$

In beliebigen Einheiten ist dies das Verhältnis (3.2.1) von synodischem zu drakonitischem Monat. In 1. Ordnung $r = 1 + \mu$ stimmt dies mit (3.2.2) überein. Für den Mond mit $\mu = 0.08085$ (siehe (3.2.3)) ergibt sich (inklusive 2. Ordnung Korrektur)

$$r = 1.08575 \quad (3.2.23)$$

in guter Übereinstimmung mit (3.2.1). Der Term (3.2.21) hat zur Folge, dass $\hat{\zeta}(\tau)$ zwar nicht mehr periodisch ist, aber eine mittlere Frequenz r besitzt. Wegen des oszillatorischen Verhaltens von (3.2.21) ist r jedoch nur in Ordnung μ^4 davon beeinflusst.

Berücksichtigung weiterer Glieder der Störungsrechnung (sowie von $\hat{z}(\tau) \neq 0$) verbessern die Übereinstimmung mit dem beobachteten Wert (3.2.1) noch weiter; z.B. in nächster Ordnung findet man

$$r = 1.08517. \quad (3.2.24)$$

4 Schwingungsprobleme

Eine weitere wichtige Klasse von mechanischen Systemen sind die Schwingungssysteme, die wir nun diskutieren wollen.

Lineare Bewegungsgleichungen entstehen oft durch Linearisierung einer nicht linearen Bewegungsgleichung um eine bekannte Lösung herum; ein Beispiel davon sahen wir bereits in Kapitel 3.1 und 3.2. Die allgemeinste lineare, inhomogene Differentialgleichung erster Ordnung für die Funktion $z(t) \in \mathbb{R}^n$ ist

$$\dot{z} = A(t)z + b(t), \quad (4.0.1)$$

wobei $b(t) \in \mathbb{R}^n$ und die lineare Abbildung (oder matrixwertige Funktion) $A(t) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ vorgegebene Funktionen von t sind. Es ist manchmal nützlich, auch komplexe Lösungen $t \mapsto z(t) \in \mathbb{C}^n$ zu betrachten: $\operatorname{Re} z(t)$ und $\operatorname{Im} z(t)$ sind dann reelle Lösungen.

Als Beispiel betrachten wir das mechanische System (siehe Kapitel 1)

$$m\ddot{x} = -fx - r\dot{x} + k(t), \quad (4.0.2)$$

das einen 1-dimensionalen Oszillator mit Reibungskraft und Anregung beschreibt. Es ist von der obigen Form (mit $n = 2$) falls wir es als 2-dimensionales System von Differentialgleichungen erster Ordnung schreiben:

$$z(t) = \begin{pmatrix} x \\ \dot{x}/\alpha \end{pmatrix}, \quad A = \begin{pmatrix} 0 & \alpha \\ -\alpha & -2\beta \end{pmatrix}, \quad b(t) = \begin{pmatrix} 0 \\ \gamma(t) \end{pmatrix}, \quad (4.0.3)$$

wobei

$$\alpha = \sqrt{\frac{f}{m}}, \quad \beta = \frac{r}{2m}, \quad \gamma(t) = \frac{k(t)}{\sqrt{fm}}. \quad (4.0.4)$$

[Die Skalierung mit α erlaubt es uns A als anti-symmetrische Matrix zu schreiben; das ist für das Folgende bequem.]

4.1 Lösungen der homogenen Gleichung

Die Lösungen der homogenen Gleichung

$$\dot{z} = A(t)z \quad (4.1.5)$$

nennt man freie Schwingungen. Sie bilden einen Vektorraum (Superpositionsprinzip). Der Raum dieser Lösungen hat die Dimension n . Dies folgt daraus, dass es n linear unabhängige Anfangsbedingungen gibt. Unter der Annahme, dass diese alle eine Lösung haben (physikalisch motiviert), folgt, dass es auch n linear unabhängige Lösungen gibt (da man aus den Lösungen umgekehrt die Anfangsbedingungen eindeutig bestimmen kann).

Die Lösungen können durch den sogenannten *Propagator* erzeugt werden; der Propagator $P(t, s)$ ist die Abbildung $\mathbb{R}^n \rightarrow \mathbb{R}^n$

$$P(t, s) : z(s) \mapsto z(t), \quad (4.1.6)$$

und ist durch die Eigenschaften

$$\begin{aligned} P(t, r)P(r, s) &= P(t, s), \\ P(s, s) &= 1, \end{aligned} \quad (4.1.7)$$

$$\frac{\partial}{\partial t} P(t, s) = A(t)P(t, s). \quad (4.1.8)$$

charakterisiert. Die letzte Gleichung garantiert, dass $P(t, s)$ zu jedem $z(s)$ eine Lösung $z(t)$ der Differentialgleichung zuordnet; die zweite Bedingung impliziert, dass $z(t)|_{t=s} = z(s)$ mit der vorgegebenen Anfangsbedingung übereinstimmt; die erste Eigenschaft ist eine Konsistenzbedingung.

Die letzten beiden Gleichungen sind äquivalent zur Integralgleichung

$$P(t, s) = 1 + \int_s^t dt_1 A(t_1)P(t_1, s), \quad (4.1.9)$$

deren Lösung durch eine konvergente Iterationsreihe dargestellt werden kann

$$P(t, s) = 1 + \sum_{n=1}^{\infty} \int_s^t dt_1 \int_s^{t_1} dt_2 \cdots \int_s^{t_{n-1}} dt_n A(t_1) \cdots A(t_n). \quad (4.1.10)$$

Das System (4.1.5) heisst *autonom*, falls A nicht von t abhängt. Da das System (4.1.5) dann invariant unter Zeit-Translation ist, hängt $P(t, s)$ nur von der Differenz $t - s$ ab, $P(t - t_0, s - t_0) = P(t, s)$. Für $P(t) \equiv P(t, 0)$ gilt dann

$$\begin{aligned} P(t)P(s) &= P(t + s), \quad P(0) = 1, \\ \frac{d}{dt} P(t) &= AP(t), \quad P(t) = e^{At} = \sum_{n=0}^{\infty} \frac{(At)^n}{n!}. \end{aligned} \quad (4.1.11)$$

Das Beispiel, das wir oben (4.0.3) behandelt haben, ist von diesem Typ (falls $k(t) = \gamma(t) = 0$). In diesem Fall gilt mit

$$\omega_0 := \sqrt{\alpha^2 - \beta^2} \quad (4.1.12)$$

$$P(t) = \begin{cases} e^{-\beta t} \left(1 \cos \omega_0 t + (A + \beta 1) \frac{1}{\omega_0} \sin \omega_0 t \right) & \text{falls } \omega_0 \neq 0 \\ e^{-\beta t} (1 + (A + \beta 1)t) & \text{falls } \omega_0 = 0. \end{cases} \quad (4.1.13)$$

Beweis. Wir definieren

$$C := (A + \beta 1) = \begin{pmatrix} \beta & \alpha \\ -\alpha & -\beta \end{pmatrix}$$

so dass,

$$C^2 = -\omega_0^2 1$$

gilt.

Für zwei beliebige Matrizen A und B die kommutieren ($[A, B] = AB - BA = 0$), gilt

$$e^A e^B = e^{A+B}.$$

Damit finden wir

$$\begin{aligned} e^{At} e^{\beta t} &= e^{At + \beta t} = e^{Ct} \\ &= 1 + Ct + \frac{1}{2!}(-\omega_0^2 1)t^2 + \frac{1}{3!}(-\omega_0^2 C)t^3 + \frac{1}{4!}(\omega_0^4 1)t^4 + \frac{1}{5!}(\omega_0^4 C)t^5 \\ &= 1 \cos \omega_0 t + C \frac{1}{\omega_0} \sin \omega_0 t, \end{aligned}$$

wobei wir im letzten Schritt die Taylor-Entwicklungen von

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} \pm \dots$$

und

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} \pm \dots$$

benutzt haben. Daraus folgt die Behauptung für $\omega_0 \neq 0$.

Im Fall $\omega_0 = 0$ gilt $\alpha = \beta$ und somit $C^2 = 0$

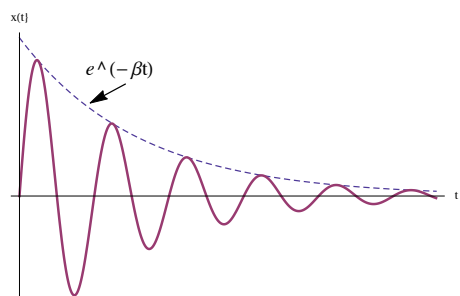
$$\Rightarrow e^{At} e^{\beta t} = e^{Ct} = 1 + Ct.$$

□

Ausgedrückt durch die ursprüngliche Variable $x(t)$ haben wir also

$$x(t) = x(0)e^{-\beta t} \left(\cos \omega_0 t + \frac{\beta}{\omega_0} \sin \omega_0 t \right) + \dot{x}(0)e^{-\beta t} \frac{1}{\omega_0} \sin \omega_0 t. \quad (4.1.14)$$

Wir wollen nun die physikalische Bedeutung von (4.1.14) verstehen. Dazu beobachten wir, dass α^2 proportional zur Federkonstante und β proportional zur Reibung ist. Wie wir im Folgenden sehen werden, ist für das Verhalten der Lösung (4.1.13) das Verhältnis zwischen α und β relevant, welches bestimmt, ob ω_0 imaginär oder reell ist.



Für $\alpha > \beta$ dominiert die Federkraft. In diesem Fall ist ω_0 reell und das System bleibt in Schwingung. Die Amplitude nimmt dabei ab.

Für $\alpha < \beta$ dominiert die Reibung. In diesem Fall ist ω_0 imaginär. Mit $\cosh(x) = \cos(ix)$ and $\sinh(x) = \sin(ix)$ ist $\cos(\omega_0 t) = \cosh(|\omega_0|t)$ and $\frac{1}{\omega_0} \sin(\omega_0 t) = \frac{1}{|\omega_0|} \sinh(|\omega_0|t)$. Die Lösungen von (4.1.14) sind dann von der Form $e^{-\beta t} e^{|\omega_0|t}$ und $e^{-\beta t} e^{-|\omega_0|t}$. Da $|\omega_0| < \beta$ ist der Exponent immer negativ falls $\beta \geq 0$.

Die Dämpfung ist maximal (d.h. das System kommt am schnellsten zum Stillstand) falls $\alpha = \beta$ und nicht, wie man vielleicht erwarten würde, falls β beliebig grösser als α ist. Falls die Reibung nämlich zu stark ist, bewegt sich das System nur langsam zum Nullpunkt. Im Fall $\alpha \approx \beta$ wird der Nullpunkt schnell erreicht und es überschießt diesen auch nicht.

Im autonomen Fall kann man die homogene Gleichung auch durch das Auffinden der *Eigenschwingungen* lösen. Die Eigenschwingungen sind die Lösungen der Form

$$z(t) = a e^{\lambda t}, \quad (\lambda \in \mathbb{C}, a \in \mathbb{C}^n). \quad (4.1.15)$$

Sie entsprechen den Lösungen des Eigenwertproblems

$$Aa = \lambda a. \quad (4.1.16)$$

Falls die Eigenvektoren von A ganz \mathbb{C}^n aufspannen, ist jede Lösung von (4.1.5) eine Linearkombination von Eigenschwingungen (4.1.15). A heisst dann diagonalisierbar, denn in einer Basis von Eigenvektoren ist A durch eine Diagonalmatrix

$$A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \quad (\lambda_1, \dots, \lambda_n : \text{Eigenwerte}) \quad (4.1.17)$$

dargestellt. Die Menge $\sigma(A) = \{\lambda_1, \dots, \lambda_n\}$ der Eigenwerte ist das *Spektrum* von A . Ohne Bezugnahme auf eine spezielle Basis kann man (4.1.17) als Spektralzerlegung von

A schreiben:

$$A = \sum_{\lambda \in \sigma(A)} \lambda P_\lambda. \quad (4.1.18)$$

Die lineare Abbildung $P_\lambda : \mathbb{C}^n \rightarrow \mathbb{C}^n$ ist die Eigenprojektion zum Eigenwert λ ; sie ist definiert durch

$$P_\lambda a = \begin{cases} a & \text{falls } Aa = \lambda a \\ 0 & \text{falls } Aa = \mu a \text{ mit } \mu \neq \lambda. \end{cases} \quad (4.1.19)$$

Die Eigenprojektoren sind paarweise orthogonal, und summieren sie sich zur Identität

$$P_\lambda P_\mu = \delta_{\lambda\mu} P_\lambda, \quad \sum_{\lambda \in \sigma(A)} P_\lambda = 1. \quad (4.1.20)$$

Weiterhin gilt

$$e^{At} = \sum_{\lambda \in \sigma(A)} e^{\lambda t} P_\lambda, \quad z(t) = \sum_{\lambda \in \sigma(A)} e^{\lambda t} P_\lambda z(0). \quad (4.1.21)$$

Für das Beispiel, das wir oben betrachtet haben, ist A (4.0.3) (generisch, d.h. für alle α und β) diagonalisierbar. Seine Eigenwerte sind

$$\lambda_{1,2} = -\beta \pm i\sqrt{\alpha^2 - \beta^2} \quad (4.1.22)$$

mit den zugehörigen Eigenvektoren $a_{1,2} = \begin{pmatrix} \alpha \\ \lambda_{1,2} \end{pmatrix}$, die linear unabhängig sind, solange $\lambda_1 \neq \lambda_2$, d.h. für $\omega_0 \neq 0$. Dann ist die freie Schwingung von der Form $x(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}$, und $c_{1,2}$ lassen sich durch $x(0)$, $\dot{x}(0)$ ausdrücken. So findet man wieder (4.1.14).

Das schwingende System (4.1.5) heisst *stabil*, falls keine Lösung $z(t)$ für $t \rightarrow +\infty$ unbeschränkt wächst. Falls A diagonalisierbar ist, ist das äquivalent zu der Bedingung

$$\operatorname{Re} \lambda \leq 0, \quad \text{alle } \lambda \in \sigma(A), \quad (4.1.23)$$

wie direkt aus (4.1.21) folgt.

Das System heisst *dissipativ*, falls es eine positiv definite quadratische Form (z, z) in \mathbb{R}^n gibt, so dass

$$\frac{d}{dt}(z, z) \leq 0 \quad (4.1.24)$$

für jede Lösung $z(t) \in \mathbb{R}^n$. [In diesem Fall gilt dann dasselbe für die entsprechende definierte quadratische Form auf \mathbb{C}^n]. Insbesondere ist das System dann stabil: $z(t)$ ist beschränkt, da $(z(t), z(t)) \leq (z(0), z(0))$ für $t \geq 0$. [Alternativ zeigt man, dass $\operatorname{Re} \lambda \leq 0$, da $d(z, z)/dt = (Az, z) + (z, Az) = 2(\operatorname{Re} \lambda)(z, z)$ für $z = ae^{\lambda t}$.]

Für das Beispiel des Oszillators ohne Anregung ist diese positiv definite quadratische Form zum Beispiel die Gesamtenergie

$$(z, z) = \frac{1}{2}(m\dot{x}^2 + fx^2), \quad z = \begin{pmatrix} x \\ \dot{x}/\alpha \end{pmatrix}. \quad (4.1.25)$$

Man rechnet leicht nach, dass (4.0.2) mit $k(t) = 0$ impliziert, dass

$$\frac{d}{dt}(z, z) = -r\dot{x}^2 \leq 0, \quad (4.1.26)$$

für $\beta = \frac{r}{2m} \geq 0$ (Reibung).

4.2 Erzwungene Schwingungen

Die allgemeine Lösung des inhomogenen Systems (4.0.1) ist

$$z(t) = P(t, s)z(s) + \int_s^t d\tau P(t, \tau)b(\tau) \quad (4.2.27)$$

(Duhamel-Formel). In der Tat rechnet man leicht nach, dass

$$\begin{aligned} \dot{z}(t) &= A(t)P(t, s)z(s) + P(t, t)b(t) + \int_s^t d\tau A(t)P(t, \tau)b(\tau) \\ &= A(t)z(t) + b(t), \end{aligned} \quad (4.2.28)$$

wobei wir die Eigenschaften (4.1.7) und (4.1.8) des Propagators (sowie den Fundamentalsatz der Differential- und Integralrechnung) benützt haben.

Lösungen für Differentialgleichungen erster Ordnung sind eindeutig (und damit (4.1.5)), falls gewisse Stetigkeitsbedingungen erfüllt sind. In der Praxis ist dies im Allgemeinen der Fall. Nicht-stetige Funktionen können durch stetige Funktionen approximiert werden. Zum Beispiel kann die Dirac-Delta Funktion, welche wir im Folgenden betrachten werden, durch eine Gauss-Funktion angenähert werden. Solche Approximationen sind physikalisch gerechtfertigt, weil auf der Grössenskala wo sie nicht mehr stimmen, das System gar nicht mehr durch die klassische Mechanik beschrieben werden kann, sondern die Quantenmechanik relevant ist. Für die mathematische Bahndlung kann man den Formalismus im nicht-physikalischen Bereich beliebig fortsetzen.

Bei autonomen Systemen vereinfacht sich (4.2.27) zu

$$z(t) = e^{At}z(0) + \int_0^t d\tau e^{A(t-\tau)}b(\tau). \quad (4.2.29)$$

Beispiel (Stossanregung). Für eine Stossanregung $b(t) = \delta(t - t_0)b$, die auf die Anfangsbedingung $z(s) = 0$ mit $(s < t_0)$ folgt, ergibt sich daraus die Stossantwort

$$z(t) = \theta(t - t_0)P(t, t_0)b, \quad (4.2.30)$$

($\theta(t) = 1$ für $t \geq 0$, und $\theta(t) = 0$ für $t < 0$).

Beispiel (Harmonische Anregung). Wir betrachten eine harmonische Anregung für ein autonomes System durch eine externe Kraft, die ebenfalls oszilliert

$$b(t) = b e^{i\omega t}, \quad (b \in \mathbb{C}^n, \omega \in \mathbb{R}). \quad (4.2.31)$$

1. Fall: $i\omega \notin \sigma(A)$

Man erhält aus (4.2.29) (oder direkt)⁸

$$\begin{aligned} z(t) &= e^{At} z(0) + \int_0^t d\tau e^{\tau(i\omega - A)} e^{At} b \\ &= e^{At} z(0) + (i\omega - A)^{-1} (e^{t(i\omega - A)} - 1) e^{At} b \\ &= e^{At} [z(0) - a(\omega)] + e^{i\omega t} a(\omega), \end{aligned} \quad (4.2.32)$$

wobei

$$a(\omega) = (i\omega - A)^{-1} b. \quad (4.2.33)$$

Die Inverse $(i\omega - A)^{-1}$ existiert da $i\omega \notin \sigma(A)$. Falls alle Eigenwerte von A einen negativen Realteil haben (also für gedämpfte Schwingungen), nimmt der erste Term exponentiell in t ab. Für grosse t bleibt also nur noch eine Schwingung mit Frequenz ω . Dies ist die erzwungene Schwingung

$$z(t) = e^{i\omega t} a(\omega) \quad (4.2.34)$$

mit der Anregungsfrequenz ω .

2. Fall: $i\omega \in \sigma(A)$

Falls $i\omega = \lambda_i$ für einen Eigenwert λ_i von A ist, verschwindet $i\omega - A$ in dem Eigenraum mit Projektor P_{λ_i} und die Inverse $(i\omega - A)^{-1}$ existiert in diesem Eigenraum nicht. Wir betrachten⁹

$$P_{\lambda_i} z(t) = P_{\lambda_i} z(0) + \underbrace{\int_0^t d\tau P_{\lambda_i} e^{At} b}_{\text{proportional in } t}.$$

In diesem Unterraum nimmt $z(t)$ also linear in t zu: Der Oszillator wird aufgeschaukelt. In der Praxis wird er natürlich nicht bis zu beliebig grossen Amplituden verstärkt, weil für grosse Auslenkungen die linearisierte Gleichung (4.0.2) nicht mehr stimmt.

Zusätzlich betrachten wir den Spezialfall von Fall 1, in dem $\exists \lambda \in \sigma(A)$ mit $\lambda \approx i\omega$ gilt.

Für diagonalisierbare A folgt aus (4.1.18)

$$(i\omega - A)^{-1} = \sum_{\lambda \in \sigma(A)} \frac{P_{\lambda}}{i\omega - \lambda}. \quad (4.2.35)$$

⁸Wir schreiben $e^{t(i\omega - A)}$ für $e^{t(i\omega 1 - A)}$.

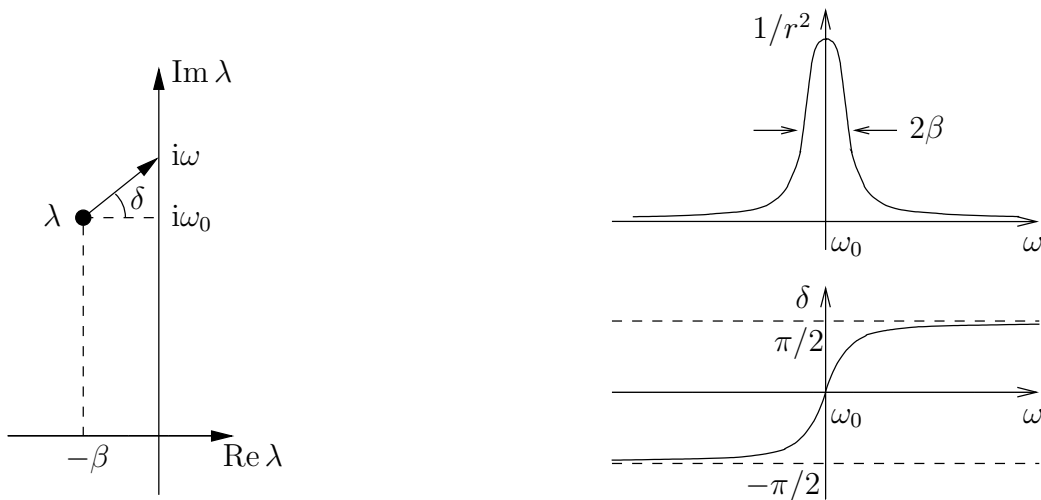
⁹Die Bedeutung von $P_{\lambda_i} z(t)$ folgt aus dem Superpositionsprinzip, das besagt, dass wir für jeden Eigenwert in dem dazugehörigen Unterraum eine separate Schwingung haben.

Liegt nun $i\omega$ nahe bei einem Eigenwert $\lambda \equiv -\beta + i\omega_0$, so ist der entsprechende Term in (4.2.33) dominant, also die erzwungene Schwingung

$$z(t) \cong \frac{e^{i\omega t}}{i\omega - \lambda} P_\lambda b = \frac{1}{r} e^{i(\omega t - \delta)} P_\lambda b, \quad (4.2.36)$$

wobei wir $i\omega - \lambda = \beta + i(\omega - \omega_0) \equiv r e^{i\delta}$ definiert haben (r, δ reell), also

$$\frac{1}{r^2} = \frac{1}{(\omega - \omega_0)^2 + \beta^2}, \quad \delta = \arctg \frac{\omega - \omega_0}{\beta}. \quad (4.2.37)$$



Die Amplitude $1/r$ und die Phase δ zeigen als Funktion von ω nahe bei ω_0 ein typisches Resonanzverhalten. Dabei ist die Breite der Resonanz ist proportional zur Dämpfung β der resonanten Eigenschwingung.

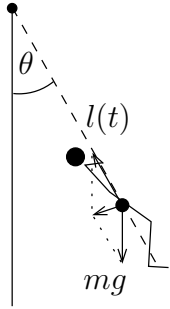
4.3 Parametrische Resonanz

Eine andere Art von Resonanz als die erzwungener Schwingungen zeigen frei schwingende Systeme mit einer periodischen Zeitabhängigkeit:

$$\dot{z} = A(t)z, \quad A(t + T) = A(t). \quad (4.3.38)$$

Solche Systeme kommen häufig durch Linearisierung um eine periodische Lösung eines nicht linearen Problems zustande (siehe z.B. (3.2.20)). Gleichung (4.3.38) beschreibt aber auch zum Beispiel das Schaukeln im Stehen: Der Abstand vom Aufhängepunkt zum Schwerpunkt wird periodisch längs dem Seil verändert, $l(t) = l(t + T)$. Der Drehimpulssatz sagt, dass die Zeitänderung des Drehimpulses gleich dem Drehmoment ist

$$\frac{d}{dt}(ml^2\dot{\theta}) = -mgl \sin \theta.$$



Dies ist eine Differentialgleichung zweiter Ordnung. Damit wir sie mit den im letzten Kapitel besprochenen Methoden lösen können, müssen wir sie als Differentialgleichung erster Ordnung schreiben. Für kleine Auslenkungen ($\sin \theta = \theta$) gilt

$$\frac{d}{dt}(l^2 \dot{\theta}) = l \frac{d^2}{dt^2}(l\theta) - l \ddot{l} \theta = -g l \theta$$

und daher

$$\frac{d^2}{dt^2}(l\theta) = -\frac{g - \ddot{l}}{l}(l\theta). \quad (4.3.39)$$

Dies lässt sich mit $x = l\theta$ schreiben als

$$\ddot{x} = -\omega^2(t)x, \quad \text{wobei} \quad \omega^2(t) = \frac{g - \ddot{l}}{l}, \quad \omega(t+T) = \omega(t). \quad (4.3.40)$$

Dies hat die Form (4.1.5) (ist also eine Differentialgleichung erster Ordnung) für

$$z = \begin{pmatrix} x \\ \dot{x} \end{pmatrix}, \quad A(t) = \begin{pmatrix} 0 & 1 \\ -\omega^2(t) & 0 \end{pmatrix}. \quad (4.3.41)$$

Es sei $P(t)$ der Propagator von (4.3.38) für das Zeitintervall von 0 nach t . Die Periodizität von A impliziert dann sofort, dass

$$P(nT) = P(T)^n, \quad (n = 1, 2, 3, \dots). \quad (4.3.42)$$

Exponentiell wachsende Lösungen treten also dann auf, wenn $P(T)$ einen Eigenwert $\lambda \in \mathbb{C}$ hat mit $|\lambda| > 1$; umgekehrt sind alle Lösungen beschränkt, falls $P(T)$ diagonalisierbar ist mit Eigenwerten $|\lambda_k| \leq 1$. Die Eigenwerte selber sind schwierig zu bestimmen. Was wir berechnen können, ist die Determinante, welche uns ebenfalls zu entscheiden erlaubt, ob die Lösungen unbeschränkt wachsen.

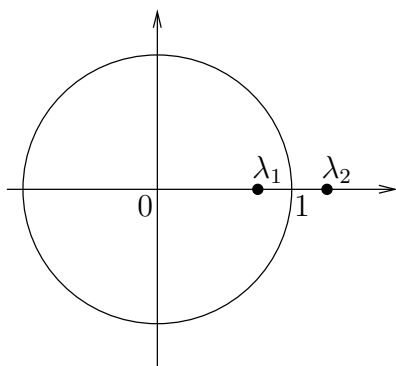
Da die Spur von A verschwindet, $\text{sp}A(t) = 0$, folgt [zu niedrigster Ordnung in δP gilt $\det(P + \delta P) = \det(1 + \delta P P^{-1}) \det(P) = \det(P) + \text{sp}(\delta P P^{-1}) \det(P)$] gilt

$$\frac{d}{dt} \det P(t) = \frac{\det(P + \delta P) - \det P}{\delta t} = \text{sp} \left(\underbrace{\frac{dP}{dt} P(t)^{-1}}_{A(t)} \right) \det P(t) = 0, \quad (4.3.43)$$

und daher $\det P(t) = \det P(0) = 1$.

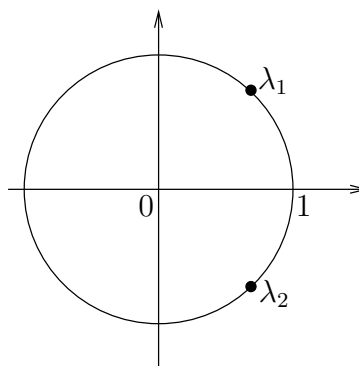
Die beiden Eigenwerte λ_1, λ_2 erfüllen also $\lambda_1 \lambda_2 = 1$. Weiterhin ist $P(t)$ reell, und falls λ_i ein Eigenwert ist, so ist auch $\bar{\lambda}_i$ ein Eigenwert. Es gibt daher also zwei Möglichkeiten (siehe Skizze). Im ersten Fall ist das System instabil ($|\lambda_2| > 1$), im zweiten hingegen stabil ($|\lambda_1| = |\lambda_2| = 1$). [Der Grenzfall $\lambda_1 = \lambda_2 = 1$ ist stabil.]

- $\lambda_i = \bar{\lambda}_i$: λ_i reell



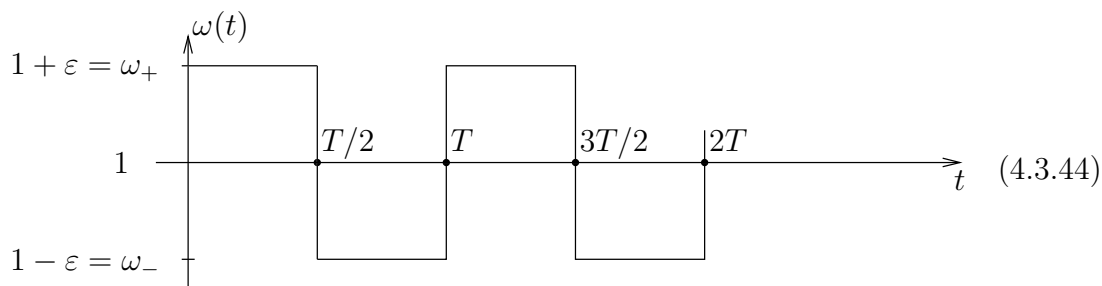
$$|\text{sp}P(T)| = |\lambda_1 + \lambda_2| > 2^{10}$$

- $\lambda_1 = \bar{\lambda}_2$: $\lambda_i \bar{\lambda}_i = 1$



$$|\text{sp}P(T)| = |\lambda_1 + \lambda_2| \leq 2$$

Als Beispiel für eine Anregung wählen wir nun $\omega(t)$ als Stufenfunktion. Die Frequenz ω



variiert also um den Wert $\omega = 1$. Für konstantes $\omega(t) \equiv \omega$ wäre die Lösung von (4.3.40) einfach durch den Propagator $e^{At} = p(t; \omega)$ gegeben. Indem wir wie im vorherigen Kapitel die Entwicklung von $\sin x$ und $\cos x$ betrachten, finden wir

$$p(t; \omega) = \begin{pmatrix} \cos \omega t & \omega^{-1} \sin \omega t \\ -\omega \sin \omega t & \cos \omega t \end{pmatrix}. \quad (4.3.45)$$

Insbesondere wäre für $\omega(t) \equiv 1$ die Periode des Oszillators einfach $T_0 = 2\pi$. Im Beispiel (4.3.44) ist deshalb

$$\begin{aligned} P(T) &= P\left(T, \frac{T}{2}\right)P\left(\frac{T}{2}, 0\right) = p\left(\frac{T}{2}; \omega_-\right) p\left(\frac{T}{2}; \omega_+\right) \\ &= \begin{pmatrix} \cos \omega_- T/2 & \omega_-^{-1} \sin \omega_- T/2 \\ -\omega_- \sin \omega_- T/2 & \cos \omega_- T/2 \end{pmatrix} \begin{pmatrix} \cos \omega_+ T/2 & \omega_+^{-1} \sin \omega_+ T/2 \\ -\omega_+ \sin \omega_+ T/2 & \cos \omega_+ T/2 \end{pmatrix}. \end{aligned} \quad (4.3.46)$$

¹⁰Dies folgt aus $\lambda_1 \lambda_2 = 1$.

Daher gilt

$$\begin{aligned}
 \operatorname{sp}P(T) &= \cos \frac{\omega_+ T}{2} \cos \frac{\omega_- T}{2} - \frac{\omega_+}{\omega_-} \sin \frac{\omega_+ T}{2} \sin \frac{\omega_- T}{2} \\
 &\quad + \cos \frac{\omega_+ T}{2} \cos \frac{\omega_- T}{2} - \frac{\omega_-}{\omega_+} \sin \frac{\omega_+ T}{2} \sin \frac{\omega_- T}{2} \\
 &= (\cos T + \cos \varepsilon T) + \frac{1}{2} \underbrace{\left(\frac{\omega_+}{\omega_-} + \frac{\omega_-}{\omega_+} \right)}_{\frac{1+\varepsilon^2}{1-\varepsilon^2}} (\cos T - \cos \varepsilon T) \\
 &= \frac{2}{1-\varepsilon^2} (\cos T - \varepsilon^2 \cos \varepsilon T) \\
 &= 2 - \frac{4}{1-\varepsilon^2} \left(\sin^2 \frac{T}{2} - \varepsilon^2 \sin^2 \frac{\varepsilon T}{2} \right) \tag{4.3.47}
 \end{aligned}$$

$$= -2 + \frac{4}{1-\varepsilon^2} \left(\cos^2 \frac{T}{2} - \varepsilon^2 \cos^2 \frac{\varepsilon T}{2} \right). \tag{4.3.48}$$

[In der zweiten Gleichung haben wir die Identitäten $\cos(a+b) + \cos(a-b) = 2 \cos a \cos b$ und $\cos(a+b) - \cos(a-b) = -2 \sin a \sin b$ benützt; in der vierten bzw. fünften die Identität $\cos(2a) = 1 - 2 \sin^2 a$ und $\cos(2a) = 2 \cos^2 a - 1$.]

Wir fassen nun T als variabel auf und suchen zu gegebenem ε die T -Bereiche, in denen das System instabil ist (Resonanzbereiche). Wie zuvor besprochen ist das System instabil falls $|\operatorname{sp}P(T)| > 2$. Für $\varepsilon = 0$ ist $|\operatorname{sp}P(T)| \in [-2, 2]$ und es gibt keine Resonanz. Allgemein ergeben sich zwei Möglichkeiten:

$$\begin{aligned}
 \text{(a)} \quad \operatorname{sp}P(T) > 2 : \quad & \sin^2 \frac{T}{2} < \varepsilon^2 \sin^2 \frac{\varepsilon T}{2}, \\
 \text{(b)} \quad \operatorname{sp}P(T) < -2 : \quad & \cos^2 \frac{T}{2} < \varepsilon^2 \cos^2 \frac{\varepsilon T}{2}.
 \end{aligned} \tag{4.3.49}$$

Wir diskutieren die beiden Fälle für $\varepsilon \ll 1$:

(a) Für $\varepsilon = 0$ ist $\operatorname{sp}P(T) = 2$, falls $T = 2\pi n$, ($n = 1, 2, 3, \dots$). Für kleine ε suchen wir also Lösungen von (4.3.49) der Form $T = 2\pi n + x$ mit $|x| \ll 1$. Näherungsweise ist die Bedingung dann

$$\underbrace{\sin^2 \frac{x}{2}}_{\approx \left(\frac{x}{2}\right)^2} < \varepsilon^2 \underbrace{\sin^2 \frac{\varepsilon(2n\pi + x)}{2}}_{\approx (\varepsilon n\pi)^2}$$

$$\left| \frac{x}{2} \right| < \varepsilon^2 n\pi \quad \Leftrightarrow \quad |x| < 2\varepsilon^2 n\pi$$

In diesem Fall gibt es Resonanz auch falls wir mit einem Vielfachen der Eigenfrequenz anregen. Ausserdem ist das System umso toleranter, je schneller die Anregung ist (die Länge der Resonanzintervalle bei festem ε wächst proportional zu n und verschwindet $\sim \varepsilon^2$ für $\varepsilon \rightarrow 0$).

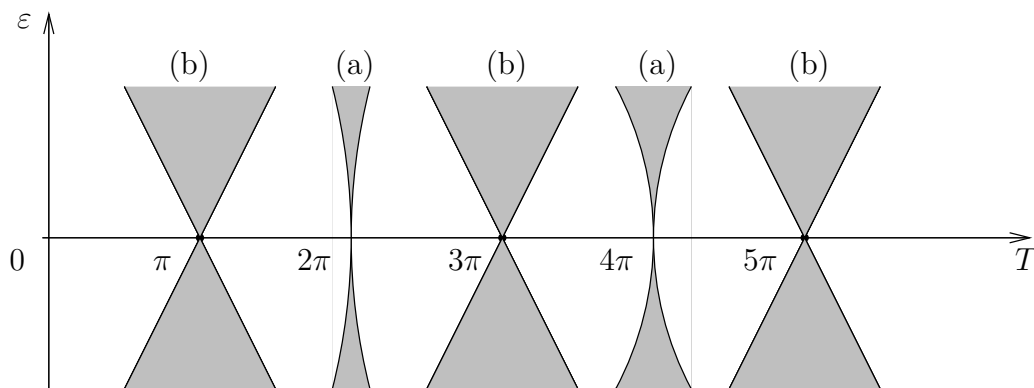
(b) Für $\varepsilon = 0$ ist $\text{sp}P(T) = -2$ falls $T = (2n + 1)\pi$, ($n = 0, 1, 2, \dots$). Für kleine ε setzen wir $T = (2n + 1)\pi + x$, und (4.3.49) lautet dann

$$\sin^2 \frac{x}{2} < \varepsilon^2 \underbrace{\cos^2 \frac{\varepsilon(2n + 1)\pi + \varepsilon x}{2}}_{\approx 1}$$

$$\left| \frac{x}{2} \right|^2 < \varepsilon^2 \quad \Leftrightarrow \quad |x| < 2\varepsilon$$

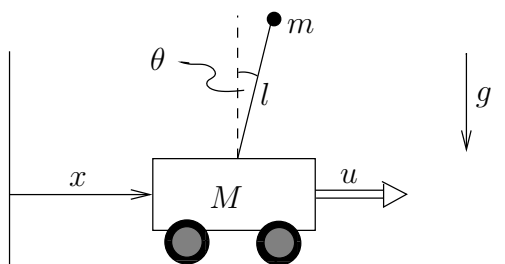
In diesem Fall ist die Länge der Resonanzintervalle unabhängig von n und verschwindet $\sim \varepsilon$ für $\varepsilon \rightarrow 0$.

In beliebigen Einheiten ist $T/2\pi$ das Verhältnis der Perioden der Störung und der freien Schwingung.



4.4 Stabilisierung linearer Systeme

Als Beispiel betrachten wir das umgekehrte Pendel



Koordinaten (x, θ)
horizontale Kraft $u(t)$

Wir werden sehen, dass es möglich ist einen Stab zu stabilisieren. Für einen zweiten Stab ist dies dagegen nicht möglich. Entscheidend ist dabei, wieviele Freiheitsgrade man zum stabilisieren hat (in diesem Fall einen u) und wieviele man stabilisieren muss.

Die linearisierten Bewegungsgleichungen (θ und $\dot{\theta}$ klein) des Systems sind [wir werden diese in Kapitel 5.4.1 im Euler-Lagrange-Formalismus einfach herleiten können]

$$(1) \quad (M + m)\ddot{x} + ml\ddot{\theta} = u$$

und

$$(2) \quad m\ddot{x} + ml\ddot{\theta} = mg\theta.$$

(1) – (2) liefert

$$M\ddot{x} = u - ml\theta$$

und $(1)m - (2)(M + m)$ gibt

$$Ml\ddot{x} = (M + m)l\theta - u.$$

Damit kann das System in folgender allgemeinen Form geschrieben werden

$$\dot{z} = Az + bu(t) \tag{4.4.1}$$

($A : n \times n$ Matrix, $b \in \mathbb{R}^n$), und zwar mit $n = 4$,

$$z = \begin{pmatrix} l\theta \\ l\dot{\theta} \\ x \\ \dot{x} \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 + m & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -m & 0 & 0 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}. \tag{4.4.2}$$

Hier haben wir Einheiten gewählt, in denen $M = 1$ und $g = l$.

Falls die Kraft u von den Koordinaten (und Ableitungen) abhängt, spricht man von Rückkopplung. Wir betrachten den Fall, wo u eine lineare Funktion der Komponenten z ist,¹¹ d.h.

$$u = r^T z, \quad r \in \mathbb{R}^n. \tag{4.4.3}$$

Falls $r \in \mathbb{R}^n$ so gewählt werden kann, dass das resultierende (homogene) System mit Matrix

$$A + br^T \tag{4.4.4}$$

stabil ist, so nennen wir (4.4.1) stabilisierbar.

Satz. Das System (4.4.4) ist stabilisierbar, falls

$$b, Ab, \dots, A^{n-1}b$$

den ganzen \mathbb{R}^n aufspannen.

¹¹Das ist gerechtfertigt, weil wir das System stabilisieren, wenn es noch nicht zu stark gestört ist.

Diese Bedingung verlangt, dass alle Freiheitsgrade direkt (b) oder indirekt ($Ab, \dots, A^{n-1}b$) durch die Rückkopplung r beeinflusst werden. Höhere Potenzen $A^k b$, $k \geq n$ spielen übrigens keine Rolle: wegen des Satzes von Cayley-Hamilton ist $p_n(A) = 0$, wobei p_n das charakteristische Polynom von A ist

$$p_n(\lambda) = \det(\lambda - A) \equiv \lambda^n + \alpha_1 \lambda^{n-1} + \dots + \alpha_{n-1} \lambda + \alpha_n. \quad (4.4.5)$$

Die höheren Potenzen $A^k b$ sind daher Linearkombinationen der ersten $n - 1$ Potenzen.

Beweis. Um die obige Behauptung zu beweisen, definieren wir rekursiv $p_k(\lambda)$, $k = n - 1, n - 2, \dots, 0$ durch

$$p_k(\lambda) = \lambda p_{k-1}(\lambda) + \alpha_k, \quad (k = 1, \dots, n), \quad (4.4.6)$$

wobei $p_0(\lambda) = 1$. Unter der Annahme des Satzes ist dann auch

$$e_0 = b = p_0(A)b, \quad e_1 = p_1(A)b, \quad \dots, \quad e_{n-1} = p_{n-1}(A)b$$

eine Basis des \mathbb{R}^n . Die obige Rekursionsrelation für p_k impliziert dann

$$A e_k = A p_k(A)b = p_{k+1}(A)b - \alpha_{k+1} b = e_{k+1} - \alpha_{k+1} e_0, \quad (k = 0, \dots, n - 1). \quad (4.4.7)$$

Bezüglich dieser Basis hat daher A und b die Komponenten

$$A = \begin{pmatrix} -\alpha_1 & -\alpha_2 & \cdots & -\alpha_n \\ 1 & 0 & & \vdots \\ 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & & & 1 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad (4.4.8)$$

wobei wir wiederum $p_n(A) = 0$ benützt haben. Für jede Wahl von r ist nun

$$b r^T = \begin{pmatrix} r_1 & \cdots & r_n \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & & 0 \end{pmatrix}. \quad (4.4.9)$$

Die Matrix $A + b r^T$, welche dem zu stabilisierenden System entspricht, hat genau dieselbe Struktur wie A , abgesehen davon, dass $\alpha_k \mapsto \alpha_k - r_k$. Das charakteristische Polynom von $A + b r^T$ ist.

$$p(\lambda; r) = \lambda^n + (\alpha_1 - r_1) \lambda^{n-1} + \dots + (\alpha_n - r_n). \quad (4.4.10)$$

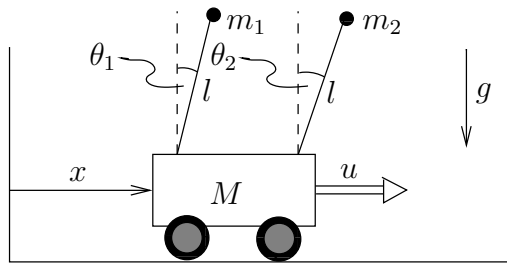
Das System ist stabil, wenn es nur noch beschränkte Lösungen hat, wenn also die Eigenwerte (die Lösungen von $p(\lambda; r) = 0$) von $A + b r^T$ alle einen negativen Realteil haben. Da r_1, \dots, r_n beliebig gewählt werden können, kann diese Bedingung immer erfüllt werden.

Im obigen Beispiel ist

$$\det(b, Ab, A^2b, A^3b) = \begin{vmatrix} 0 & -1 & 0 & -(1+m) \\ -1 & 0 & -(1+m) & 0 \\ 0 & 1 & 0 & m \\ 1 & 0 & m & 0 \end{vmatrix} = 1 \neq 0, \quad (4.4.11)$$

und das System ist somit stabilisierbar. \square

Zum Abschluss betrachten wir noch das Beispiel zweier umgekehrter Pendel gleicher Länge



Die linearisierten Bewegungsgleichungen sind nun

$$\begin{aligned} (M + m_1 + m_2)\ddot{x} + m_1l\ddot{\theta}_1 + m_2l\ddot{\theta}_2 &= u \\ m_1\ddot{x} + m_1l\ddot{\theta}_1 &= m_1g\theta_1 \\ m_2\ddot{x} + m_2l\ddot{\theta}_2 &= m_2g\theta_2. \end{aligned}$$

Die letzten beiden Gleichungen implizieren, dass

$$l(\ddot{\theta}_1 - \ddot{\theta}_2) = g(\theta_1 - \theta_2), \quad (4.4.12)$$

und daher wird die Bewegung von $\theta_1 - \theta_2$ nicht stabilisiert!

Das kann auch im Rahmen der obigen Analyse verstanden werden: benützt man die Kombinationen $z_1 = l(\theta_1 - \theta_2)$, $z_2 = l(\dot{\theta}_1 - \dot{\theta}_2)$ als zwei der sechs Komponenten von z , so ist

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ g/l & 0 & 0 & 0 & 0 & 0 \\ \vdots & & & & & \vdots \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \\ \vdots \end{pmatrix}, \quad (4.4.13)$$

und daher verschwinden die ersten beiden Komponenten aller Vektoren $A^k b$, ($k = 1, \dots, 5$). Die Bedingung des obigen Satzes ist daher nicht erfüllt und das System damit nicht stabilisierbar.

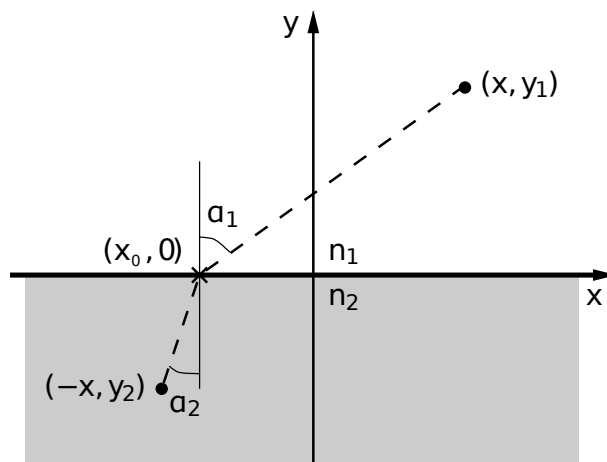
5 Lagrange Formulierung

Der Lagrange-Formalismus ist eine Formulierung der klassischen Mechanik, in der die Dynamik eines Systems durch eine einzige skalare Funktion, die *Lagrange-Funktion*, beschrieben wird. Im Gegensatz zu den Newton'schen Bewegungsgleichungen ist sie unabhängig von den Koordinaten und auch in beschleunigten Bezugssystemen gültig.

Weiter hat der Formalismus den Vorteil, dass sich Zwangsbedingungen einfach durch eine geeignete Wahl der Koordinaten integrieren lassen. Die Lagrange-Formulierung folgt aus dem *Prinzip der kleinsten Wirkung*. Dieses Prinzip postuliert, dass die Teilchenbahnen eines Systems gerade so sind, dass die Wirkung minimiert ist.

Motiviert ist das Prinzip unter anderem durch Beobachtungen aus der Optik. Zum Beispiel kann man das Reflexionsgesetz und Brechungsgesetz als Extremalprinzip formulieren: Der vom Licht eingeschlagene Weg ist derjenige, bei dem die Zeit um vom Start- zum Endpunkt zu gelangen extremal ist (*Fermat'sches Prinzip*).

Betrachte zum Beispiel die Konfiguration zweier optischer Medien mit Brechungsindizes n_1 und n_2 . Die Koordinaten der beiden Punkte in der $x - y$ -Ebene seien (nach geeigneter Verschiebung des Ursprungs) (x, y_1) und $(-x, y_2)$, wobei $y_1 > 0$ und $y_2 < 0$. (Beide Punkte haben dieselbe z -Komponente, die wir deshalb unterdrücken. Die Trennebene zwischen den beiden Medien ist die $x - z$ -Ebene $y = 0$, siehe Diagramm.)



Um den Weg kürzester Zeit zu suchen, können wir ohne Beschränkung der Allgemeinheit annehmen, dass das Licht auf gerader Linie in Medium 1 von (x, y_1) zu dem Punkt auf der Trennebene $(x_0, 0)$ fliegt (im Diagramm mit \times gekennzeichnet), und von dort in Medium 2 zum Punkt $(-x, y_2)$. In einem Medium mit Brechungsindex n hat das Licht die Geschwindigkeit c/n , wobei c die Lichtgeschwindigkeit ist. Die Zeit T , die das Licht auf einer dieser Probekurven benötigt, ist deshalb

$$cT = n_1 \sqrt{(x - x_0)^2 + y_1^2} + n_2 \sqrt{(x + x_0)^2 + y_2^2}. \quad (5.0.1)$$

Wir suchen den Punkt x_0 , für den diese Zeit minimal wird. Dazu betrachten wir die

Ableitung

$$\frac{d}{dx_0} cT = -\frac{n_1(x-x_0)}{\sqrt{(x-x_0)^2 + y_1^2}} + \frac{n_2(x_0+x)}{\sqrt{(x+x_0)^2 + y_2^2}} = 0. \quad (5.0.2)$$

Wenn wir mit α_i den Einfallswinkel des Strahls relativ zur Normalen bezeichnen, dann ist

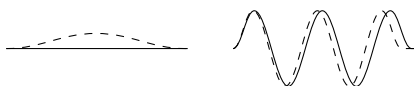
$$\sin \alpha_1 = \frac{(x-x_0)}{\sqrt{(x-x_0)^2 + y_1^2}}, \quad \sin \alpha_2 = \frac{(x+x_0)}{\sqrt{(x+x_0)^2 + y_2^2}}, \quad (5.0.3)$$

und daher ist die Bedingung (5.0.2) gerade das Snellius'sche Brechungsgesetz,

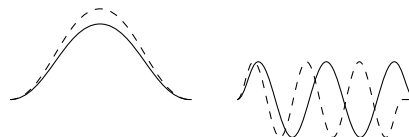
$$n_1 \sin \alpha_1 = n_2 \sin \alpha_2. \quad (5.0.4)$$

Dass sich Licht innerhalb eines Mediums geradlinig bewegt, lässt sich mit Hilfe der Quantenmechanik verstehen: Das Teilchen nimmt tatsächlich alle Wege. Mit dem Welleilchen-Dualismus folgt, dass sich das Teilchen auch als Welle mit einer bestimmten Wellenlänge beschreiben lässt. Betrachten wir nun verschiedene Teilchenbahnen, welche nur leicht von einem geraden Weg abweichen, so ändert sich die effektive Weglänge nur leicht und die Wellen interferieren konstruktiv. Im Gegensatz dazu unterscheiden sich die Pfadlängen in einer Umgebung eines gekrümmten Weges stärker und es kommt zu destruktiver Interferenz.

Pfade in der Umgebung eines geraden Weges führen zu konstruktiver Interferenz.



Pfade in der Umgebung eines gekrümmten Weges führen zu destruktiver Interferenz.



5.1 Das Brachistochronenproblem

Ein ein wenig komplizierteres Problem ist das sogenannte Brachistochronenproblem: bestimme diejenige Kurve in einer vertikalen Ebene, in der ein reibungsfreier Körper unter dem Einfluss eines homogenen Schwerfeldes in der kürzesten Zeit von einem vorgegebenen Punkt zu einem anderen gelangt. Wie beschreiben die Kurve C durch eine Funktion

$$x = f(z), \quad (5.1.1)$$

wobei x die horizontale und z die vertikale Koordinate (nach unten gerichtet) ist. Die Randbedingungen seien

$$f(0) = 0, \quad f(a) = b, \quad (5.1.2)$$

d.h. die beiden Punkte haben die Koordinaten $(0, 0)$ und (b, a) in der $x - z$ -Ebene. Wir nehmen an, dass $b > 0$, und dass der Körper an der Stelle $x = z = 0$ in Ruhe beginnt. Wegen der Energieerhaltung ist der Betrag seiner Geschwindigkeit in der Höhe z dann

$$v = \sqrt{2gz}. \quad \left[\frac{1}{2}mv^2 = mgz. \right] \quad (5.1.3)$$

Zum Durchlaufen der Kurve benötigt er daher die Zeit

$$T(f) = \int_C \frac{ds}{v} = \int_0^a dz \sqrt{\frac{1 + f'(z)^2}{2gz}}. \quad (5.1.4)$$

[Hier ist $ds^2 = dx^2 + dz^2 = (1 + f'(z)^2)dz^2$.] Man muss jetzt diejenige *Funktion* $f(z)$ finden, für die diese Zeit minimal ist. Dies ist ein Extremalproblem in dem unendlich dimensionalen Raum der (genügend oft) differenzierbaren Funktionen. Es lässt sich in der folgenden Weise auf ein Extremalproblem in einer Dimension zurückführen.

Sei f die Lösung des Extremalproblems, und sei h eine beliebige unendlich oft differenzierbare Funktion mit $h(0) = h(a) = 0$. Dann wird durch $f + \alpha h$ für jedes $\alpha \in \mathbb{R}$ eine andere Kurve beschrieben, die auch die richtigen Randbedingungen erfüllt. Da nach Voraussetzung $T(f)$ minimal ist, muss $T(f + \alpha h)$ grösser oder gleich $T(f)$ sein. Insbesondere muss daher die Ableitung von $T(f + \alpha h)$ nach α verschwinden,

$$0 = \left. \frac{d}{d\alpha} T(f + \alpha h) \right|_{\alpha=0}. \quad (5.1.5)$$

Nach Vertauschen der Integration und Differentiation ergibt sich daraus

$$\begin{aligned} 0 &= \int_0^a dz \left. \frac{\partial}{\partial y} \sqrt{\frac{1 + y^2}{2gz}} \right|_{y=f'(z)} h'(z) \\ &= \int_0^a dz \frac{f'(z)}{\sqrt{2gz(1 + f'(z)^2)}} h'(z). \end{aligned} \quad (5.1.6)$$

Schliesslich integrieren wir partiell und erhalten dann

$$0 = \left. \frac{f'(z)}{\sqrt{2gz(1 + f'(z)^2)}} h(z) \right|_{z=0}^{z=a} - \int_0^a dz \left(\frac{d}{dz} \frac{f'(z)}{\sqrt{2gz(1 + f'(z)^2)}} \right) h(z). \quad (5.1.7)$$

Die Randterme verschwinden nach Voraussetzung an die Funktion $h(z)$. Da $h(z)$ ansonsten beliebig ist, kann das Integral nur dann verschwinden, falls der Integrand verschwindet, d.h.

$$\frac{d}{dz} \frac{f'(z)}{\sqrt{2gz(1 + f'(z)^2)}} = 0. \quad (5.1.8)$$

Damit haben wir eine Differentialgleichung abgeleitet, die von der Lösung des Brachistochronenproblems erfüllt werden muss!

Um die Funktion f tatsächlich zu bestimmen, definieren wir jetzt

$$c = \frac{f'(z)}{\sqrt{z(1+f'(z)^2)}}. \quad (5.1.9)$$

Die obige Gleichung besagt, dass c konstant ist. Falls $c = 0$, dann gilt $f'(z) = 0$, d.h. die Kurve ist eine vertikale Gerade. [Dies ist natürlich nur dann die richtige Lösung, falls $b = 0$.] Sei nun $c > 0$, was $b = f(a) > 0$ entspricht. Auflösen der obigen Gleichung nach $f'(z)$ ergibt

$$f'(z) = \sqrt{\frac{c^2 z}{1 - c^2 z}}, \quad \text{wobei } z < \frac{1}{c^2}. \quad (5.1.10)$$

Wir substituieren jetzt

$$z = \frac{1}{c^2} \sin^2 \frac{\varphi}{2} = \frac{1}{2c^2} (1 - \cos \varphi), \quad \varphi \in [0, \pi]. \quad (5.1.11)$$

so dass

$$c^2 z = \sin^2 \frac{\varphi}{2}, \quad 1 - c^2 z = \cos^2 \frac{\varphi}{2}. \quad (5.1.12)$$

Weiterhin ist

$$dz = \frac{1}{c^2} \sin \frac{\varphi}{2} \cos \frac{\varphi}{2} d\varphi, \quad (5.1.13)$$

und daher

$$\begin{aligned} x(\varphi) &= f(z(\varphi)) = \int_0^{z(\varphi)} dz' f'(z') = \frac{1}{c^2} \int_0^\varphi d\varphi' \sin \frac{\varphi'}{2} \cos \frac{\varphi'}{2} \frac{\sin \frac{\varphi'}{2}}{\cos \frac{\varphi'}{2}} \\ &= \frac{1}{c^2} \int_0^\varphi d\varphi' \sin^2 \frac{\varphi'}{2} = \frac{1}{2c^2} \int_0^\varphi d\varphi' (1 - \cos \varphi') = \frac{1}{2c^2} (\varphi - \sin \varphi). \end{aligned} \quad (5.1.14)$$

Die gesuchte Kurve schreibt sich mit Hilfe des Parameters φ also als

$$\begin{aligned} x(\varphi) &= \frac{1}{2c^2} (\varphi - \sin \varphi) \\ z(\varphi) &= \frac{1}{2c^2} (1 - \cos \varphi). \end{aligned} \quad (5.1.15)$$

Dies ist gerade die *Zykloide*, also die Bahn eines Punktes auf der Peripherie eines Kreises mit Radius $\frac{1}{2c^2}$, der auf der Geraden $z = 0$ abrollt. [Dies folgt einfach daraus, dass die obige Kurve die Gleichung

$$(x(\varphi) - R\varphi)^2 + (z(\varphi) - R)^2 = R^2, \quad R = \frac{1}{2c^2} \quad (5.1.16)$$

erfüllt.]

Schliesslich bestimmt man die Konstante c aus der Bedingung, dass es ein $\varphi \in [0, \pi]$ gibt, so dass

$$b = \frac{1}{2c^2} (\varphi - \sin \varphi), \quad a = \frac{1}{2c^2} (1 - \cos \varphi). \quad (5.1.17)$$

5.2 Die Euler-Lagrange Gleichungen

Das Brachistochronenproblem, das von den schweizer Mathematikern Johann und Jakob Bernoulli 1697 gelöst wurde, war historisch der Ausgangspunkt für eine allgemeine Untersuchung von Extremalproblemen für Funktionale S der Form

$$S(f) = \int_{x_1}^{x_2} dx s(f(x), f'(x), x), \quad (5.2.1)$$

wobei s eine zweimal stetig differenzierbare Funktion von drei reellen Variablen, und f eine zweimal stetig differenzierbare Funktion auf dem Intervall $[x_1, x_2]$ ist, deren Werte bei x_1 und x_2 vorgegeben sind. Mit denselben Überlegungen wie im Kapitel 5.1 können wir eine Bedingung an die Funktion f , die das Funktional $S(f)$ extremal macht, bestimmen: sei f eine solche Funktion, dann muss gelten

$$\left. \frac{d}{d\alpha} S(f + \alpha h) \right|_{\alpha=0} = 0, \quad (5.2.2)$$

wobei h eine beliebige Funktion mit $h(x_1) = h(x_2) = 0$ ist.¹² Mit der Notation

$$\frac{\partial s}{\partial f}(f(x), f'(x), x) \equiv \left. \frac{\partial s}{\partial y}(y, z, x) \right|_{(y,z,x)=(f(x),f'(x),x)}, \quad (5.2.3)$$

$$\frac{\partial s}{\partial f'}(f(x), f'(x), x) \equiv \left. \frac{\partial s}{\partial z}(y, z, x) \right|_{(y,z,x)=(f(x),f'(x),x)}, \quad (5.2.4)$$

erhält man aus (5.2.2)

$$\begin{aligned} 0 &= \int_{x_1}^{x_2} dx \left(\frac{\partial s}{\partial f}(f(x), f'(x), x) h(x) + \frac{\partial s}{\partial f'}(f(x), f'(x), x) h'(x) \right) \\ &= \int_{x_1}^{x_2} dx \left(\frac{\partial s}{\partial f}(f(x), f'(x), x) - \frac{d}{dx} \left[\frac{\partial s}{\partial f'}(f(x), f'(x), x) \right] \right) h(x), \end{aligned} \quad (5.2.5)$$

wobei wir in der letzten Gleichung wiederum partiell integriert haben. (Die Randterme tragen nicht bei, da $h(x_1) = h(x_2) = 0$.) Da diese Gleichung für beliebige $h(x)$ gelten muss, folgt daraus, dass eine Funktion f , die das Extremalproblem löst, die Differentialgleichung

$$\frac{\partial s}{\partial f}(f(x), f'(x), x) - \frac{d}{dx} \left[\frac{\partial s}{\partial f'}(f(x), f'(x), x) \right] = 0 \quad (5.2.6)$$

erfüllt. Diese Gleichung wird *Euler-Lagrange Gleichung* genannt. Falls s von mehreren Funktionen $f_1(x), \dots, f_n(x)$ sowie ihrer Ableitungen abhängt,

$$s \equiv s(f_1(x), f_1'(x), \dots, f_n(x), f_n'(x); x),$$

so bestehen die Euler-Lagrange Gleichungen aus n verschiedenen Gleichungen der Form (5.2.6), je einer für jedes $f_i(x)$.

¹²Da der Funktionenraum ein Vektorraum ist, kann jede Funktion in der Form $f + \alpha h$ geschrieben werden, also genügt es Funktionen dieser Form zu betrachten.

5.3 Beispiele zur Variationsrechnung

Bevor wir besprechen, wie sich die Gesetze der Mechanik als Extremalprinzip beschreiben lassen, wollen wir noch ein paar Beispiele zur Variationsrechnung besprechen, um uns damit besser vertraut zu machen. Insbesondere wollen wir auch die Methode der Lagrange'schen Multiplikatoren dabei erklären.

5.3.1 Kürzeste Verbindungslinie zweier Punkte

Seien (x_0, y_0) und (x_1, y_1) zwei Punkte in der euklidischen Ebene. Jede beliebige Kurve zwischen diesen zwei Punkten lässt sich durch

$$C : [0, 1] \ni t \mapsto (x(t), y(t)) \quad (5.3.1)$$

beschreiben, wobei $(x(0), y(0)) = (x_0, y_0)$ und $(x(1), y(1)) = (x_1, y_1)$. Die Länge der Kurve ist

$$l(C) = \int_0^1 dt \sqrt{\dot{x}^2 + \dot{y}^2}. \quad (5.3.2)$$

Das Problem, die kürzeste Verbindungsstrecke zu finden, führt dann zu den Euler-Lagrange Gleichungen

$$\begin{aligned} \frac{\partial}{\partial x} \sqrt{\dot{x}^2 + \dot{y}^2} - \frac{d}{dt} \frac{\partial}{\partial \dot{x}} \sqrt{\dot{x}^2 + \dot{y}^2} &= 0, \\ \frac{\partial}{\partial y} \sqrt{\dot{x}^2 + \dot{y}^2} - \frac{d}{dt} \frac{\partial}{\partial \dot{y}} \sqrt{\dot{x}^2 + \dot{y}^2} &= 0. \end{aligned} \quad (5.3.3)$$

Der erste Term auf der linken Seite verschwindet jeweils (da die Wurzel nur von \dot{x} und \dot{y} abhängt, nicht aber von x und y), und daher folgt, dass

$$\frac{\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2}} \quad \text{und} \quad \frac{\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2}} \quad (5.3.4)$$

konstant sind. Diese beiden Größen sind gerade die Komponenten des Einheitsvektors in Richtung der Tangente der Kurve; die Kurve ist daher also eine Gerade.

Als zweites Beispiel analysieren wir die kürzeste Verbindungslinie zwischen zwei Punkten auf einer Kugeloberfläche. Wie parametrisieren die Kugeloberfläche durch die Kugelkoordinaten

$$(x, y, z) = r(\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta), \quad (5.3.5)$$

wobei $\theta \in [0, \pi)$ und $\varphi \in [0, 2\pi)$. (Der Radius r ist natürlich konstant.) Die Kurve wird wie zuvor durch

$$C : [0, 1] \ni t \mapsto (\theta(t), \varphi(t)) \quad (5.3.6)$$

beschrieben, und da $\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} = \sqrt{\dot{\theta}^2 + \sin^2 \theta \dot{\varphi}^2}$ ist die Länge der Kurve gleich

$$l(C) = r \int_0^1 \sqrt{\dot{\theta}^2 + \sin^2 \theta \dot{\varphi}^2}. \quad (5.3.7)$$

Die Euler-Lagrange Gleichungen lauten nun

$$\frac{d}{dt} \frac{\sin^2 \theta \dot{\varphi}}{\sqrt{\dot{\theta}^2 + \sin^2 \theta \dot{\varphi}^2}} = 0, \quad (5.3.8)$$

sowie

$$\frac{d}{dt} \frac{\dot{\theta}}{\sqrt{\dot{\theta}^2 + \sin^2 \theta \dot{\varphi}^2}} = \frac{\sin \theta \cos \theta \dot{\varphi}^2}{\sqrt{\dot{\theta}^2 + \sin^2 \theta \dot{\varphi}^2}}. \quad (5.3.9)$$

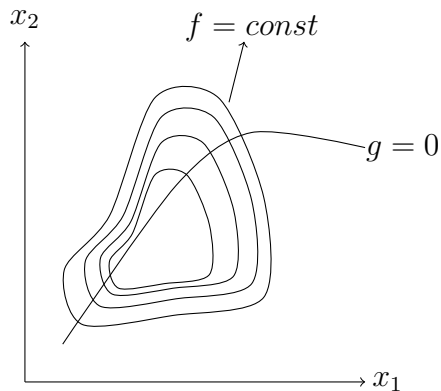
Ohne Beschränkung der Allgemeinheit wählen wir als Anfangspunkt der Kurve C den Nordpol ($\theta = 0$). Dann impliziert die erste Gleichung, dass

$$\frac{\sin^2 \theta \dot{\varphi}}{\sqrt{\dot{\theta}^2 + \sin^2 \theta \dot{\varphi}^2}} = 0 \quad (5.3.10)$$

auf der ganzen Kurve, also $\dot{\varphi} = 0$. [Bei $\theta = 0$ ist φ nicht eindeutig definiert. Für jede differenzierbare Kurve ist jedoch $r |\sin \theta \dot{\varphi}| \leq |\dot{\mathbf{x}}|$, d.h. $\sin^2 \theta \dot{\varphi}$ konvergiert gegen Null, wenn man sich dem Nordpol nähert.] Die kürzeste Verbindungslinie zu einem Punkt mit den Koordinaten (θ_1, φ_1) verläuft daher auf dem Grosskreis $\varphi = \varphi_1$.

5.3.2 Die Kettenlinie und Lagrange Multiplikatoren

Die Methode der Lagrange Multiplikatoren erlaubt es, das Maximum einer Funktion $f(x_1, \dots, x_n)$, $x_1, \dots, x_n \in \mathbb{R}$ unter der Nebenbedingung (NB), dass $g(x_1, \dots, x_n) = 0$, zu finden.



Für $n = 2$ kann man die Höhenlinien von f betrachten: Wenn man dem durch $g = 0$ beschriebenen Weg entlang geht, muss beim Maximum von f auf diesem Weg der Gradient von f orthogonal zu g liegen (um maximale Höhe zu gewinnen, müsste man orthogonal vom Weg abweichen).

Das heisst das Maximum ist erreicht wenn die Gradienten von f und g parallel sind, wenn also

$$\nabla f + \lambda \nabla g = 0$$

für ein geeignetes λ gilt. Der Parameter λ wird *Lagrange-Multiplikator* genannt. Zur Lösung des Problems definiert man eine neue Funktion

$$F(x_1, \dots, x_n) := f(x_1, \dots, x_n) + \lambda g(x_1, \dots, x_n)$$

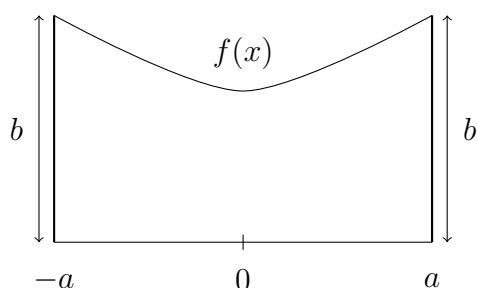
und beobachtet, dass jeder Extrempunkt von F einem Extrempunkt von f unter der NB $g = 0$ entspricht. Dies folgt daraus, dass für jeden Extrempunkt von F

$$\frac{dF}{d\lambda} = 0 \Rightarrow g = 0$$

gilt und falls $g = 0$ ist, dann folgt $F = f$.

Im Folgenden werden wir diese Methode auf die Optimierung von Funktionalen anwenden, d.h. die Koordinaten (x_1, \dots, x_n) werden durch Funktionen ersetzt (später werden es die Bahnkurven sein).

Als Beispiel betrachten wir die sogenannte Kettenlinie.

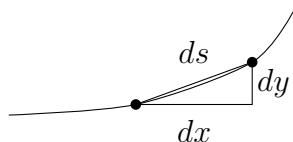


Wir betrachten eine Kette (mit homogener Massenverteilung), die unter dem Einfluss eines homogenen Schwerfeldes zwischen den Punkten $(-a, b)$ und (a, b) aufgehängt ist, und deren Länge L ist. Wir wollen berechnen, welche Form die aufgehängte Kette hat.

Wir suchen also die Funktion $f(x)$, die die Randbedingungen $f(-a) = b = f(a)$ erfüllt, und für die die potentielle Energie

$$V = g\rho_0 \int_{-a}^a ds f(x) \quad (5.3.11)$$

minimal ist, wobei ρ_0 die homogene Massendichte ist und wir über die Kettenlinie integrieren. Da g und ρ_0 konstant sind, können wir sie ohne Beschränkung der Allgemeinheit gleich 1 setzen.



Wegen

$$ds^2 = dx^2 + dy^2 \quad dy = f' dx$$

folgt

$$ds = dx \sqrt{1 + f'^2}$$

und damit

$$V = \int_{-a}^a dx f(x) \sqrt{1 + f'(x)^2}. \quad (5.3.12)$$

Natürlich hat dieses Variationsproblem nur dann eine vernünftige Lösung, falls nur Kurven betrachtet werden, deren totale Länge gerade L ist, also falls $f(x)$ die Nebenbedingung

$$L = \int_{-a}^a dx \frac{L}{2a} = \int_{-a}^a dx \sqrt{1 + f'(x)^2} \quad (5.3.13)$$

erfüllt. Um dieses Problem im Rahmen der oben angegebenen Methode behandeln zu können, führen wir einen sogenannten Lagrange'schen Multiplikator ein: wir betrachten das Problem als Variationsproblem von $f(x)$ und λ (wobei λ ein Parameter, und keine Funktion ist), und betrachten das Funktional

$$S[f(x), \lambda] = \int_{-a}^a dx \left[f(x) \sqrt{1 + f'(x)^2} + \lambda \left(\sqrt{1 + f'(x)^2} - \frac{L}{2a} \right) \right]. \quad (5.3.14)$$

Die Bedingung, dass S als Funktion von λ extremal ist, ist dann einfach die Nebenbedingung (5.3.13). Für die Kurven $f(x)$, die diese Nebenbedingung erfüllen, ist dann $S[f(x), \lambda] = V[f(x)]$, und daher minimieren wir tatsächlich (wie erwünscht) die potentielle Energie.

Für die Funktion $S[f(x), \lambda]$ ist nun die Euler-Lagrange Gleichung für $f(x)$ gerade

$$\frac{d}{dx} \left(\frac{(f(x) + \lambda) f'(x)}{\sqrt{1 + f'(x)^2}} \right) = \sqrt{1 + f'(x)^2}. \quad (5.3.15)$$

Wir definieren nun $g(x) = f(x) + \lambda$, und erhalten dann (nach Differentiation)

$$\frac{g'^2 + gg''}{\sqrt{1 + g'^2}} - \frac{g g'^2 g''}{(1 + g'^2)^{\frac{3}{2}}} = \sqrt{1 + g'^2}. \quad (5.3.16)$$

Ausmultiplizieren durch $(1 + g'^2)^{\frac{3}{2}}$ führt zu

$$g'^4 + g'^2(1 + gg'') + gg'' - g'^2 gg'' = g'^4 + 2g'^2 + 1, \quad (5.3.17)$$

was sich zu

$$g'^2 - gg'' + 1 = 0 \quad (5.3.18)$$

vereinfacht. Um die Lösung dieser Differentialgleichung zu erhalten, beobachten wir, dass sich die weitere Ableitung nach x als

$$2g'g'' - g'g'' - gg''' = -g^2 \left(\frac{g''}{g} \right)' = 0 \quad (5.3.19)$$

schreiben lässt. Falls also $g \neq 0$ (dieser Fall tritt nur ein, falls $L = 1$) folgt also

$$g'' = cg. \quad (5.3.20)$$

Die allgemeine Lösung davon ist natürlich

$$g(x) = A \cosh(\sqrt{c}x) + B \sinh(\sqrt{c}x). \quad (5.3.21)$$

Aus Symmetriegründen ist klar, dass $g(-x) = g(x)$, und daher ist $B = 0$. Einsetzen von (5.3.21) in (5.3.18) ergibt

$$A^2 c \sinh^2(\sqrt{c}x) - A^2 c \cosh^2(\sqrt{c}x) + 1 = 0, \quad (5.3.22)$$

was zu $A^2 c = 1$ führt. Die gesuchte Funktion $f(x)$ differiert von $g(x)$ um λ , und daher haben wir für $f(x)$ gerade

$$f(x) = A \cosh\left(\frac{x}{A}\right) + \lambda. \quad (5.3.23)$$

Die Werte von A und λ werden dann durch die Randbedingung $f(a) = b$ sowie durch die Nebenbedingung (5.3.13) eindeutig festgelegt.

5.4 Das Hamiltonsche Prinzip

Auch die Gesetze der Mechanik lassen sich (zumindest in einem konservativen Kraftfeld, d.h. einem Kraftfeld, das durch ein Potential beschrieben werden kann) aus einem Extremalprinzip ableiten. Dazu definieren wir die Lagrangefunktion

$$L = T - V, \quad (5.4.1)$$

wobei T die *kinetische* und V die *potentielle* Energie des Systems beschreibt. Diese sind Funktionen der Ortskoordinaten $q_\alpha(t)$, Geschwindigkeiten $\dot{q}_\alpha(t)$ und möglicherweise auch der Zeit t . Dann betrachten wir das Variationsprinzip für das Funktional

$$S[q(t)] = \int_{t_0}^{t_1} dt L(q_\alpha(t), \dot{q}_\alpha(t), t), \quad (5.4.2)$$

wobei wir den Anfangspunkt $q_\alpha(t_0)$ und Endpunkt $q_\alpha(t_1)$ festhalten. Die Euler-Lagrange Gleichungen dieses Variationsprinzips

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_\alpha} - \frac{\partial L}{\partial q_\alpha} = 0 \quad (5.4.3)$$

sind dann gerade die Newton'schen Bewegungsgleichungen — das ist der Inhalt des *Hamiltonschen Prinzips*. Die Funktion (oder besser das Funktional) $S[q(t)]$ nennt man die *Wirkung* der Bahn $q(t)$.

Im einfachsten Fall einer Bewegung einer Masse m in einer Dimension ist

$$T = \frac{m}{2} \dot{q}^2. \quad (5.4.4)$$

Dann ist

$$\frac{\partial L}{\partial \dot{q}} = m \dot{q}, \quad (5.4.5)$$

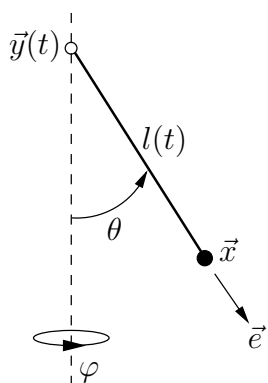
und die Euler-Lagrange-Gleichungen sind gerade

$$\frac{d}{dt} m \dot{q} = m \ddot{q} = \frac{\partial L}{\partial q} = -\frac{\partial V}{\partial q} = F, \quad (5.4.6)$$

wobei F die durch V beschriebene Kraft ist. Das ist natürlich gerade die Newton'sche Bewegungsgleichung eines Teilchens mit Masse m in einer Dimension. Es ist klar, dass dasselbe Argument auch für den Fall von N Teilchen in D Dimensionen gilt.

5.4.1 Holonome Zwangsbedingungen

Die Umformulierung der Newton'schen Bewegungsgleichungen als Variationsprinzip ist besonders nützlich für Systeme mit *Zwangsbedingungen*. In dieser Betrachtungsweise wird ein Teil der Kräfte indirekt dadurch beschrieben, dass die Bewegungsmöglichkeiten der Massepunkte eingeschränkt wird. So ist zum Beispiel beim mathematischen Pendel die Bewegung des in einer vertikalen Ebene schwingenden Massenpunktes dadurch eingeschränkt, dass der Faden den Abstand vom Aufhängepunkt konstant hält. Ein anderes Beispiel ist das sphärische Pendel, das wir nun kurz erklären wollen:



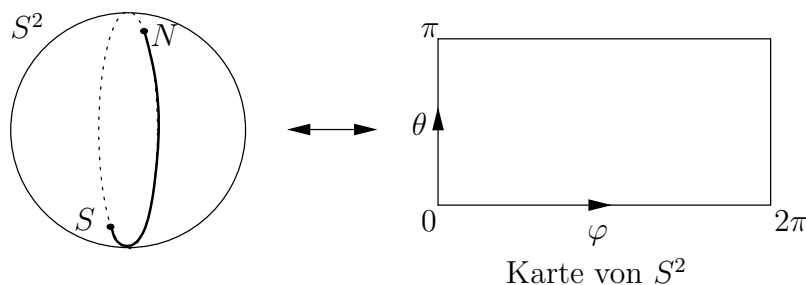
Der Aufhängepunkt $\mathbf{y}(t)$ und die Pendellänge $l(t)$ sind vorgeschriebene Funktionen der Zeit t . Die unter diesen Zwangsbedingungen noch zulässigen Lagen des Pendels sind

$$\mathbf{x} = \mathbf{y}(t) + l(t)\mathbf{e} \equiv \mathbf{x}(\mathbf{e}, t), \quad (5.4.7)$$

wobei \mathbf{e} ein Punkt auf der Einheitskugel S^2 ist (Konfigurationsraum des Pendels). Die Bewegung des Pendels ist durch eine Funktion $t \mapsto \mathbf{e}(t) \in S^2$ gegeben. Auf S^2 können Lagekoordinaten eingeführt werden, z.B. Polarkoordinaten θ, φ :

$$\mathbf{e}(\theta, \varphi) = (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta). \quad (5.4.8)$$

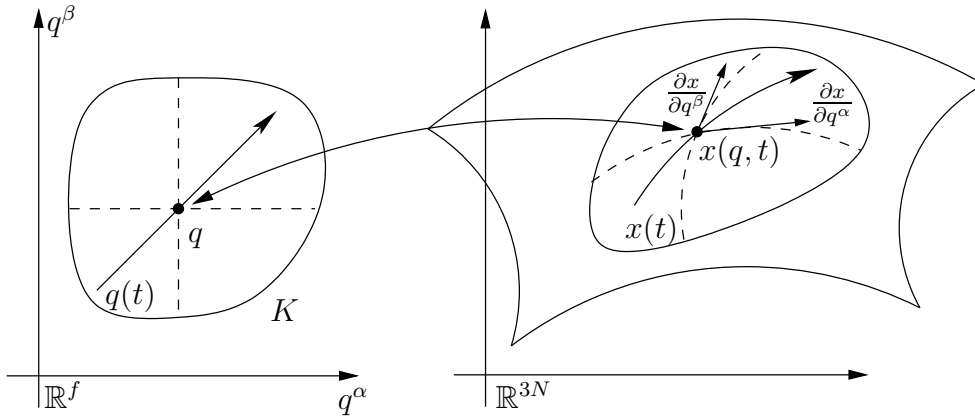
Dadurch wird der Konfigurationsraum auf ein Stück des \mathbb{R}^2 abgebildet.



Die Abbildung ist nicht bijektiv auf dem Rand des Rechtecks (bzw. auf dem Nullmeridian von S^2). Zur Darstellung von ganz S^2 benötigt man mindestens zwei Karten (einen Atlas).

Im allgemeinen haben wir ein System von N Teilchen, deren kartesische Koordinaten $x = (\mathbf{x}_1, \dots, \mathbf{x}_n) \in \mathbb{R}^{3N}$ zur Zeit t eine glatte f -dimensionale Fläche im \mathbb{R}^{3N} bilden. (Zwangsbedingungen dieser Art werden 'holonome Zwangsbedingungen' genannt.) Dieses System hat dann f Freiheitsgrade. (In dem obigen Beispiel ist $N = 1$ und $f = 2$, wir haben also zwei Freiheitsgrade.) Lokal können wir das System durch Koordinaten $q = (q_1, \dots, q_f)$ beschreiben, d.h. wir haben eine Karte $K \subset \mathbb{R}^f$,

$$K \rightarrow \mathbb{R}^{3N}, \quad q \mapsto x(q, t), \quad (5.4.9)$$



die in (q, t) differenzierbar ist. Weiterhin hat die Tangentialabbildung Rang f , d.h. die Vektoren

$$\frac{\partial x}{\partial q_\alpha} \in \mathbb{R}^{3N}, \quad (\alpha = 1, \dots, f), \quad (5.4.10)$$

sind linear unabhängig. Jede zulässige (mit den Zwangsbedingungen verträgliche) Bewegung $x(t)$ des Systems (im Bereich der Karte K) ist dann durch eine Funktion $t \mapsto q(t) \in K$ dargestellt:

$$x(t) = x(q(t), t). \quad (5.4.11)$$

Die Lagrangefunktion

$$L = \frac{1}{2} \sum_{i=1}^N m_i \dot{\mathbf{x}}_i^2 - V(\mathbf{x}_i) \quad (5.4.12)$$

wird dann als Funktion der Konfigurationsraumkoordinaten q_α , sowie ihrer Zeitableitungen (und eventuell der Zeit t) aufgefasst. Zum Beispiel haben wir

$$\frac{\partial L}{\partial \dot{q}_\alpha} = \sum_{i=1}^N m_i \dot{\mathbf{x}}_i \cdot \frac{\partial \mathbf{x}_i}{\partial q_\alpha}, \quad (5.4.13)$$

da

$$\frac{\partial \dot{\mathbf{x}}_i}{\partial \dot{q}_\alpha} = \frac{\partial \mathbf{x}_i}{\partial q_\alpha}. \quad (5.4.14)$$

Weiterhin ist

$$\frac{\partial L}{\partial q_\alpha} = \sum_{i=1}^N m_i \dot{\mathbf{x}}_i \cdot \frac{\partial \dot{\mathbf{x}}_i}{\partial q_\alpha} - \sum_j \frac{\partial V}{\partial \mathbf{x}_j} \cdot \frac{\partial \mathbf{x}_j}{\partial q_\alpha}. \quad (5.4.15)$$

Die Euler-Lagrange Gleichungen sind daher

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_\alpha} = \sum_{i=1}^N m_i \ddot{\mathbf{x}}_i \cdot \frac{\partial \mathbf{x}_i}{\partial q_\alpha} + \sum_{i=1}^N m_i \dot{\mathbf{x}}_i \cdot \frac{d}{dt} \frac{\partial \mathbf{x}_i}{\partial q_\alpha} = \frac{\partial L}{\partial q_\alpha} = \sum_{i=1}^N m_i \dot{\mathbf{x}}_i \cdot \frac{\partial \dot{\mathbf{x}}_i}{\partial q_\alpha} - \sum_j \frac{\partial V}{\partial \mathbf{x}_j} \cdot \frac{\partial \mathbf{x}_j}{\partial q_\alpha}. \quad (5.4.16)$$

Die unterstrichenen Terme heben sich gerade gegenseitig weg, und wir erhalten

$$\sum_{i=1}^N m_i \ddot{\mathbf{x}}_i \cdot \frac{\partial \mathbf{x}_i}{\partial q_\alpha} = - \sum_i \frac{\partial V}{\partial \mathbf{x}_i} \cdot \frac{\partial \mathbf{x}_i}{\partial q_\alpha}, \quad (5.4.17)$$

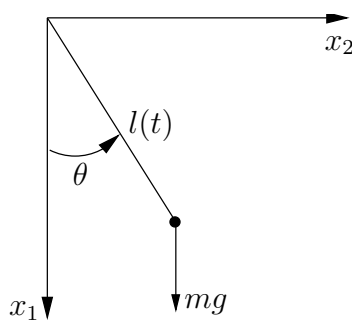
was gerade mit den üblichen Newton'schen Bewegungsgleichungen übereinstimmt!

Die Bewegungsgleichungen eines Systems mit Zwangsbedingungen kann man daher im allgemeinen wie folgt aufstellen:

- (i) Schreibe $T = \frac{1}{2} \sum_{i=1}^N m_i \dot{\mathbf{x}}_i^2$, $V = V(\mathbf{x}_i)$ in kartesischen Komponenten, ohne Berücksichtigung der Zwangsbedingungen.
- (ii) Drücke die gemäss der Zwangsbedingungen zulässigen Konfigurationen der $x = (\mathbf{x}_1, \dots, \mathbf{x}_N)$ durch unabhängige Lagekoordinaten $q = (q_1, \dots, q_f)$ aus, $x = x(q, t)$, und berechne $\dot{x} = \dot{x}(q, \dot{q}, t)$.
- (iii) Setze x, \dot{x} in $L = T - V$ ein.
- (iv) Stelle die Euler-Lagrange Gleichungen (5.4.3) auf.

Diese Vorgehensweise soll nun an ein paar Beispielen illustriert werden.

Beispiel 1 (siehe Kapitel 4.3)



Ebenes Pendel mit festem Aufhängepunkt und vorgeschriebener, zeitabhängiger Länge $l(t)$. Die Lagekoordinate sei θ . Aus

$$\begin{aligned} \mathbf{x} &= l(\cos \theta, \sin \theta) \\ \dot{\mathbf{x}} &= \dot{l}(\cos \theta, \sin \theta) + l\dot{\theta}(-\sin \theta, \cos \theta) \end{aligned}$$

berechnet man

$$T = \frac{m}{2} (\dot{l}^2 + (l\dot{\theta})^2), \quad V = -mgx_1 = -mgl \cos \theta,$$

und daraus

$$\frac{\partial L}{\partial \dot{\theta}} = ml^2 \dot{\theta}, \quad \frac{\partial L}{\partial \theta} = -mgl \sin \theta.$$

Damit lautet (5.4.3)

$$\frac{d}{dt}(ml(t)^2 \dot{\theta}) + mgl(t) \sin \theta = 0.$$

was gerade die in 4.3 benützte Bewegungsgleichung ist.

Beispiel 2 (Umgekehrtes Pendel — siehe Kapitel 4.4).

Die Koordinaten und die Geschwindigkeiten der Masse m sind

$$\mathbf{x} = (x + l \sin \theta, l \cos \theta), \quad \dot{\mathbf{x}} = (\dot{x} + l\dot{\theta} \cos \theta, -l\dot{\theta} \sin \theta).$$

Damit ist

$$T = \frac{1}{2}M\dot{x}^2 + \frac{1}{2}m(\dot{x}^2 + l^2\dot{\theta}^2 + 2l\dot{x}\dot{\theta} \cos \theta), \quad V = mgl \cos \theta - ux,$$

also

$$\begin{aligned} \frac{\partial L}{\partial \dot{x}} &= (M + m)\dot{x} + ml\dot{\theta} \cos \theta, & \frac{\partial L}{\partial x} &= u, \\ \frac{\partial L}{\partial \dot{\theta}} &= m(l^2\dot{\theta} + l\dot{x} \cos \theta), & \frac{\partial L}{\partial \theta} &= ml \sin \theta (g - \dot{x}\dot{\theta}) \end{aligned}$$

und die Gl. (5.4.3) lauten

$$\begin{aligned} (M + m)\ddot{x} + ml(\ddot{\theta} \cos \theta - \dot{\theta}^2 \sin \theta) &= u, \\ l\ddot{\theta} + \ddot{x} \cos \theta - \dot{x}\dot{\theta} \sin \theta &= \sin \theta (g - \dot{x}\dot{\theta}). \end{aligned}$$

Daraus erhält man durch Linearisierung in θ und $\dot{\theta}$ die in Kapitel 4.4 benützten Bewegungsgleichungen.

Beispiel 3 (Geladenes Teilchen im elektromagnetischen Feld.)

Aus den homogenen Maxwell Gleichungen

$$\operatorname{div} \mathbf{B} = 0, \quad \operatorname{rot} \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} = 0$$

folgt, dass sich das magnetische und elektrische Feld ($\mathbf{B}(\mathbf{x}, t)$ und $\mathbf{E}(\mathbf{x}, t)$) durch elektromagnetische Potentiale φ, \mathbf{A} darstellen lassen:

$$\mathbf{B} = \operatorname{rot} \mathbf{A}, \quad \mathbf{E} = -\nabla \varphi - \frac{1}{c} \frac{\partial \mathbf{A}}{\partial t}. \quad (5.4.18)$$

Die Bewegung eines Teilchens (in kartesischen Koordinaten $\mathbf{x} \in \mathbb{R}^3$) ist bestimmt durch die Lagrangefunktion

$$L(\mathbf{x}, \dot{\mathbf{x}}, t) = \frac{1}{2}m\dot{\mathbf{x}}^2 - e[\varphi(\mathbf{x}, t) - \frac{\dot{\mathbf{x}}}{c} \cdot \mathbf{A}(\mathbf{x}, t)], \quad (5.4.19)$$

wobei m, e die Masse und die Ladung des Teilchens sind. In der Tat ist

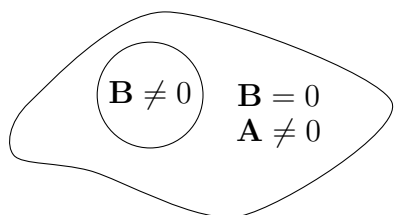
$$\begin{aligned} \frac{\partial L}{\partial x_k} &= -e \frac{\partial \varphi}{\partial x_k} + \frac{e}{c} \sum_{i=1}^3 \dot{x}_i \frac{\partial A_i}{\partial x_k}, & \frac{\partial L}{\partial \dot{x}_k} &= m\dot{x}_k + \frac{e}{c} A_k, \\ \frac{d}{dt} \frac{\partial L}{\partial \dot{x}_k} &= m\ddot{x}_k + \frac{e}{c} \left(\frac{\partial A_k}{\partial t} + \sum_{i=1}^3 \frac{\partial A_k}{\partial x_i} \cdot \dot{x}_i \right), \end{aligned}$$

so dass (5.4.3) lautet

$$m\ddot{x}_k = e \underbrace{\left(-\frac{\partial\varphi}{\partial x_k} - \frac{1}{c} \frac{\partial A_k}{\partial t}\right)}_{E_k} + \frac{e}{c} \underbrace{\sum_{i=1}^3 \dot{x}_i \left(\frac{\partial A_i}{\partial x_k} - \frac{\partial A_k}{\partial x_i}\right)}_{(\dot{\mathbf{x}} \wedge \mathbf{B})_k},$$

was mit (1.4.7) übereinstimmt.

Innerhalb der Quantenmechanik ist das Vektorpotential \mathbf{A} direkt physikalisch relevant. Dies lässt sich zum Beispiel mit Hilfe des Aharonov-Bohm Experiments verstehen.



Ein Teilchen bewegt sich in einem Gebiet, welches isoliert ist von einem externen magnetischen Feld. Obwohl hier $\mathbf{B} = 0$ (aber nicht $\mathbf{A} = 0$) gilt, reagiert das Teilchen auf Änderungen des Feldes. Das Vektorpotential $\mathbf{A} = 0$ ist also nicht nur ein mathematisches Konstrukt sondern von physikalischer Relevanz.

Beispiel 4 (Doppelpendel)

Beim (ebenen) Doppelpendel bewegen sich zwei Massepunkte unter dem Einfluss eines homogenen Schwerfeldes in einer vertikalen Ebene, der erste mit einem festen Abstand l_1 von einem festgehaltenen Punkt (den wir als den Ursprung wählen), und der zweite mit einem festen Abstand l_2 vom ersten Massepunkt. Wir bezeichnen die Auslenkungen von der Vertikalen vom jeweiligen Aufhängepunkt durch φ_1 und φ_2 . Dann sind die Ortskoordinaten (in der Ebene)

$$\mathbf{x}_1 = l_1(\cos \varphi_1, \sin \varphi_1), \quad \mathbf{x}_2 = \mathbf{x}_1 + l_2(\cos \varphi_2, \sin \varphi_2). \quad (5.4.20)$$

Für die kinetische Energie findet man ($\dot{\mathbf{x}}_1 = l_1\dot{\varphi}_1(-\sin \varphi_1, \cos \varphi_1)$)

$$T = \frac{m_1 + m_2}{2} l_1^2 \dot{\varphi}_1^2 + m_2 l_1 l_2 \dot{\varphi}_1 \dot{\varphi}_2 \cos(\varphi_1 - \varphi_2) + \frac{m_2}{2} l_2^2 \dot{\varphi}_2^2, \quad (5.4.21)$$

während die potentielle Energie durch

$$V = -m_1 l_1 g \cos \varphi_1 - m_2 g (l_1 \cos \varphi_1 + l_2 \cos \varphi_2) \quad (5.4.22)$$

gegeben ist. Die Lagrangefunktion ist $L = T - V$, und daraus kann man dann leicht die Bewegungsgleichungen ableiten. Das Doppelpendel ist jedoch *nicht integrabel*: die Bewegungsgleichungen können nicht auf die Berechnung von Integralen zurückgeführt werden. In der Tat weisen numerische Berechnungen (und Experimente) auf ein ‘chaotisches Verhalten’ hin: das Verhalten des Doppelpendels hängt nach endlicher Zeit bereits in starker Weise von minimalen Änderungen der Anfangsbedingungen ab!

5.4.2 Äquivalente Lagrangefunktionen

Zwei Lagrange-Funktionen L_1 und L_2 , die sich nur um die totale zeitliche Ableitung einer Funktion $F(q, t)$ unterscheiden, d.h.

$$L_1 - L_2 = \frac{d}{dt}F(q, t) \equiv \sum_{\alpha} \frac{\partial F}{\partial q_{\alpha}} \dot{q}_{\alpha} + \frac{\partial F}{\partial t}, \quad (5.4.23)$$

beschreiben dieselbe Dynamik, d.h. führen zu denselben Euler-Lagrange Gleichungen. Bei festen Endpunkten ist nämlich

$$\delta \int_{(1)}^{(2)} \frac{d}{dt}F(q, t) dt = \delta F(q, t) \Big|_{(1)}^{(2)} = 0.$$

Natürlich kann man auch direkt zeigen, dass L_1 und L_2 zu denselben Euler-Lagrange Gleichungen führen.

Bemerkung (zur Variation δF):



q_{α} : Schar von Kurven mit $q^0 = id$.

Dann ist

$$\delta F = \left. \frac{d}{d\alpha} F(q_{\alpha}) \right|_{\alpha=0}. \quad (5.4.24)$$

In der vorherigen Rechnung haben wir also

$$\delta F(q, t) \Big|_{(1)}^{(2)} = \underbrace{\delta F(q(t_2))}_{=0} - \underbrace{\delta F(q(t_1))}_{=0} = 0$$

Beispiel 3 (Fortsetzung). Ändert man die Potentiale (φ, \mathbf{A}) durch eine *Eichtransformation*

$$\varphi \rightarrow \varphi - \frac{1}{c} \frac{\partial \chi}{\partial t}, \quad \mathbf{A} \rightarrow \mathbf{A} + \nabla \chi, \quad (5.4.25)$$

wobei χ eine beliebige skalare Funktion ist, so ändert sich das elektromagnetische Feld (5.4.18) nicht. Daher sollten auch die Bewegungsgleichung (1.4.7) invariant sein. In der Tat geht L (s. (5.4.19)) unter (5.4.25) in die äquivalente Lagrange-Funktion

$$L + \frac{e}{c} \left(\frac{\partial \chi}{\partial t} + \dot{x} \cdot \nabla \chi \right) = L + \frac{e}{c} \frac{d\chi}{dt} \quad (5.4.26)$$

über.

5.5 Zyklische Koordinaten und konjugierte Impulse

Sei L die Lagrangefunktion eines Lagrangeschen Systems. Falls L von der Koordinate q_α unabhängig ist, $\frac{\partial L}{\partial q_\alpha} = 0$, dann nennt man die Koordinate q_α zyklisch. Zu jeder Koordinate q_α definiert man den *konjugierten Impuls* durch

$$p_\alpha = \frac{\partial L}{\partial \dot{q}_\alpha}. \quad (5.5.1)$$

Ist q_α zyklisch, dann ist wegen der Euler-Lagrange-Gleichung

$$\frac{d}{dt} p_\alpha = \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_\alpha} = \frac{\partial L}{\partial q_\alpha} = 0 \quad (5.5.2)$$

der konjugierte Impuls eine *Erhaltungsgrösse*.

Als Beispiel betrachten wir das sphärischen Pendel, das zu Beginn von 5.3.1 kurz erwähnt wurde. Der Einfachheit halber wählen wir $\mathbf{y}(t) = \mathbf{0}$ und $l(t) \equiv l$. Dann ist

$$\mathbf{x}(\theta, \varphi) = l(\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta). \quad (5.5.3)$$

Die Geschwindigkeit $\dot{\mathbf{x}}$ ist daher

$$\begin{aligned} \dot{\mathbf{x}} &= \frac{\partial \mathbf{x}}{\partial \theta} \dot{\theta} + \frac{\partial \mathbf{x}}{\partial \varphi} \dot{\varphi} \\ &= l \dot{\theta} (\cos \theta \cos \varphi, \cos \theta \sin \varphi, -\sin \theta) + l \sin \theta \dot{\varphi} (-\sin \varphi, \cos \varphi, 0). \end{aligned}$$

Damit ergibt sich die kinetische Energie zu

$$T = \frac{1}{2} l^2 (\dot{\theta}^2 + \sin^2 \theta \dot{\varphi}^2). \quad (5.5.4)$$

Andererseits ist die potentielle Energie

$$V = -g l \cos \theta. \quad (5.5.5)$$

Die Lagrangefunktion $L = T - V$ hängt nicht von φ ab. Die zugehörigen Erhaltungsgrösse ist dann

$$p_\varphi = l^2 \sin^2 \theta \dot{\varphi}, \quad (5.5.6)$$

was gerade der Drehimpuls um die vertikale Achse ist; er ist erhalten, da in vertikaler Richtung kein Drehmoment wirkt.

Auf ähnliche Weise kann man zeigen, dass falls L (explizit) unabhängig von der Zeit t ist, $\frac{\partial L}{\partial t} = 0$, dass dann die Grösse

$$E = \sum_{\alpha=1}^f p_\alpha \dot{q}_\alpha - L = \sum_{\alpha=1}^f \frac{\partial L}{\partial \dot{q}_\alpha} \dot{q}_\alpha - L \quad (5.5.7)$$

erhalten ist. In der Tat rechnet man leicht nach, dass

$$\frac{dE}{dt} = \sum_{\alpha=1}^f \left(\ddot{q}_{\alpha} \frac{\partial L}{\partial \dot{q}_{\alpha}} + \dot{q}_{\alpha} \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_{\alpha}} - \ddot{q}_{\alpha} \frac{\partial L}{\partial \dot{q}_{\alpha}} - \dot{q}_{\alpha} \frac{\partial L}{\partial q_{\alpha}} \right) - \frac{\partial L}{\partial t} = -\frac{\partial L}{\partial t} = 0, \quad (5.5.8)$$

vorausgesetzt, dass $q(t)$ die Euler-Lagrange-Gleichung löst. Ist $L = T - V$, wobei

$$T(q, \dot{q}) = \frac{1}{2} \sum_{\alpha, \beta=1}^f a_{\alpha\beta}(q) \dot{q}_{\alpha} \dot{q}_{\beta} \quad (5.5.9)$$

und V ist geschwindigkeitsunabhängig (d.h. unabhängig von den \dot{q}_{α}), dann gilt

$$\sum_{\alpha=1}^f \dot{q}_{\alpha} \frac{\partial L}{\partial \dot{q}_{\alpha}} = \sum_{\alpha, \beta=1}^f a_{\alpha\beta}(q) \dot{q}_{\alpha} \dot{q}_{\beta} = 2T, \quad (5.5.10)$$

und E ist einfach

$$E = T + V. \quad (5.5.11)$$

E hat in diesem Fall die Interpretation der Energie.

Als Beispiel betrachten wir ein geladenes Teilchen in einem *statischen* (zeit-unabhängigen) Feld (Fortsetzung von **Beispiel 3**). In diesem Fall sind die Maxwell Gleichungen

$$\mathbf{E} = -\nabla\varphi(\mathbf{x}), \quad \mathbf{B} = \text{rot}\mathbf{A}(\mathbf{x}). \quad (5.5.12)$$

Die Lagrangefunktion ist dann auch zeit-unabhängig, und der zu \mathbf{x} konjugierte Impuls ist

$$\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{x}}} = m\dot{\mathbf{x}} + \frac{e}{c}\mathbf{A}, \quad (5.5.13)$$

und daher ist die Grösse

$$\mathbf{p} \cdot \dot{\mathbf{x}} - L = \frac{1}{2}m\dot{\mathbf{x}}^2 + e\varphi(\mathbf{x}) \quad (5.5.14)$$

erhalten.

Ein anderes Beispiel ist das sphärische Pendel, für das (wie wir gerade gesehen haben) die Lagrangefunktion von t unabhängig ist. In diesem Fall ist die erhaltene Energie

$$E = \dot{\theta} \frac{\partial L}{\partial \dot{\theta}} + \dot{\varphi} \frac{\partial L}{\partial \dot{\varphi}} - L = \frac{l^2}{2}\dot{\theta}^2 + V_{\text{eff}}(\theta), \quad (5.5.15)$$

wobei

$$V_{\text{eff}}(\theta) = \frac{p_{\varphi}^2}{2l^2 \sin^2 \theta} - gl \cos \theta \quad (5.5.16)$$

ist. Zusammen mit dem erhaltenen Drehimpuls p_{φ} kann man dann die Bewegungsgleichungen durch Separation der Variablen lösen. Unter Benützung der Energieerhaltung finden wir nämlich

$$t = \int d\theta \frac{l}{\sqrt{2(E - V_{\text{eff}}(\theta))}}. \quad (5.5.17)$$

Nun wenden wir die Substitution $w = 1 - \cos \theta$ an, und finden

$$t(\theta) - t(\theta_1) = \sqrt{\frac{l}{2g}} \int_{w_1}^{1-\cos \theta} \frac{dw}{\sqrt{w^3 - aw^2 + bw + c}}, \quad (5.5.18)$$

wobei

$$a = -(2 + w_0), \quad b = 2w_0, \quad c = -\frac{p_\varphi^2}{2gl^3}, \quad w_0 = 1 + \frac{E}{gl}. \quad (5.5.19)$$

Das Polynom unter der Wurzel im Nenner des Integranden hat 3 reelle Nullstellen, $w_1 < w_2 < w_3$, d.h. ist von der Form $(w - w_1)(w - w_2)(w - w_3)$. Für $w_1 < w < w_2$ ist dieser Ausdruck positiv. Wir substituieren $w = w_1 + (w_3 - w_1)u^2$ mit $u > 0$ und finden dann

$$t(\theta) - t(\theta_1) = \sqrt{\frac{2l}{g(w_2 - w_1)}} \int_0^{\sqrt{\frac{w-w_1}{w_3-w_1}}} \frac{du}{\sqrt{(1-u^2) \left(1 - \frac{w_3-w_1}{w_2-w_1}u^2\right)}}, \quad (5.5.20)$$

und schliesslich ersetzen wir $u = \sin \xi$, $du = \cos \xi d\xi = \sqrt{1-u^2} d\xi$ und erhalten damit

$$t(\theta) - t(\theta_1) = \sqrt{\frac{2l}{g(w_2 - w_1)}} \int_0^{\arcsin \sqrt{\frac{w-w_1}{w_3-w_1}}} \frac{d\xi}{\sqrt{\left(1 - \frac{w_3-w_1}{w_2-w_1} \sin^2 \xi\right)}}. \quad (5.5.21)$$

Für kleine $\delta = \frac{w_3-w_1}{w_2-w_1}$ können wir den Integranden bis zur ersten Ordnung entwickeln,

$$\frac{1}{\sqrt{1 - \delta \sin^2 \xi}} = 1 + \frac{\delta}{2} \sin^2 \xi + \mathcal{O}(\delta^2),$$

und finden den Näherungsausdruck

$$t(\theta) - t(\theta_1) = \sqrt{\frac{2l}{g(w_2 - w_1)}} \left[\xi \left(1 + \frac{\delta}{4}\right) - \frac{\delta}{8} \sin 2\xi \right], \quad (5.5.22)$$

wobei $\xi = \arcsin \sqrt{\frac{1-\cos \theta - w_1}{w_3 - w_1}}$.

5.6 Noether's Theorem

Das Noether'sche Theorem ermöglicht eine konzeptionell einfache und allgemeine Ableitung von Erhaltungssätzen aus Symmetrien. Dabei betrachten wir nur kontinuierliche Symmetrien und keine Spiegelung von Raum oder Zeit. Zur Vereinfachung betrachten wir den Fall, wo der Konfigurationsraum einfach (global) \mathbb{R}^f ist. Der allgemeine Fall, in dem der Konfigurationsraum aus einem Atlas von mehreren Karten besteht, lässt sich daraus ableiten.

Definition. Ein Fluss ist eine Schar ϕ^λ , ($\lambda \in \mathbb{R}$), von Abbildungen des Konfigurationsraums auf sich, so dass

$$\phi^0 = \text{id}, \quad \phi^\lambda \circ \phi^\mu = \phi^{\lambda+\mu}. \quad (5.6.1)$$

Definition. Jeder Fluss hat ein erzeugendes Vektorfeld $v(q)$ auf \mathbb{R}^f , das durch

$$v(q) = \left. \frac{\partial}{\partial \lambda} \phi^\lambda(q) \right|_{\lambda=0} \quad (5.6.2)$$

definiert ist.

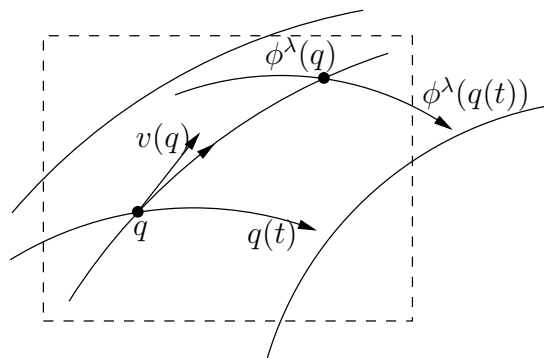
Wegen der Gruppeneigenschaft des Flusses gilt (5.6.2) für alle $(\lambda; q) \in \mathbb{R} \times \mathbb{R}^f$

$$\left. \frac{\partial}{\partial \lambda} \phi^\lambda(q) \right|_{\lambda=\varepsilon} = \left. \frac{\partial}{\partial \varepsilon} \phi^{\lambda+\varepsilon}(q) \right|_{\varepsilon=0} = \left. \frac{\partial}{\partial \varepsilon} \phi^\varepsilon(\phi^\lambda(q)) \right|_{\varepsilon=0} = v(\phi^\lambda(q)). \quad (5.6.3)$$

Die Funktion $q(\lambda) = \phi^\lambda(q)$ ist also die Lösung der Differentialgleichung 1. Ordnung

$$\frac{dq}{d\lambda} = v(q(\lambda)) \quad \text{mit der Anfangsbedingung } q(0) = q. \quad (5.6.4)$$

Damit bestimmt umgekehrt das Vektorfeld den Fluss, vorausgesetzt, dass (5.6.4) für jeden Anfangswert q eine globale Lösung $q(\lambda)$ besitzt. Der Begriff des Flusses überträgt sich auf einen allgemeinen Konfigurationsraum: In einer Karte ist jeweils nur ein Ausschnitt des Flusses ϕ^λ darstellbar; im Allgemeinen benötigt man mehrere Karten (einen Atlas), um den ganzen Konfigurationsraum zu überdecken.



Sei $q(t)$ nun eine Kurve im Konfigurationsraum. Der Fluss ϕ^λ bildet die Kurve $q(t)$ auf $\phi^\lambda(q(t))$ ab. Wir nennen den Fluss ϕ^λ eine *kontinuierliche Symmetrie* einer Lagrange-Funktion $L(q, \dot{q}, t)$, falls

$$L(\phi^\lambda(q(t)), \frac{\partial}{\partial t} \phi^\lambda(q(t)), t) = L(q(t), \dot{q}(t), t) \quad (5.6.5)$$

für alle $\lambda \in \mathbb{R}$ und für jede Kurve $t \mapsto q(t) \in \mathbb{R}^f$.

Noether's Theorem sagt nun aus, dass zu jeder kontinuierlichen Symmetrie eine Erhaltungsgrösse gehört. Explizit gilt: sei ϕ^λ eine kontinuierliche Symmetrie für L , so ist

$$\langle p, v(q) \rangle \equiv \sum_{\alpha=1}^f p_\alpha v_\alpha(q), \quad (5.6.6)$$

wobei $p_\alpha = \frac{\partial L}{\partial \dot{q}_\alpha}$ der zu q_α konjugierte Impuls ist, erhalten, d.h. es gilt

$$\frac{d}{dt} \langle p, v(q) \rangle = 0.$$

[Diese Aussage gilt nur auf der tatsächlichen Bahn, d.h. falls $q(t)$ eine Lösung der Euler-Lagrange Gleichungen ist!]

Der Beweis ist sehr einfach: wir berechnen

$$\begin{aligned} \frac{d}{dt} \left(\sum_{\alpha=1}^f \frac{\partial L}{\partial \dot{q}_\alpha} v_\alpha(q) \right) &= \sum_{\alpha=1}^f \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_\alpha} \right) v_\alpha(q) + \sum_{\alpha=1}^f \frac{\partial L}{\partial \dot{q}_\alpha} \frac{d}{dt} v_\alpha(q) \\ &= \sum_{\alpha=1}^f \frac{\partial L}{\partial q_\alpha} \frac{\partial}{\partial \lambda} \phi^\lambda(q(t)) \Big|_{\lambda=0} + \sum_{\alpha=1}^f \frac{\partial L}{\partial \dot{q}_\alpha} \frac{\partial}{\partial \lambda} \frac{d}{dt} \phi^\lambda(q(t)) \Big|_{\lambda=0} \\ &= \frac{d}{d\lambda} L \left(\phi^\lambda(q(t)), \frac{d}{dt} \phi^\lambda(q(t)), t \right) \Big|_{\lambda=0} = 0 \end{aligned} \quad (5.6.7)$$

wobei wir in der zweiten Zeile die Euler-Lagrange Gleichungen sowie die Definition von v_α , (5.6.3), benützt haben. Die letzte Zeile verschwindet wegen der Annahme, dass ϕ^λ eine kontinuierliche Symmetrie definiert, siehe (5.6.5).

Zum Beispiel kann man die Erhaltung des Impulses einer zyklischen Koordinate (siehe Kapitel 5.4) auf diese Weise verstehen: falls q_α zyklisch ist, so definiert der Fluss

$$\phi^\lambda(q_\beta) = q_\beta + \delta_{\beta\alpha} \lambda \quad (5.6.8)$$

eine kontinuierliche Symmetrie der Lagrangefunktion. Das zugehörige Vektorfeld hat dann nur eine nicht-triviale Komponente,

$$v^\beta = \delta_{\alpha\beta} \quad (5.6.9)$$

und die Erhaltungsgrösse, die durch (5.6.6) definiert wird, ist einfach der konjugierte Impuls p_α .

Noether's Theorem lässt sich in zweierlei Weise verallgemeinern:

- (i) Die Lagrange-Funktion ist nicht notwendigerweise invariant unter der kontinuierlichen Symmetrie (Bedingung (5.6.5)), sondern geht stattdessen nur in eine äquivalente

Lagrangefunktion über (die sich um die totale Zeitableitung von der ursprünglichen Lagrangefunktion unterscheidet)

$$L(\phi^\lambda(q(t)), \frac{\partial}{\partial t} \phi^\lambda(q(t)), t) = L(q(t), \dot{q}(t), t) + \frac{d}{dt} F(q, t, \lambda). \quad (5.6.10)$$

In diesem Fall ist die Erhaltungsgrösse

$$\langle p, v(q) \rangle - \delta F, \quad (5.6.11)$$

wobei die Varianz δF in (5.4.24) definiert ist.

Beweis. Wir berechnen zuerst

$$\begin{aligned} \left. \frac{d}{d\lambda} L(\phi^\lambda(q(t)), \frac{\partial}{\partial t} \phi^\lambda(q(t)), t) \right|_{\lambda=0} &= \left. \frac{d}{d\lambda} L(q(t), \dot{q}(t), t) \right|_{\lambda=0} + \left. \frac{d}{d\lambda} \frac{d}{dt} F(q, t, \lambda) \right|_{\lambda=0} \\ &= \left. \frac{d}{dt} \frac{d}{d\lambda} F(q, t, \lambda) \right|_{\lambda=0} \\ &= \frac{d}{dt} \delta F. \end{aligned}$$

Von vorher wissen wir (siehe (5.6.7)), dass

$$\begin{aligned} \frac{d}{dt} \langle p, v(q) \rangle &= \frac{d}{dt} \left(\sum_{\alpha=1}^f \frac{\partial L}{\partial \dot{q}_\alpha} v_\alpha(q) \right) \\ &= \left. \frac{d}{d\lambda} L \left(\phi^\lambda(q(t)), \frac{d}{dt} \phi^\lambda(q(t)), t \right) \right|_{\lambda=0}. \end{aligned}$$

Also ist in diesem Fall

$$\begin{aligned} \frac{d}{dt} (\langle p, v(q) \rangle - \delta F) &= \left. \frac{d}{d\lambda} L \left(\phi^\lambda(q(t)), \frac{d}{dt} \phi^\lambda(q(t)), t \right) \right|_{\lambda=0} - \frac{d}{dt} \delta F \\ &= 0. \end{aligned}$$

□

- (ii) Der Fluss transformiert auch die Zeit, d.h. er wirkt auf dem erweiterten Konfigurationsraum $\mathbb{R}^f \times \mathbb{R} \ni (q, t)$,

$$(q, t) \mapsto \psi^\lambda(q, t) = (\phi^\lambda(q, t), \tau^\lambda(q, t)), \quad (5.6.12)$$

wobei das zugehörige Vektorfeld gerade $(v(q, t), \delta\tau(q, t))$ ist. Eine Kurve $q(t)$ mit $t_1 \leq t \leq t_2$ wird dann auf $q^\lambda(\cdot)$ abgebildet, wobei

$$q^\lambda(\tau^\lambda(q(t), t)) = \phi^\lambda(q(t), t). \quad (5.6.13)$$

Die relevante Invarianzeigenschaft der Lagrangefunktion (statt (5.6.10)) ist in diesem Fall

$$L\left(q^\lambda(\tau), \frac{d}{d\tau}q^\lambda(\tau), \tau\right)\Big|_{\tau=\tau^\lambda(q(t),t)} \cdot \frac{d\tau^\lambda}{dt} = L\left(q(t), \frac{dq}{dt}, t\right) + \frac{d}{dt}F(q(t), t, \lambda) \quad (5.6.14)$$

für alle $\lambda \in \mathbb{R}$ und jede Kurve $q(\cdot)$. Motiviert ist diese Bedingung dadurch, dass dann (wegen $(d\tau^\lambda/dt)dt = d\tau$)

$$\begin{aligned} S[q^\lambda] &= \int_{\tau^\lambda(t_1)}^{\tau^\lambda(t_2)} d\tau L\left(q^\lambda(\tau), \frac{d}{d\tau}q^\lambda(\tau), \tau\right) \\ &= \int_{t_1}^{t_2} dt \left\{ L\left(q(t), \frac{d}{dt}q(t), t\right) + \frac{d}{dt}F(q(t), t, \lambda) \right\} = S[q] + F(q, t, \lambda)\Big|_{(1)}^{(2)} \end{aligned}$$

gilt. Insbesondere werden unter (5.6.13) stationäre Punkte der Wirkung (d.h. die physikalischen Lösungen) in solche übergeführt. Die zugehörige Erhaltungsgrösse ist dann

$$K = \langle p, v(q) \rangle - (\langle p, \dot{q} \rangle - L)\delta\tau - \delta F, \quad (5.6.15)$$

wobei $\delta F = \frac{dF}{d\lambda}\Big|_{\lambda=0}$.

Bemerkung: Mit $\delta\tau = 0$ folgt (i).

Beweis. Dies folgt aus denselben Argumenten wie zuvor, wobei man jedoch berücksichtigen muss, dass jetzt

$$\begin{aligned} v(q) &= \frac{d}{d\lambda}q^\lambda(\tau^\lambda)\Big|_{\lambda=0} \\ &= \frac{\partial}{\partial\lambda}q^\lambda(\tau)\Big|_{\lambda=0} + \frac{\partial}{\partial\tau^\lambda}q^\lambda(\tau^\lambda)\frac{\partial\tau^\lambda}{\partial\lambda}\Big|_{\lambda=0} \\ &= \frac{\partial}{\partial\lambda}q^\lambda(\tau)\Big|_{\lambda=0} + \dot{q}\delta\tau \\ \Rightarrow \frac{d}{d\lambda}q^\lambda(\tau)\Big|_{\lambda=0} &= \frac{\partial}{\partial\lambda}q^\lambda(\tau)\Big|_{\lambda=0} = v(q) - \dot{q}\delta\tau. \end{aligned}$$

Im Folgenden schreiben wir q^λ für $q^\lambda(\tau)$. Es ist daher

$$K = \frac{d}{d\lambda}\langle p, q^\lambda \rangle\Big|_{\lambda=0} + L\delta\tau - \delta F, \quad (5.6.16)$$

und die Zeitableitung von K ist (beachte, dass $L = L\left(q^\lambda(\tau), \frac{d}{d\tau}q^\lambda(\tau), \tau\right)$)¹³

$$\frac{dK}{dt} = \left(\frac{\partial L}{\partial[q^\lambda]_\alpha} \frac{d}{d\lambda}[q^\lambda]_\alpha + \frac{\partial L}{\partial[\dot{q}^\lambda]_\alpha} \frac{d}{d\lambda}[\dot{q}^\lambda]_\alpha + \frac{dL}{dt} \frac{\partial\tau^\lambda}{\partial\lambda} + L \frac{d}{d\lambda} \frac{d\tau^\lambda}{dt} - \frac{d}{d\lambda} \frac{d}{dt} F \right)\Big|_{\lambda=0}$$

¹³summiert wird über α

wobei

$$\frac{dL}{dt} = \frac{\partial L}{\partial [q^\lambda]_\alpha} \frac{d[q^\lambda]_\alpha}{d\tau^\lambda} \frac{d\tau^\lambda}{dt} + \frac{\partial L}{\partial [\dot{q}^\lambda]_\alpha} \frac{d[\dot{q}^\lambda]_\alpha}{d\tau^\lambda} \frac{d\tau^\lambda}{dt} + \frac{\partial L}{\partial \tau^\lambda} \frac{d\tau^\lambda}{dt}$$

Durch nachrechnen folgt

$$\frac{dK}{dt} = \frac{d}{d\lambda} \left(L \left(q^\lambda(\tau), \frac{d}{d\tau} q^\lambda(\tau), \tau \right) \Big|_{\tau=\tau^\lambda(q(t),t)} \cdot \frac{d\tau^\lambda}{dt} - \frac{d}{dt} F \right)_{\lambda=0} = 0$$

denn rechnen ergibt

$$\begin{aligned} & \frac{d}{d\lambda} \left(L \left(q^\lambda(\tau), \frac{d}{d\tau} q^\lambda(\tau), \tau \right) \Big|_{\tau=\tau^\lambda(q(t),t)} \cdot \frac{d\tau^\lambda}{dt} \right)_{\lambda=0} \\ &= \left[\left(\frac{\partial L}{\partial [q^\lambda]_\alpha} \frac{d[q^\lambda]_\alpha}{d\lambda} + \frac{\partial L}{\partial [\dot{q}^\lambda]_\alpha} \frac{d[\dot{q}^\lambda]_\alpha}{d\lambda} + \frac{\partial L}{\partial [q^\lambda]_\alpha} \frac{d[q^\lambda]_\alpha}{d\tau^\lambda} \frac{d\tau^\lambda}{d\lambda} + \frac{\partial L}{\partial [\dot{q}^\lambda]_\alpha} \frac{d[\dot{q}^\lambda]_\alpha}{d\tau^\lambda} \frac{d\tau^\lambda}{d\lambda} + \frac{\partial L}{\partial \tau^\lambda} \frac{d\tau^\lambda}{d\lambda} \right) \cdot \underbrace{\frac{d\tau^\lambda}{dt}}_{=1} \right. \\ & \quad \left. + L \frac{d}{d\lambda} \frac{d\tau^\lambda}{dt} \right]_{\lambda=0}. \end{aligned}$$

Wobei wir benutzt haben, dass $\frac{d\tau^\lambda}{dt} \Big|_{\lambda=0} = 1$. Vergleichen der Terme liefert die gewünschte Behauptung zusammen mit der Invarianz der Lagrangefunktion (5.6.14). \square

Ein einfaches Beispiel ist eine Lagrangefunktion, die zeitunabhängig ist (autonomes System). Dann ist die Zeittranslation $\phi^\lambda = \text{id}$, $\tau^\lambda(t) = t + \lambda$ eine Symmetrie im obigen Sinn (mit $F = 0$). Die erhaltene Grösse ist dann

$$L - \langle p, \dot{q} \rangle = L - \sum_{\alpha=1}^f \frac{\partial L}{\partial \dot{q}_\alpha} \dot{q}_\alpha. \quad (5.6.17)$$

Das stimmt natürlich (bis auf ein Vorzeichen) mit der ‘Energie’ überein (siehe (5.5.7)).

5.7 Erhaltungssätze revisited

Für ein (abgeschlossenes) System, dessen Kräfte durch ein Potential beschrieben werden (Lagrangesches System) das unter Galileitransformationen invariant ist, haben wir bereits in Kapitel 1.6 10 Erhaltungsgrössen abgeleitet. Wir wollen diese nun auf die 10 kontinuierlichen Parametern der Galilei-Gruppe (1.2.3) zurückführen.

Wegen der Galielinvarianz ist die Lagrangefunktion eines solchen Systems von der Form

$$L(\underbrace{\mathbf{x}_1, \dots, \mathbf{x}_N}_{=x}; \dot{\mathbf{x}}_1, \dots, \dot{\mathbf{x}}_N) = \frac{1}{2} \sum_{i=1}^N m_i \dot{\mathbf{x}}_i^2 - V(\mathbf{x}_1, \dots, \mathbf{x}_N) \equiv T - V, \quad (5.7.1)$$

wobei

$$V(R\mathbf{x}_1 + \mathbf{a}, \dots, R\mathbf{x}_N + \mathbf{a}) = V(\mathbf{x}_1, \dots, \mathbf{x}_N), \quad (R \in SO(3), \mathbf{a} \in \mathbb{R}^3). \quad (5.7.2)$$

Die Impulse sind dann gerade durch

$$\mathbf{p}_i = \frac{\partial L}{\partial \dot{\mathbf{x}}_i} = m_i \dot{\mathbf{x}}_i \quad (5.7.3)$$

gegeben.

(i) **Zeittranslationen:** Wie wir bereits oben (mehrfach) diskutiert haben, ist die erhaltene Grösse die Energie $E = T + V$.

(ii) **Räumliche Translationen:** L ist invariant unter der gemeinsamen Translation

$$\phi^\lambda(\mathbf{x}_1, \dots, \mathbf{x}_N) = (\mathbf{x}_1 + \lambda \mathbf{a}, \dots, \mathbf{x}_N + \lambda \mathbf{a}).$$

Das zugehörige Vektorfeld ist $v(x) = (\mathbf{a}, \dots, \mathbf{a})$. Die zugehörige Erhaltungsgrösse ist

$$\sum_{i=1}^N m_i \dot{\mathbf{x}}_i \cdot \mathbf{a} \equiv \mathbf{P} \cdot \mathbf{a}. \quad (5.7.4)$$

Da \mathbf{a} beliebig ist, ist der **Gesamtimpuls \mathbf{P}** erhalten.

(iii) **Drehungen:** L ist invariant unter den Drehungen

$$\phi^\lambda(\mathbf{x}_1, \dots, \mathbf{x}_N) = (R(\mathbf{e}, \lambda) \mathbf{x}_1, \dots, R(\mathbf{e}, \lambda) \mathbf{x}_N),$$

wobei $R(\mathbf{e}, \lambda)$ die Drehung vom Winkel λ um \mathbf{e} ist. Das zugehörige Vektorfeld ist

$$v(\mathbf{x}_1, \dots, \mathbf{x}_N) = (\mathbf{e} \wedge \mathbf{x}_1, \dots, \mathbf{e} \wedge \mathbf{x}_N), \quad (5.7.5)$$

und die Erhaltungsgrösse ist daher

$$\sum_{i=1}^N m_i \dot{\mathbf{x}}_i \cdot (\mathbf{e} \wedge \mathbf{x}_i) = \mathbf{e} \cdot \sum_{i=1}^N \mathbf{x}_i \wedge m_i \dot{\mathbf{x}}_i \equiv \mathbf{e} \cdot \mathbf{L}. \quad (5.7.6)$$

Da \mathbf{e} beliebig ist, führt das zur Erhaltung des **Gesamtdrehimpuls \mathbf{L}** .

(iv) Schliesslich betrachten wir die Invarianz des Systems unter **speziellen Galilei-Transformationen**, die die Transformation in ein mit gleichförmiger Geschwindigkeit \mathbf{v} bewegtes Koordinatensystem beschreiben. Dies entspricht der Schar

$$\phi^\lambda(\mathbf{x}_1, \dots, \mathbf{x}_N, t) = (\mathbf{x}_1 + \lambda \mathbf{v} t, \dots, \mathbf{x}_N + \lambda \mathbf{v} t), \quad \tau^\lambda(t) = t. \quad (5.7.7)$$

Das zugehörige Vektorfeld ist

$$v(\mathbf{x}_1, \dots, \mathbf{x}_N) = (\mathbf{v} t, \dots, \mathbf{v} t), \quad \delta \tau = 0.$$

Die Lagrangefunktion transformiert sich dann wie

$$\begin{aligned} L(x^\lambda(t), \dot{x}^\lambda(t)) &= \frac{1}{2} \sum_{i=1}^N m_i (\dot{\mathbf{x}}_i + \lambda \mathbf{v})^2 - V(\mathbf{x}_1 + \lambda \mathbf{v}t, \dots, \mathbf{x}_N + \lambda \mathbf{v}t) \\ &= L(x(t), \dot{x}(t)) + \sum_{i=1}^N m_i \left(\lambda \dot{\mathbf{x}}_i \cdot \mathbf{v} + \frac{\lambda^2}{2} \mathbf{v}^2 \right), \end{aligned}$$

und daher gilt (5.6.14) mit

$$\begin{aligned} F(x, t, \lambda) &= \sum_{i=1}^N m_i \left(\lambda \mathbf{x}_i \cdot \mathbf{v} + \frac{\lambda^2}{2} \mathbf{v}^2 t \right), \\ \delta F(x, t) &= \sum_{i=1}^N m_i \mathbf{x}_i \cdot \mathbf{v}. \end{aligned}$$

[Die Variation $\delta F = \frac{\partial F}{\partial \lambda}$ wird bei $\lambda = 0$ ausgewertet.] Die Erhaltungsgrösse (5.6.15) ist das **Schwerpunktsintegral**

$$\sum_{i=1}^N m_i \dot{\mathbf{x}}_i \cdot \mathbf{v}t - \sum_{i=1}^N m_i \mathbf{x}_i \cdot \mathbf{v} = -(\mathbf{M}\mathbf{X} - \mathbf{P}t) \cdot \mathbf{v}. \quad (5.7.8)$$

Wie wir später sehen werden (siehe Kapitel 6.7) führt jede Erhaltungsgrösse auch umgekehrt zu einer kontinuierlichen Symmetrie.

5.8 Das Prinzip von Maupertuis

Wir betrachten nun ein System für das die kinetische und potentielle Energie von der Form

$$T = \frac{1}{2} \sum_{\alpha, \beta=1}^f g_{\alpha\beta}(q) \dot{q}_\alpha \dot{q}_\beta, \quad V \equiv V(q) \quad (5.8.1)$$

sind, wobei die kinetische Energie positiv definit ist, d.h. $T(q, \dot{q}) > 0$ für $\dot{q} \neq 0$. Für ein solches System ist die Energie

$$E = \sum_{\alpha=1}^f p_\alpha \dot{q}_\alpha - L = T + V \quad (5.8.2)$$

erhalten (vgl. Kapitel 5.4).

Im Konfigurationsraum definieren wir eine *Riemannsche Metrik* durch das Linienelement

$$ds^2 = \frac{1}{2} \sum_{\alpha, \beta} g_{\alpha\beta}(q) dq_\alpha dq_\beta. \quad (5.8.3)$$

Für jede (differenzierbare) Kurve $\tau \mapsto q(\tau)$, (wobei $\tau \in [\tau_1, \tau_2]$ ein beliebiger Kurvenparameter mit $dq/d\tau \neq 0$ ist) ist das Längenquadrat des Tangentialvektors $dq/d\tau$ durch

$$\frac{1}{2} \sum_{\alpha, \beta} g_{\alpha\beta}(q(\tau)) \frac{dq_\alpha}{d\tau} \frac{dq_\beta}{d\tau} = T\left(q(\tau), \frac{dq}{d\tau}\right) = \left(\frac{ds}{d\tau}\right)^2 \quad (5.8.4)$$

definiert. Wir wählen die positive Wurzel für $ds/d\tau > 0$; dann ist die *Bogenlänge* $s(\tau)$ bis auf eine additive Konstante eindeutig bestimmt, und die Länge der Kurve beträgt

$$s(\tau_2) - s(\tau_1) = \int_{\tau_1}^{\tau_2} d\tau \sqrt{T\left(q(\tau), \frac{dq}{d\tau}\right)}. \quad (5.8.5)$$

Diese Grösse ist unabhängig von der Wahl der Lagekoordinaten q und des Kurvenparameters τ . Insbesondere kann man wegen $ds/d\tau > 0$ auch s als Kurvenparameter benutzen; für diese Parametrisierung gilt dann natürlich

$$\frac{1}{2} \sum_{\alpha, \beta} g_{\alpha\beta}(q(s)) \frac{dq_\alpha}{ds} \frac{dq_\beta}{ds} = 1. \quad (5.8.6)$$

Es sei nun $t \mapsto q(t)$ eine Lösung der Bewegungsgleichung mit der (festen) Energie E . Die geometrische Gestalt der Kurve, die *Bahnkurve*, ist durch eine beliebige Parametrisierung der Kurve gegeben, d.h. ohne Angabe ihres zeitlichen Ablaufs. Die Bewegung des Systems längs der Bahnkurve ist dann bis auf die Wahl der Richtung (Vorzeichen von ds/dt) und einer Anfangslage bestimmt durch

$$\left(\frac{ds}{dt}\right)^2 = T(q, \dot{q}) = E - V(q) \geq 0. \quad (5.8.7)$$

Die Bahnkurven zur Energie E sind charakterisiert durch das Variationsprinzip von Maupertuis:¹⁴

$$\delta \int_{(1)}^{(2)} \sqrt{E - V(q)} ds = 0 \quad (5.8.8)$$

für jede Variation der Bahnkurve bei festen Endpunkten $q^{(i)} = q(s_i)$, $i = 1, 2$.

Um diese Behauptung zu beweisen, legen wir auf jeder Kurve der Variationsschar die Bewegung durch (5.8.7) fest. Zusammen mit der Anfangsbedingung $q = q^{(1)}$ für $t = t^{(1)} = 0$ bestimmt dies die Bewegung (d.h. $q(t)$) eindeutig. Für diese Wahl gilt dann

$$\sqrt{E - V} ds = \sqrt{T} \frac{ds}{dt} dt = T dt. \quad (5.8.9)$$

¹⁴“In beinahe allen Lehrbüchern, sogar in den besten, wird dieses Prinzip so dargestellt, dass man es nicht verstehen kann.” (C. Jacobi, Vorlesungen über Dynamik, 1842-1843). Es ist nicht beabsichtigt, mit dieser Tradition zu brechen.” [A, Seite 249].

Andererseits hängt die ‘Ankunftszeit’ $t^{(2)}$ dann von der betrachteten Kurve der Schar ab. In der Ableitung der Euler-Lagrange Gleichungen aus der Variation von

$$\int_{t_0}^{t_1} L dt$$

hatten wir jedoch angenommen, dass $q(t_0)$ und $q(t_1)$ vorgegeben sind. Wir betrachten nun eine Schar $q(t, \lambda)$, bei der auch die Endpunkte $q(t^{(i)}(\lambda), \lambda)$ von dem Variationsparameter λ abhängen. Die Variation der Endpunkte ist dann

$$\Delta t^{(i)} = \left. \frac{dt^{(i)}}{d\lambda} \right|_{\lambda=0}$$

und

$$\Delta q^{(i)} = \left. \frac{d}{d\lambda} q(t^{(i)}(\lambda), \lambda) \right|_{\lambda=0} = \delta q^{(i)} + \dot{q} \Delta t^{(i)}.$$

Damit ergibt sich

$$\begin{aligned} \delta \int_{(1)}^{(2)} L dt &= \left. \frac{d}{d\lambda} \int_{t^{(1)}(\lambda)}^{t^{(2)}(\lambda)} L(q(t, \lambda), \dot{q}(t, \lambda), t) dt \right|_{\lambda=0} \\ &= \int_{t^{(1)}}^{t^{(2)}} dt \left(\frac{\partial L}{\partial \dot{q}} \frac{\partial}{\partial \lambda} \dot{q} + \frac{\partial L}{\partial q} \frac{\partial q}{\partial \lambda} \right) + L \Delta t^{(i)} \Big|_{(1)}^{(2)} \\ &= \left\{ \frac{\partial L}{\partial \dot{q}} \delta q^{(i)} + L \Delta t^{(i)} \right\} \Big|_{(1)}^{(2)} \\ &= \left\{ \frac{\partial L}{\partial \dot{q}} \Delta q^{(i)} - \left(\frac{\partial L}{\partial \dot{q}} \dot{q} - L \right) \Delta t^{(i)} \right\} \Big|_{(1)}^{(2)} \end{aligned} \quad (5.8.10)$$

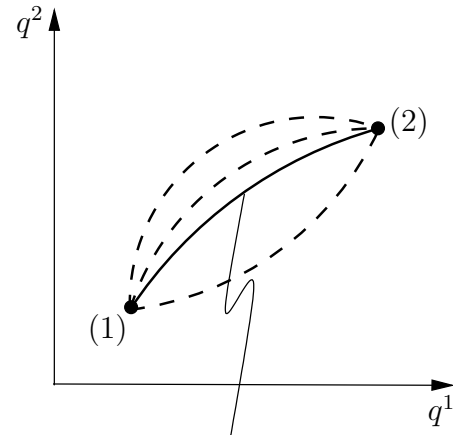
wobei wir in der dritten Zeile die Euler-Lagrange Gleichungen benützt haben (und die Summation über α der besseren Lesbarkeit halber unterdrückt haben). Wegen $\Delta q^{(i)} = 0$ fällt der erste Term weg, und der zweite Term ist gerade

$$\frac{\partial L}{\partial \dot{q}} \dot{q} - L = 2T - L = T + V = E. \quad (5.8.11)$$

Da $\Delta t \Big|_{(1)}^{(2)} = \delta \int_{(1)}^{(2)} dt$ gilt also

$$0 = \delta \int_{(1)}^{(2)} (L + E) dt = \delta \int_{(1)}^{(2)} 2T dt = \delta \int_{(1)}^{(2)} 2\sqrt{E - V} ds. \quad (5.8.12)$$

Dies beweist das Maupertuische Prinzip.



Bahnkurve zur
Energie E im Gebiet
 $E - V(q) > 0$.

Als Beispiel betrachten wir den Fall eines freien Teilchens, $V(q) \equiv 0$. Dann lautet das Variationsprinzip (5.8.8)

$$\delta \int_{(1)}^{(2)} ds = 0, \quad (5.8.13)$$

d.h. die Bahnkurven zu jeder Energie $E > 0$ sind *Geodäten* (in der Metrik (5.8.3)), d.h. Kurven extremaler Länge. Eine Illustration dazu liefert ein Massenpunkt, der reibungsfrei auf einer 2-dimensionalen Fläche im \mathbb{R}^3 gleitet. In diesem Fall ist

$$T = \frac{m}{2} \dot{\mathbf{x}}^2 = \frac{m}{2} \left(\frac{ds}{dt} \right)^2, \quad (5.8.14)$$

wobei s die durch die euklidische Metrik des \mathbb{R}^3 definierte Bogenlänge der Bahnkurve bezeichnet. Alle Bahnkurven sind also Geodäten im euklidischen Sinn auf der Fläche, z.B. die Grosskreise auf der Kugel.

Bei der Herleitung der Differentialgleichung der Geodäten ist zu beachten, dass bei der Variation $q(s, \lambda)$ (wobei λ wie oben der Scharparameter ist) nicht die Bogenlänge s_i ($i = 1, 2$), sondern die Endpunkte $q^{(i)}$ festgehalten werden. Um eine formale Übereinstimmung mit dem Hamiltonschen Prinzip zu erzielen, müssen wir statt s einen Kurvenparameter τ einführen, dessen Grenzen τ_i fest sind:

$$\delta \int_{\tau_1}^{\tau_2} \frac{ds}{d\tau} d\tau = \delta \int_{\tau_1}^{\tau_2} \underbrace{\sqrt{\frac{1}{2} \sum_{\alpha\beta} g_{\alpha\beta}(q) q'_\alpha q'_\beta}}_{=: L(q, q')} d\tau = 0, \quad (5.8.15)$$

wobei $' = d/d\tau$. Man darf bloss annehmen, dass $\tau = s$, und somit $L = 1$, auf der Geodäten ($\lambda = 0$). Die Differentialgleichung ist nun die Euler-Lagrange Gleichung für $L(q, q')$. Da die Variation für $\lambda = 0$ verschwindet, wobei $L = 1$ ist, gilt aber nun

$$\delta \int_{\tau_1}^{\tau_2} f(L) d\tau = \int_{\tau_1}^{\tau_2} f'(L) \delta L d\tau = f'(1) \delta \int_{\tau_1}^{\tau_2} L d\tau = 0$$

für jede Funktion f . Insbesondere erhalten wir dasselbe Variationsprinzip falls wir L durch $f(L) = L^2$ ersetzen, d.h. durch

$$L(q, q') = \frac{1}{2} \sum_{\alpha\beta} g_{\alpha\beta}(q) q'_\alpha q'_\beta.$$

Also lautet die *Geodätengleichung* (mit $_{,\beta} = \partial/\partial q_\beta$):

$$\begin{aligned} 0 &= \frac{d}{d\tau} \frac{\partial L}{\partial q'_\beta} - \frac{\partial L}{\partial q_\beta} = \frac{d}{d\tau} \sum_{\alpha} g_{\alpha\beta} q'_\alpha - \frac{1}{2} \sum_{\alpha\gamma} g_{\alpha\gamma,\beta} q'_\alpha q'_\gamma \\ &= \sum_{\alpha} g_{\alpha\beta} q''_\alpha + \frac{1}{2} \sum_{\alpha\gamma} (g_{\alpha\beta,\gamma} + g_{\gamma\beta,\alpha} - g_{\alpha\gamma,\beta}) q'_\alpha q'_\gamma, \end{aligned} \quad (5.8.16)$$

wobei die ersten beiden Terme in der letzten Summe explizit symmetrisiert sind.

Zum Beispiel ist für einen Massenpunkt im \mathbb{R}^3 mit

$$L = \frac{m}{2} \dot{\mathbf{x}}^2 - V(\mathbf{x}) \quad (5.8.17)$$

das Variationsprinzip

$$\delta \int_{(1)}^{(2)} ds \sqrt{E - V(\vec{x})} = 0, \quad (5.8.18)$$

wobei ds das euklidische Bogenelement bezeichnet.

Dieses Variationsprinzip ist zum Fermat'schen Prinzip analog, das wir kurz zu Beginn von Kapitel 5 angesprochen haben, wobei wir die formale Identifikation

$$n(\mathbf{x}) \propto \sqrt{E - V(\mathbf{x})} \quad (5.8.19)$$

haben. Versuchte man aber, diese Analogie zwischen geometrischer Optik und Mechanik auf die Bewegung längs dieser Bahnkurven zu erweitern, so wäre die Lichtgeschwindigkeit proportional zu

$$v = \frac{p(\mathbf{x})}{m} \equiv \sqrt{\frac{2}{m}(E - V(\mathbf{x}))} \propto n(\mathbf{x}), \quad (5.8.20)$$

wogegen die tatsächliche (Phasen-)Geschwindigkeit $c/n(\mathbf{x})$ beträgt. Dies spricht gegen eine reine Teilchennatur des Lichts. Für eine *Welle* zu fester Frequenz ν ist die Wellenlänge ortsabhängig

$$\lambda(\mathbf{x}) = \frac{c}{\nu n(\mathbf{x})} \propto \frac{1}{p(\mathbf{x})}. \quad (5.8.21)$$

Dies legt nahe, dass der Mechanik eine Wellenmechanik zugrunde liegt nach dem Schema

$$\begin{array}{ccc} \text{geometrische Optik} & \longleftrightarrow & \text{Mechanik} \\ \uparrow & & \uparrow \\ \text{Wellenoptik} & \longleftrightarrow & \text{Wellenmechanik} \end{array}$$

In Anlehnung an (5.8.21) postulierte de Broglie (1923), dass einem Teilchen mit Impuls p eine Welle der Wellenlänge λ zugeordnet ist gemäss

$$\lambda = \frac{h}{p}, \quad (h : \text{Plancksche Konstante}). \quad (5.8.22)$$

Darauf aufbauend entwickelte Schrödinger (1926) die Wellenmechanik.

6 Hamiltonsche Systeme

Für verschiedene Aspekte der Mechanik (zum Beispiel ihre Relation zur Quantenmechanik) ist eine ein wenig andere Beschreibung der Mechanik wichtig. Die Grundidee dabei ist, dass man zur Beschreibung des Systems nicht die Lagekoordinaten und ihre Geschwindigkeiten benützt, sondern statt der Geschwindigkeiten mit den konjugierten Impulsen arbeitet. Dies soll nun im Detail erklärt werden.

6.1 Die Legendretransformation

Gegeben sei ein Lagrangesches System, für das die Lagrangefunktion L von den Lagekoordinaten q_α , ihren Geschwindigkeiten \dot{q}_α sowie gegebenenfalls der Zeit t abhängt, $L \equiv L(q, \dot{q}, t)$. Die konjugierten Impulse sind durch

$$p_\alpha = \frac{\partial L}{\partial \dot{q}_\alpha} \equiv p_\alpha(q, \dot{q}, t), \quad (\alpha = 1, \dots, f) \quad (6.1.1)$$

definiert, und die verallgemeinerte Kraft ist

$$K_\alpha = \frac{\partial L}{\partial q_\alpha}. \quad (6.1.2)$$

Die Euler-Lagrange Gleichungen sind daher einfach die Newton'schen Gleichungen

$$\frac{dp_\alpha}{dt} = K_\alpha. \quad (6.1.3)$$

Sie stellen eine einfache Beziehung zwischen den Ableitungen von L nach q_α und \dot{q}_α her. Wie wir später sehen werden, lässt sich die Bewegung eines mechanischen Systems oft natürlicher im *Phasenraum* beschreiben, dessen Koordinaten die Lagekoordinaten q_α sowie die zugehörigen Impulse p_α sind. Um zu dieser Beschreibung zu gelangen, müssen wir also die Geschwindigkeiten \dot{q}_α durch p_α ausdrücken. Wir suchen dann ein Potential (als Funktion der Phasenraumkoordinaten (q_α, p_β)), aus dem sich die Bewegungsgleichungen in analoger Weise zu oben erhalten lassen.

Probleme dieser Art treten in der Physik an vielen Stellen auf (insbesondere in der Thermodynamik). Man löst sie mit Hilfe der sogenannten *Legendretransformation* (siehe Anhang A). Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ eine zweimal stetig differenzierbare Funktion mit $f'' > 0$ und sei $u = f'(x)$. Dann ist $u = f'(x)$ invertierbar (da $u' = f'' > 0$ und daher u monoton wächst), und es gilt

$$x(u) = (f')^{-1}(u). \quad (6.1.4)$$

Wir suchen nun ein Potential für x als Funktion von u , d.h. eine Funktion $g(u)$, so dass $\frac{dg}{du} = (f')^{-1}(u) = x$. Man sieht sofort, dass das gesuchte Potential durch

$$g(u) = x(u)u - f(x(u)) = x f'(x) - f \quad (6.1.5)$$

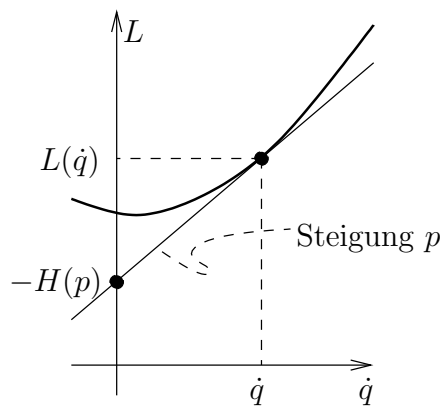
definiert ist, da dann

$$\frac{dg}{du} = \frac{dx}{du} u + x - f'(x(u)) \frac{dx}{du} = x, \quad (6.1.6)$$

da $f'(x(u)) = u$. Hängt die Funktion f noch von weiteren Variablen ab, muss man entsprechend die totalen Ableitungen durch partielle Ableitungen ersetzen.

In dem uns interessierenden Fall der Lagrangefunktion $L(= f)$ wollen wir die Legendretransformation bezüglich $x = \dot{q}_\alpha$ betrachten. Die resultierende Funktion wird die *Hamiltonfunktion* $H(q, p, t)$ genannt.

Geometrisch:



Analytisch:

Wir betrachten $x = \dot{q}$ als Funktion von $p = \frac{\partial L}{\partial \dot{q}}$

$$\dot{q} = \dot{q}(q, p, t) \quad (6.1.7)$$

und definieren dann das Analog von (6.1.5)

$$\sum_{\alpha} p_{\alpha} \dot{q}_{\alpha} - L(q, \dot{q}, t) =: H(q, p, t) \quad (6.1.8)$$

(vgl. (5.5.7)).

Das totale Differential von H ist dann

$$\begin{aligned} dH &= \sum_{\alpha} \left(\frac{\partial H}{\partial p_{\alpha}} dp_{\alpha} + \frac{\partial H}{\partial q_{\alpha}} dq_{\alpha} \right) + \frac{\partial H}{\partial t} dt \\ &= \sum_{\alpha} \left(\dot{q}_{\alpha} dp_{\alpha} + p_{\alpha} d\dot{q}_{\alpha} - \frac{\partial L}{\partial q_{\alpha}} dq_{\alpha} - \frac{\partial L}{\partial \dot{q}_{\alpha}} d\dot{q}_{\alpha} \right) - \frac{\partial L}{\partial t} dt. \end{aligned}$$

Da $\frac{\partial L}{\partial \dot{q}_{\alpha}} = p_{\alpha}$ kürzen sich der zweite und vierte Term weg, und wir erhalten

$$dH = \sum_{\alpha} \left(\dot{q}_{\alpha} dp_{\alpha} - \frac{\partial L}{\partial q_{\alpha}} dq_{\alpha} \right) - \frac{\partial L}{\partial t} dt, \quad (6.1.9)$$

was durch Koeffizientenvergleich zu

$$\frac{\partial H}{\partial p_{\alpha}} = \dot{q}_{\alpha}, \quad \frac{\partial H}{\partial q_{\alpha}} = -\frac{\partial L}{\partial q_{\alpha}}, \quad \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \quad (6.1.10)$$

führt.

6.2 Die Hamiltonschen Gleichungen

Die Bewegungsgleichungen für die Phasenraumkoordinaten können nun leicht aus den Euler-Lagrange Gleichungen $\dot{p}_\alpha = \frac{\partial L}{\partial q_\alpha}$ abgeleitet werden. Die resultierenden Hamiltonschen (kanonischen) Bewegungsgleichungen sind

$$\dot{q}_\alpha = \frac{\partial H}{\partial p_\alpha}, \quad \dot{p}_\alpha = -\frac{\partial H}{\partial q_\alpha}, \quad (\alpha = 1, \dots, f). \quad (6.2.1)$$

Dies sind $2f$ Differentialgleichungen erster Ordnung für die $2f$ unbekannt Funktionen $q_1(t), \dots, q_f(t), p_1(t), \dots, p_f(t)$.

Beispiele:

(i) Massepunkt im äusseren Potential $V(\mathbf{x})$:

$$\begin{aligned} L &= \frac{m}{2} \dot{\mathbf{x}}^2 - V(\mathbf{x}), & \mathbf{p} &= m\dot{\mathbf{x}}, \\ H &= \frac{\mathbf{p}^2}{2m} + V(\mathbf{x}). \end{aligned} \quad (6.2.2)$$

(ii) Dasselbe System in Kugelkoordinaten

$$\begin{aligned} L &= \frac{m}{2} (\dot{r}^2 + r^2 \dot{\varphi}^2) - V(r), & p_r &= m\dot{r}, & p_\varphi &= mr^2 \dot{\varphi}, \\ H &= \frac{1}{2m} (p_r^2 + r^{-2} p_\varphi^2) + V(r). \end{aligned} \quad (6.2.3)$$

(iii) Der allgemeine Fall (vgl. Kapitel 5.4)

$$\begin{aligned} L &= \frac{1}{2} \sum_{\alpha, \beta} g_{\alpha\beta}(q) \dot{q}_\alpha \dot{q}_\beta - V(q), \\ p_\alpha &= \sum_{\beta} g_{\alpha\beta}(q) \dot{q}_\beta, \\ H &= \frac{1}{2} \sum_{\alpha, \beta} g_{\alpha\beta}^{-1}(q) p_\alpha p_\beta + V(q). \end{aligned} \quad (6.2.4)$$

In diesen Fällen ist also die Hamiltonfunktion einfach die Energie (vgl. (5.5.7)). Es ist jedoch darauf zu achten, dass H eine Funktion von (q, p, t) ist, d.h. dass \dot{q} durch p ausgedrückt werden muss.

(iv) Geladenes Teilchen in äusserem elektromagnetischen Feld:

$$\begin{aligned} L &= \frac{m}{2} \dot{\mathbf{x}}^2 - e \left(\varphi(\mathbf{x}, t) - \frac{\dot{\mathbf{x}}}{c} \cdot \mathbf{A}(\mathbf{x}, t) \right), \\ \mathbf{p} &= m\dot{\mathbf{x}} + \frac{e}{c} \mathbf{A}, \\ H &= \frac{1}{2m} \left(\mathbf{p} - \frac{e}{c} \mathbf{A}(\mathbf{x}, t) \right)^2 + e\varphi(\mathbf{x}, t). \end{aligned} \quad (6.2.5)$$

[Wegen (6.3.2) kann man das auch kompakter als

$$\frac{d}{dt}F(x(t)) = \sum_k \frac{\partial F}{\partial x_k} \dot{x}_k = \sum_{i,k} \frac{\partial H}{\partial x_i} \varepsilon_{ik} \frac{\partial F}{\partial x_k} \quad (6.3.6)$$

schreiben.] Für zwei beliebige Funktionen des Phasenraums F und G definiert man nun den Ausdruck

$$\{F, G\} = \sum_{\alpha=1}^f \left(-\frac{\partial F}{\partial p_\alpha} \frac{\partial G}{\partial q_\alpha} + \frac{\partial F}{\partial q_\alpha} \frac{\partial G}{\partial p_\alpha} \right), \quad (6.3.7)$$

als die *Poissonklammer* von F und G . Kompakter lässt sich das als

$$\{F, G\} = (\nabla F)^T \varepsilon^T (\nabla G) \quad (6.3.8)$$

schreiben. Die Zeitableitung von F (entlang einer physikalischen Bahn) schreibt sich dann als Poissonklammer mit der Hamiltonfunktion (siehe Gleichung (6.3.5))

$$\frac{d}{dt}F(p(t), q(t)) = \{F, H\}(p(t), q(t)). \quad (6.3.9)$$

Besonders einfache Poissonklammern gibt es für Impuls- und Ortskoordinaten untereinander. In der Tat rechnet man leicht nach, dass die ‘kanonischen Poissonklammern’ gelten

$$\{q_\alpha, q_\beta\} = 0, \quad \{q_\alpha, p_\beta\} = \delta_{\alpha\beta}, \quad \{p_\alpha, p_\beta\} = 0, \quad (6.3.10)$$

wobei $\delta_{\alpha\beta}$ das Kroneckersymbol ist.

Die Poissonklammer gibt der Algebra \mathcal{A} der (unendlich oft differenzierbaren) Funktionen auf dem Phasenraum die Struktur einer *Poissonalgebra*. Die übliche Algebrastruktur von \mathcal{A} ist durch punktweise Addition und Multiplikation definiert

$$\begin{aligned} (\lambda F + \mu G)(x) &= \lambda F(x) + \mu G(x) \\ (F \cdot G)(x) &= F(x) \cdot G(x). \end{aligned}$$

\mathcal{A} ist daher eine kommutative Algebra, d.h. ein Vektorraum zusammen mit einem assoziativen kommutativen Produkt, das das Distributionsgesetz erfüllt.

Durch die Poissonklammer wird ein zweites, nicht-kommutatives Produkt auf \mathcal{A} eingeführt. Es ist offensichtlich anti-symmetrisch

$$\{F, G\} = -\{G, F\}, \quad (6.3.11)$$

und erfüllt die Verträglichkeitsbedingungen mit den anderen Verknüpfungen

$$\begin{aligned} \{\lambda F_1 + \mu F_2, G\} &= \lambda \{F_1, G\} + \mu \{F_2, G\} \\ \{F, G_1 \cdot G_2\} &= \{F, G_1\} \cdot G_2 + G_1 \cdot \{F, G_2\}. \end{aligned}$$

Die letzte Bedingung sagt, dass die Poissonklammer mit F eine ‘Derivation’ der Algebra darstellt. Zum besseren Verständnis dieser Eigenschaft ordnen wir F den Differentialoperator 1. Ordnung zu

$$X_F = \sum_{\alpha=1}^f \left(\frac{\partial F}{\partial q_\alpha} \frac{\partial}{\partial p_\alpha} - \frac{\partial F}{\partial p_\alpha} \frac{\partial}{\partial q_\alpha} \right), \quad (6.3.12)$$

so dass

$$\{F, G\} = X_F G. \quad (6.3.13)$$

Die Derivationseigenschaft schreibt sich dann als

$$X_F(G_1 \cdot G_2) = (X_F G_1)G_2 + G_1(X_F G_2), \quad (6.3.14)$$

was gerade die Produktregel der Differentiation ist. Das durch die Poissonklammer definierte Produkt ist nicht assoziativ; stattdessen gilt die sogenannte Jacobi-Identität

$$\{\{F_1, F_2\}, F_3\} + \{\{F_2, F_3\}, F_1\} + \{\{F_3, F_1\}, F_2\} = 0. \quad (6.3.15)$$

Mit Hilfe der Notation (6.3.6) ist der erste Term nämlich

$$\begin{aligned} \{\{F_1, F_2\}, F_3\} &= \sum_{ikjl} \frac{\partial}{\partial x_j} \left(\frac{\partial F_1}{\partial x_i} \varepsilon_{ik} \frac{\partial F_2}{\partial x_k} \right) \varepsilon_{jl} \frac{\partial F_3}{\partial x_l} \\ &= \sum_{ikjl} \left(\frac{\partial^2 F_1}{\partial x_i \partial x_j} \varepsilon_{ik} \frac{\partial F_2}{\partial x_k} \varepsilon_{jl} \frac{\partial F_3}{\partial x_l} + \frac{\partial F_1}{\partial x_i} \varepsilon_{ik} \frac{\partial^2 F_2}{\partial x_j \partial x_k} \varepsilon_{jl} \frac{\partial F_3}{\partial x_l} \right). \end{aligned}$$

Nun betrachten wir den Term., bei dem zwei Ableitungen auf F_1 wirken. Ein solcher Term taucht nur noch im dritten Term von (6.3.15) auf (aber nicht im zweiten). Die Indizes der Ableitungen sind dieselben nach der Umbenennung $ikjl \rightarrow ljik$. Der gesamte Koeffizient dieses Terms in (6.3.15) ist dann $\varepsilon_{ik} \varepsilon_{jl} + \varepsilon_{lj} \varepsilon_{ik} = 0$. Entsprechendes gilt für den Term, bei dem zwei Ableitungen auf F_2 oder F_3 wirken. Dies beweist (6.3.15).

Unter Benützung der Antisymmetrie der Poissonklammer kann man die Jacobi-Identität in Differentialoperatorsprache (6.3.14) als

$$X_F X_G H - X_G X_F H = X_{\{F,G\}} H \quad (6.3.16)$$

schreiben. Die Jacobi-Identität ist typisch für nicht-assoziative Produkte. Zum Beispiel erfüllt der Kommutator zweier $n \times n$ Matrizen

$$[A, B] = AB - BA, \quad (6.3.17)$$

wobei das Produkt durch die Matrixmultiplikation definiert ist, die Jacobi-Identität. Eine kommutative Algebra mit einem zusätzlichen Produkt, das antisymmetrisch und bilinear ist, und sowohl die Derivationseigenschaft bezüglich des 1. Produktes besitzt als auch die Jacobi-Identität bezüglich des 2. Produktes erfüllt, nennt man eine Poissonalgebra. Die Poissonstruktur ist für die Quantisierung klassischer Systeme von grosser Bedeutung.

[Grob gesprochen wird bei der Quantisierung den Elementen von \mathcal{A} Operatoren auf einem Hilbertraum zugeordnet. Die Poissonklammer geht dabei in $i\hbar$ mal den Kommutator der entsprechenden Operatoren über.]

Eine wichtige Konsequenz der Jacobi-Identität ist, dass die Poissonklammern mit der Zeitentwicklung verträglich sind. Sei ϕ^t der von der zeit-unabhängigen Hamiltonfunktion H erzeugte Fluss auf dem Phasenraum

$$\phi^t(p, q) = (p(t), q(t)), \quad (6.3.18)$$

wobei $(p(t), q(t))$ die Hamiltonschen Gleichungen erfüllen. Gleichung (6.3.9) lässt sich schreiben als

$$\frac{d}{dt}F \circ \phi^t = \{F \circ \phi^t, H\}. \quad (6.3.19)$$

Für die Zeitabhängigkeit der Poissonklammer zweier Phasenraumfunktionen findet man

$$\frac{d}{dt}\{F \circ \phi^t, G \circ \phi^t\} = \{\{F \circ \phi^t, H\}, G \circ \phi^t\} + \{F \circ \phi^t, \{G \circ \phi^t, H\}\} = \{\{F \circ \phi^t, G \circ \phi^t\}, H\}, \quad (6.3.20)$$

wobei wir die Jacobi-Identität benützt haben. Daher erfüllt $\{F \circ \phi^t, G \circ \phi^t\}$ dieselbe (gewöhnliche) Differentialgleichung in t wie $\{F, G\} \circ \phi^t$. Weiterhin stimmen die Anfangswerte überein, und daher gilt also

$$\{F \circ \phi^t, G \circ \phi^t\} = \{F, G\} \circ \phi^t. \quad (6.3.21)$$

6.4 Kanonische Transformationen

Wie wir zu Beginn von Kapitel 6.3 gesehen haben, sind die Hamiltonschen Bewegungsgleichungen im Phasenraum von der einfachen Form

$$\sum_{k=1}^{2f} \varepsilon_{ik} \dot{x}_k = \frac{\partial H}{\partial x_i}, \quad (6.4.1)$$

wobei x summarisch die q_α and p_β bezeichnet (siehe (6.3.1)) und ε_{ij} in (6.3.2) definiert ist. In dieser Notation ist die Poissonklammer einfach durch

$$\{F, G\} = - \sum_{i,j=1}^{2f} \frac{\partial F}{\partial x_i} \varepsilon_{ij} \frac{\partial G}{\partial x_j} \quad (6.4.2)$$

gegeben. Wir wollen nun die Struktur der bijektiven Koordinatentransformationen

$$x_i = x_i(\bar{x}_1, \dots, \bar{x}_{2f}) \quad (6.4.3)$$

studieren, welche die kanonischen Gleichungen (6.4.1) forminvariant lassen. Dies soll für jede Hamiltonfunktion H gelten, d.h. die Bewegungsgleichungen

$$\sum_{k=1}^{2f} \varepsilon_{ik} \dot{\bar{x}}_k = \frac{\partial \bar{H}}{\partial \bar{x}_i}, \quad (6.4.4)$$

wobei $\bar{H}(\bar{x}) = H(x)$ sollen für beliebiges H zu (6.4.1) äquivalent sein. Transformationen mit dieser Eigenschaft werden *kanonische Transformationen* genannt.

Aus (6.4.1) und (6.4.3) folgt

$$\sum_{k,l} \varepsilon_{ik} \frac{\partial x_k}{\partial \bar{x}_l} \dot{\bar{x}}_l = \sum_l \frac{\partial \bar{H}}{\partial \bar{x}_l} \frac{\partial \bar{x}_l}{\partial x_i}. \quad (6.4.5)$$

Wir definieren

$$A_{ij}(\bar{x}) = \frac{\partial x_i}{\partial \bar{x}_j} \quad (6.4.6)$$

(Jacobi-Matrix). Es gilt

$$(A^{-1})_{ij} = \frac{\partial \bar{x}_i}{\partial x_j},$$

was sich durch Nachrechnen überprüfen lässt. Weiter gilt

$$\dot{x} = A \dot{\bar{x}} \quad \text{bzw.} \quad \dot{x}_i = \frac{\partial x_i}{\partial \bar{x}_j} \dot{\bar{x}}_j \quad (6.4.7)$$

und

$$\nabla H = (A^{-1})^T \nabla \bar{H} \quad \text{bzw.} \quad \frac{\partial H}{\partial x_i} = \frac{\partial \bar{H}}{\partial \bar{x}_k} \frac{\partial \bar{x}_k}{\partial x_i}. \quad (6.4.8)$$

Bemerkung. Es gilt

$$(A^{-1})^T = (A^T)^{-1}$$

da $AA^{-1} = \text{id} \Rightarrow (A^{-1})^T A^T = \text{id}$ also ist $(A^{-1})^T$ das Inverse vom A^T .

Aus der Bewegungsgleichung $\dot{x} = \varepsilon^T \nabla H$ und (6.4.8) und (6.4.7) folgt damit

$$A \dot{\bar{x}} = \varepsilon^T (A^{-1})^T \nabla \bar{H} \Leftrightarrow A^T \varepsilon A \dot{\bar{x}} = \nabla \bar{H}. \quad (6.4.9)$$

Damit die Transformation kanonisch ist, muss das obige äquivalent sein zu

$$\dot{\bar{x}} = \varepsilon^T \nabla \bar{H},$$

womit folgt:

Satz. Eine Transformation ist kanonisch genau dann wenn

$$A^T \varepsilon A = \varepsilon \quad \text{bzw.} \quad \sum_{j,k} A_{ji} \varepsilon_{jk} A_{kl} = \varepsilon_{il}. \quad (6.4.10)$$

Bemerkung. Das ist äquivalent zu

$$A\varepsilon A^T = \varepsilon \quad \text{bzw.} \quad \sum_{j,k} A_{ij}\varepsilon_{jk}A_{lk} = \varepsilon_{il}. \quad (6.4.11)$$

Um dies zu sehen, multiplizieren wir Gleichung (6.4.10) von links mit $\varepsilon(A^T)^{-1}$ und von rechts mit εA^T .

Diese Bedingung impliziert, dass die kanonischen Transformationen auch die Poissonklammer erhalten, d.h.

$$\{F, G\} = \{\bar{F}, \bar{G}\}, \quad \text{wobei } F(x) = \bar{F}(\bar{x}), G(x) = \bar{G}(\bar{x}). \quad (6.4.12)$$

Dies folgt mit Gleichung (6.4.11)

$$\begin{aligned} \{\bar{F}, \bar{G}\} &= - \sum_{jk} \frac{\partial \bar{F}}{\partial \bar{x}_j} \varepsilon_{jk} \frac{\partial \bar{G}}{\partial \bar{x}_k} \\ &= - \sum_{ijkl} \frac{\partial F}{\partial x_i} A_{ij} \varepsilon_{jk} A_{lk} \frac{\partial G}{\partial x_l} \\ &= \{F, G\}. \end{aligned} \quad (6.4.13)$$

Bemerkung. Obige Rechnung lässt sich mit $\nabla F = (A^{-1})^T \nabla \bar{F}$ auch schreiben als

$$\begin{aligned} \{F, G\} &= -\nabla F^T \varepsilon \nabla G \\ &= \nabla \bar{F}^T \underbrace{A^{-1} \varepsilon (A^{-1})^T}_{=\varepsilon} \nabla \bar{G} \\ &= \{\bar{F}, \bar{G}\} \end{aligned}$$

Umgekehrt definiert jede Transformation, die die Poissonklammer erhält (6.4.12), eine kanonische Transformation. Dies folgt aus obiger Rechnung:

$$-\nabla \bar{F}^T A^{-1} \varepsilon (A^{-1})^T \nabla \bar{G} = -\nabla \bar{F} \varepsilon \nabla \bar{G}$$

kann nur für alle F, G gelten wenn

$$A^{-1} \varepsilon (A^{-1})^T = \varepsilon.$$

Die linearen Abbildungen $A : \mathbb{R}^{2f} \rightarrow \mathbb{R}^{2f}$ mit der Eigenschaft $A^T \varepsilon A = \varepsilon$ heißen *symplektisch* und bilden die Gruppe $\text{Sp}(2f)$. A ist invertierbar, da

$$\underbrace{\det(A^T \varepsilon A)}_{\det(A^T) \det(\varepsilon) \det(A)} = \det(\varepsilon) \Rightarrow \det(A)^2 = 1 \quad (6.4.14)$$

(Es gilt sogar $\det A = +1$, was wir aber nicht zeigen wollen.) Dass A^{-1} symplektisch ist folgt durch invertieren von Gleichung (6.4.10). Auch die Inverse einer kanonischen Transformation ist deshalb kanonisch.

Beispiel. Wenn die Orts-Koordinaten beliebig transformiert werden, dann gibt es immer eine Transformation der Impulse, so dass die Gesamttransformation kanonisch ist:

$$q_\alpha = q_\alpha(\bar{q}^1, \dots, \bar{q}^f) \quad \text{beliebig}$$

wird kanonisch durch Ergänzung der Transformation auf den ganzen Phasenraum durch

$$\bar{p}_\alpha = \sum_\beta p_\beta \frac{\partial q_\beta}{\partial \bar{q}_\alpha}. \quad (6.4.15)$$

Dies lässt sich durch Einsetzen in $A^T \varepsilon A$ überprüfen. [Diese Bedingung wird am Ende von Kapitel 6.5 abgeleitet werden.]

Die Gruppe der kanonischen Transformationen ist jedoch wesentlich grösser, sie erlauben es auch Orts- und Impulskoordinaten zu mischen.

Beispiel. Die Transformation

$$\bar{q}_\alpha = p_\alpha, \quad \bar{p}_\alpha = -q_\alpha, \quad (\alpha = 1, \dots, f) \quad (6.4.16)$$

ist eine kanonische Transformation.

Beispiel. Ein anderes Beispiel ist der durch die Hamiltonfunktion erzeugte Fluss: wegen (6.3.21) erhält er die Poissonklammer, und definiert daher eine kanonische Transformation.

6.5 Symplektische Geometrie

Die Struktur des Phasenraums, die wir jetzt beschrieben haben, kann man noch ein wenig abstrakter charakterisieren. Der Raum der unendlich oft differenzierbaren Funktionen auf dem Phasenraum (der Raum der ‘Observablen’) bildet eine Poissonalgebra. Die Automorphismen dieser Poissonalgebra sind die kanonischen Transformationen. Die Zeitentwicklung wird durch eine 1-Parametergruppe von kanonischen Transformationen beschrieben, deren infinitesimaler Generator die Poissonklammer mit einem ausgezeichneten Element der Poissonalgebra, der Hamiltonfunktion, ist.

Statt dieser algebraischen Beschreibung der Struktur der Mechanik kann man auch die geometrische Struktur des Phasenraums zum Ausgangspunkt nehmen. Dies führt dann auf eine Beschreibung der Mechanik im Rahmen der *Symplektischen Geometrie*. Um diesen Gesichtspunkt zu erklären wollen wir zunächst der Poissonklammer eine geometrische Interpretation geben.

Jeder Phasenraumfunktion F ist mittels der Poissonklammer ein Differentialoperator X_F auf dem Phasenraum zugeordnet (siehe (6.3.12)). Andererseits können (wie aus der Differentialgeometrie bekannt) Differentialoperatoren 1. Ordnung mit Vektorfeldern identifiziert werden, und man nennt X_F das Hamiltonsche Vektorfeld zu F . Mit Hilfe der Poissonklammer können wir nun jedem Paar von Vektorfeldern ein Skalarfeld durch

$$\omega(X_F(p, q), X_G(p, q)) = \{F, G\}(p, q) \quad (6.5.1)$$

zuordnen. Diese Zuordnung ist bilinear und antisymmetrisch in den beiden Vektoren; man nennt sie eine alternierende Differentialform vom Grad 2, oder kurz eine 2-Form.

Differentialformen sind eine mathematische Präzisierung der in der Physik oft auftretenden infinitesimalen orientierten Weg-, Flächen und Volumenelemente. Sie lassen sich am einfachsten als Multilinearformen über Vektorfeldern beschreiben. Hierbei sind 0-Formen einfach gewöhnliche Funktionen, eine 1-Form ist eine Linearform auf dem Raum der Vektorfelder V , wohingegen eine 2-Form eine anti-symmetrische bilineare Funktion auf dem kartesischen Produkt $V \times V$ ist. Eine k -Form φ ist daher eine multi-lineare, total antisymmetrische Funktion

$$\chi : (X_1, \dots, X_k) \mapsto \varphi(X_1, \dots, X_k) \in \mathbb{R}, \quad (6.5.2)$$

wobei die X_1, \dots, X_k Vektorfelder sind. Ein Beispiel für eine 1-Form ist die infinitesimale Arbeit $F_1 dx^1 + F_2 dx^2 + F_3 dx^3$ in einem ortsabhängigen Kraftfeld; ein Beispiel für eine 2-Form ist der infinitesimale magnetische Fluss $B_1 dx^2 \wedge dx^3 + B_2 dx^3 \wedge dx^1 + B_3 dx^1 \wedge dx^2$.

Auf dem Raum der l -Formen ist das wedge-Produkt \wedge definiert; es ordnet einer k -Form φ und einer l -Form ψ eine $(k+l)$ -Form $\varphi \wedge \psi$ zu. Diese Zuordnung ist assoziativ und distributiv, und es gilt

$$\varphi \wedge \psi = (-1)^{kl} \psi \wedge \varphi. \quad (6.5.3)$$

Für die Multiplikation einer k -Form φ mit einer 0-Form f (einer Funktion) gilt dann

$$(f \wedge \varphi)(X_1, \dots, X_k) = f \varphi(X_1, \dots, X_k). \quad (6.5.4)$$

Das Symbol \wedge wird in diesem Fall üblicherweise weggelassen. Für k 1-Formen ψ_1, \dots, ψ_k gilt andererseits

$$(\psi_1 \wedge \dots \wedge \psi_k)(X_1, \dots, X_k) = \begin{vmatrix} \varphi_1(X_1) & \dots & \varphi_k(X_1) \\ \dots & \dots & \dots \\ \varphi_k(X_1) & \dots & \varphi_k(X_k) \end{vmatrix}, \quad (6.5.5)$$

d.h. es gleicht der Determinante der Evaluationsmatrix.

Die äussere Differentiation d ist eine Fortsetzung des auf Funktionen erklärten totalen Differentials, $dF(X) = X(F)$. Konkret definieren wir die 1-Formen dx^i , $i = 1, \dots, n$ (der zu Grunde liegende Raum habe Dimension n) durch

$$dx^i \left(\frac{\partial}{\partial x^j} \right) = \delta_j^i. \quad (6.5.6)$$

Dann ist

$$dF = \sum_{i=1}^n \frac{\partial F}{\partial x^i} dx^i. \quad (6.5.7)$$

Auf Differentialformen wird nun d so fortgesetzt, dass für das wedge-Produkt die Produktregel

$$d(\varphi \wedge \psi) = (d\varphi) \wedge \psi + (-1)^k \varphi \wedge (d\psi), \quad (\varphi \text{ ist } k\text{-Form}) \quad (6.5.8)$$

gilt. Weiterhin wird verlangt, dass das zweimalige Anwenden von d immer Null ergibt: $d^2\varphi = 0$ für alle Differentialformen φ . Hieraus ergibt sich die Formel

$$d(F_0 dF_1 \wedge \cdots \wedge dF_k) = dF_0 \wedge dF_1 \wedge \cdots \wedge dF_k, \quad (6.5.9)$$

wobei alle F_i Null-Formen (Funktionen) sind.

Eine Differentialform φ nennt man *geschlossen*, falls $d\varphi = 0$. Eine solche Differentialform ist sicherlich geschlossen, falls es eine Differentialform ψ (deren Grad um eins niedriger als jener von φ ist) gibt, so dass $\varphi = d\psi$. [Denn dann gilt $d\varphi = d^2\psi = 0$.] In diesem Fall nennt man φ eine *exakte* Form. In jedem kontrahierbaren Gebiet gilt das Lemma von Poincaré: jede geschlossene Form ist exakt. Zum Beispiel ist die oben angegebene 1-Form der infinitesimalen Arbeit genau dann geschlossen, falls die Rotation des Kraftfeldes verschwindet. In diesem Fall ist das Kraftfeld konservativ und die infinitesimale Arbeit ist minus das totale Differential der potentiellen Energie V .

Mit diesen Vorbereitungen können wir nun die durch die Poissonklammer erklärte 2-Form ω durch die Differentiale der Koordinaten des Phasenraumes ausdrücken. Da (siehe (6.3.12))

$$\frac{\partial}{\partial q_\alpha} = -X_{p_\alpha}, \quad \frac{\partial}{\partial p_\alpha} = X_{q_\alpha}, \quad (6.5.10)$$

haben wir

$$\begin{aligned} \omega \left(\frac{\partial}{\partial q_\alpha}, \frac{\partial}{\partial q_\beta} \right) &= \{p_\alpha, p_\beta\} = 0 \\ \omega \left(\frac{\partial}{\partial p_\alpha}, \frac{\partial}{\partial q_\beta} \right) &= -\{q_\alpha, p_\beta\} = -\delta_{\alpha\beta} \\ \omega \left(\frac{\partial}{\partial p_\alpha}, \frac{\partial}{\partial p_\beta} \right) &= \{q_\alpha, q_\beta\} = 0. \end{aligned}$$

Dies können wir daher als

$$\omega = - \sum_{\alpha=1}^f dp_\alpha \wedge dq_\alpha \quad (6.5.11)$$

schreiben. Insbesondere ist daher die 2-Form ω geschlossen. Ausserdem ist sie nicht-entartet in dem folgenden Sinn: falls $\omega(X, Y) = 0$ für alle Y , so folgt daraus, dass $X = 0$. Eine geschlossene nicht-entartete 2-Form nennt man *symplektisch*.

Nach einem Satz von Darboux kann man jede symplektische 2-Form lokal (d.h. in der Umgeben jedes beliebigen Punktes) in Normalform bringen, d.h. man kann (lokal) Koordinaten finden, in denen ω gerade von der Form (6.5.11) ist. Die zugehörigen Karten bilden einen *kanonischen Atlas* für den Phasenraum. Für einen solchen Atlas sind dann die Koordinatentransformationen kanonisch.

Die Benützung von Differentialformen vereinfacht die Bestimmung von kanonischen Transformationen. Zum Beispiel ist die Transformation, die durch (6.4.15) beschrieben wird, kanonisch, da dann

$$\sum_{\alpha} \bar{p}_\alpha d\bar{q}_\alpha = \sum_{\beta} p_\beta dq_\beta \quad (6.5.12)$$

impliziert, dass

$$\sum_{\alpha} d\bar{p}_{\alpha} d\bar{q}_{\alpha} = \sum_{\beta} dp_{\beta} dq_{\beta}. \quad (6.5.13)$$

6.6 Kanonische Flüsse

Seien $x = (x_1, \dots, x_{2f})$ Phasenkoordinaten auf dem Phasenraum $\Gamma = \mathbb{R}^{2f}$. Sei $\phi^{\lambda} : x \mapsto y(x, \lambda)$ ein Fluss kanonischer Abbildungen. Durch Ableiten nach λ erhalten wir das erzeugende Vektorfeld $v(x)$ (Flussrichtung),

$$\left. \frac{dy_i}{d\lambda} \right|_{\lambda=0} = v_i(y) \quad (6.6.1)$$

zur Anfangsbedingung $y(x, 0) = x$. Wir wollen die Vektorfelder $v(x)$ charakterisieren, die *kanonische* Flüsse erzeugen, d.h. Flüsse, für die die Jacobi-Matrix

$$A_{ik}(x, \lambda) = \frac{\partial y_i}{\partial x_k}(x, \lambda) \quad (6.6.2)$$

symplektisch ist für alle (x, λ) . Es gilt

$$\begin{aligned} \frac{\partial}{\partial \lambda} A_{ik}(x, \lambda) &= \frac{\partial}{\partial \lambda} \frac{\partial y_i}{\partial x_k}(x, \lambda) \\ &= \frac{\partial}{\partial x_k} \frac{\partial y_i}{\partial \lambda}(x, \lambda). \end{aligned}$$

Beachte, dass der Fluss wie in (6.6.1) definiert, an der Stelle $\lambda = 0$ ausgewertet ist. Wie in Gleichung (5.6.3) gezeigt, gilt aber

$$\frac{dy}{d\lambda}(x, \lambda) = v(y(x, \lambda))$$

für beliebiges λ . Damit haben wir für $y = y(x, \lambda)$

$$\begin{aligned} \frac{\partial}{\partial \lambda} A_{ik}(x, \lambda) &= \frac{\partial}{\partial x_k} v_i(y) \\ &= \underbrace{\frac{\partial v_i}{\partial y_l}}_{:=V_{il}(y)} \underbrace{\frac{\partial y_l}{\partial x_k}}_{:=A_{lk}} \end{aligned}$$

also in Matrixnotation

$$\frac{\partial}{\partial \lambda} A(x, \lambda) = V(y) A(x, \lambda), \quad y = y(x, \lambda). \quad (6.6.3)$$

Da $A(x, 0) = 1$ symplektisch ist, ist es auch $A(x, \lambda)$ (und damit ϕ^{λ} kanonisch), falls

$$\frac{\partial}{\partial \lambda} (A^T \varepsilon A) = 0 \quad (\text{unabhängig von } \lambda). \quad (6.6.4)$$

Mit dem vorherigem ist $A(x, \lambda)$ also symplektisch, falls

$$0 = A^T(V^T\varepsilon + \varepsilon V)A. \quad (6.6.5)$$

Da A invertierbar ist und $\varepsilon^T = -\varepsilon$ ist dies äquivalent zu

$$(\varepsilon V)^T = \varepsilon V. \quad (6.6.6)$$

Damit haben wir eine Bedingung für die Symplektizität von A welche nur noch vom Vektorfeld abhängt und nicht mehr von A selber. Dies können wir nun benutzen, um eine erzeugende Funktion F für die Transformation zu definieren. Mit

$$(\varepsilon V)_{ki} = \sum_l \frac{\partial}{\partial y_i} \varepsilon_{kl} v_l(y)$$

lautet Gleichung (6.6.6) in Komponenten

$$\frac{\partial}{\partial y_k} \underbrace{\sum_l \varepsilon_{il} v_l(y)}_{=: g_i(y)} = \frac{\partial}{\partial y_i} \underbrace{\sum_l \varepsilon_{kl} v_l(y)}_{g_k(y)}. \quad (6.6.7)$$

Dies impliziert, dass die Rotation von g verschwindet, $\nabla \wedge g = 0$. Falls das Gebiet kontrahierbar ist (falls es keine Löcher im Phasenraum gibt), folgt, dass es eine Funktion F gibt, so dass g gleich dem Gradienten von F ist

$$g = \nabla F,$$

also

$$\sum_l \varepsilon_{il} v_l(y) = \frac{\partial F}{\partial y_i}. \quad (6.6.8)$$

Dies ist äquivalent zu

$$v = \varepsilon^T \nabla F, \quad (6.6.9)$$

beziehungsweise

$$\frac{\partial y}{\partial \lambda} = \varepsilon^T \nabla F. \quad (6.6.10)$$

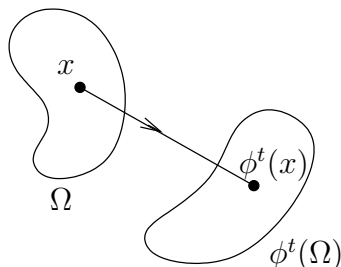
Die wichtige Aussage dieses Kapitels lässt sich wie folgt zusammenfassen. Die Differentialgleichung (6.6.10), die das Vektorfeld v definiert, haben die gleiche Form wie die Bewegungsgleichung (6.4.1), wobei λ den Platz der Zeit einnimmt und F die Hamiltonfunktion ersetzt.

Aus der Herleitung folgt auch die Umkehrung: Kanonische Gleichungen erzeugen kanonische Flüsse auf dem Phasenraum. F heisst *erzeugende Funktion* des kanonischen Flusses. Insbesondere ist die Dynamik eines *autonomen* Hamiltonschen Systems ($H = H(x)$, unabhängig von t) durch einen *kanonischen* Fluss ϕ^t gegeben:

$$x(t) = \phi^t(x_0) \quad (6.6.11)$$

ist die Lösung von (6.4.1) zur Anfangsbedingung $x(0) = x_0$. Dies hatten wir bereits zuvor (in Kapitel 6.3) gesehen.

6.6.1 Der Satz von Liouville



Satz. Das Phasenvolumen

$$\mu(\Omega) = \int_{\Omega} dx_1 \dots dx_{2f}$$

jeder Teilmenge $\Omega \subset \Gamma$ des Phasenraums ist invariant unter der Zeitentwicklung:

$$\mu(\phi^t(\Omega)) = \mu(\Omega). \quad (6.6.12)$$

Beweis. Zum Beweis berechnen wir für eine kanonische Abbildung $\phi : x \mapsto y$

$$\begin{aligned} \mu(\phi(\Omega)) &= \int_{\phi(\Omega)} dy_1 \dots dy_{2f} = \int_{\Omega} \underbrace{\left| \det \frac{\partial y_i}{\partial x_j} \right|}_{= 1, \text{ wegen (6.4.14)}} dx_1 \dots dx_{2f} = \mu(\Omega). \end{aligned} \quad (6.6.13)$$

□

Gleichung (6.6.12) besagt, dass jeder kanonische Fluss ϕ^t volumenerhaltend ist. Eine Konsequenz dieser Beobachtung ist der sogenannte Wiederkehrsatze von Poincaré.

Bemerkung. Intuitiv ist die Bedeutung des Satzes nicht trivial. Zum Beispiel können 3-Körperprobleme im Allgemeinen nur im Rahmen der Störungstheorie behandelt werden, welche versagt, wenn die Körper alle ungefähr gleich gross sind. Solche Systeme verhalten sich typischerweise chaotisch: Bei leichter Änderung der Anfangsbedingungen wird die Endsituation völlig anders aussehen. Betrachtet man einen ϵ -Ball um eine bestimmte Anfangsbedingung, enthält das Gebiet in dem die entsprechenden Endpunkte liegen, aber viele Löcher, so dass das Volumen trotzdem gleich bleibt.

6.6.2 Der Wiederkehrsatze von Poincaré

Seit ϕ^t ein *volumenerhaltender* Fluss auf dem Phasenraum Γ , und $G \subset \Gamma$ ein Gebiet. Wir nehmen an, dass

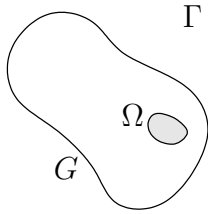
(i) G endliches Volumen hat:

$$\mu(G) < \infty \quad (6.6.14)$$

(ii) G invariant ist unter ϕ^t ,

$$\phi^t(G) \subset G, \quad (\text{alle } t). \quad (6.6.15)$$

Sei $\Omega \subset G$ eine Untermenge von G .



Definition. Einen Punkt $x \in G$ nennen wir einen *Wiederkehrpunkt* bezüglich Ω , falls seine Bahn $\phi^t(x)$ die Menge Ω für $t \rightarrow \infty$ immer wieder besucht, d.h., dass $\forall n \exists t \geq n$ so dass $\phi^t(x) \in \Omega$.

Der **Wiederkehrrsatz von Poincaré** besagt nun, dass fast alle $x \in \Omega$ Wiederkehrpunkte sind.

Satz. Die Menge der Punkte in Ω , die nicht Wiederkehrpunkte sind, ist vom Volumen Null.

Beweis. Zum Beweis definieren wir für jedes $n \in \mathbb{N}$, Ω_n durch

$$\Omega_n \equiv \bigcup_{t \geq n} \phi^{-t}(\Omega). \quad (6.6.16)$$

Falls $x \in \Omega_n$ liegt, gibt es ein $t \geq n$ so dass $\phi^t(x) \in \Omega$. Somit ist die Menge der Wiederkehrpunkte

$$W = \bigcap_{n=0,1,2,\dots} \Omega_n. \quad (6.6.17)$$

Aus (6.6.16) folgt nun

$$\Omega_0 \supset \Omega_1 \supset \Omega_2 \supset \dots \quad (6.6.18)$$

und aus (6.6.15) $\Omega_0 \subset G$, mit (6.6.14), also $\mu(\Omega_0) < \infty$. Nach einem Satz der Masstheorie (dominierte Konvergenz) ist dann

$$\mu(W) = \mu\left(\bigcap_{n=0}^{\infty} \Omega_n\right) = \lim_{n \rightarrow \infty} \mu(\Omega_n). \quad (6.6.19)$$

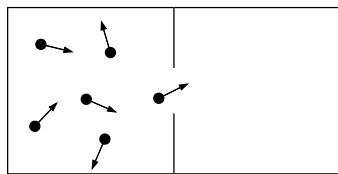
Wegen der Gruppeneigenschaft des Flusses ist aber

$$\phi^1(\Omega_n) = \bigcup_{t \geq n} \phi^{-(t-1)}(\Omega_n) = \Omega_{n-1} \quad (6.6.20)$$

und somit $\mu(\Omega_n) = \mu(\Omega_{n-1})$, da ϕ^1 volumenerhaltend ist. Daher folgt aus (6.6.19), dass

$$\mu(W) = \mu(\Omega_0). \quad (6.6.21)$$

Da aber $W \subset \Omega_0$ bedeutet dies, dass $\mu(\Omega_0 \setminus W) = 0$. Schliesslich ist aus der Definition von Ω_n klar, dass $\Omega \subseteq \Omega_0$. Daher gilt dann auch $\mu(\Omega \setminus W) = 0$. \square



System: N Teilchen mit der Hamiltonfunktion

$$H(x) = \sum_{i=1}^N \left(\frac{\mathbf{p}_i^2}{2m} + W(\mathbf{x}_i) \right) + \sum_{i < k}^{1 \dots N} V(|\mathbf{x}_i - \mathbf{x}_k|)$$

Beispiel. (Ausströmen eines Gases ins Vakuum)

$$x = (\mathbf{x}_1, \mathbf{p}_1, \dots, \mathbf{x}_N, \mathbf{p}_N).$$

Dabei beschreibt $W(\mathbf{x})$ das Potential der Wand des Gefässes ($W(\mathbf{x}) \rightarrow +\infty$ falls sich \mathbf{x} der Wand nähert) und $V(r)$ die Paarwechselwirkung zwischen den Teilchen. Die Potentiale seien nach unten beschränkt. Das Gebiet

$$G = \{x \mid H(x) \leq E\} \quad (6.6.22)$$

ist wegen der Energieerhaltung invariant unter dem durch H erzeugten Fluss ϕ^t . Ferner ist $\mu(G) < \infty$, denn für $x \in G$ sind \mathbf{p}_i^2 und $W(\mathbf{x}_i)$, ($i = 1, \dots, N$) nach oben beschränkt. $\Omega \subset G$ sei die Teilmenge der Phasenraumkonfigurationen x , bei denen sich alle Teilchen in der linken Kammer befinden. Nach dem Wiederkehrsatz sind fast alle diese x Wiederkehrpunkte: Die Bewegung $x(t)$ führt alle Teilchen gemeinsam und immer wieder in die linke Kammer zurück.

Experimentell beobachtet man dies nicht. Erstens sind die Wiederkehrzeiten sehr gross (sie verhalten sich exponentiell in der Anzahl Teilchen $\sim (\frac{1}{2})^n$). Ausserdem ist das System in Realität nicht abgeschlossen.

6.7 Erhaltungsgrössen

Wie wir in Kapitel 6.3 gesehen hatten, ist die zeitliche Änderung einer Grösse $F(x)$ längs einer tatsächlichen (physikalischen) Bahn durch

$$\frac{d}{dt}F(x(t)) = \{F, H\} \quad (6.7.1)$$

gegeben. Die Grösse F ist also genau dann erhalten, falls die Poissonklammer von F mit H verschwindet. Insbesondere ist daher H erhalten, falls es nicht explizit von t abhängt (autonomes System). Das hatten wir bereits in Kapitel 5.4 bemerkt, wo wir H mit der Energie identifiziert hatten. Seien nun F und G zwei Erhaltungsgrössen, d.h.

$$\{F, H\} = \{G, H\} = 0. \quad (6.7.2)$$

Dann folgt aus der Jacobi-Identität, dass auch $\{F, G\}$ verschwindende Poissonklammer mit H hat, und daher auch erhalten ist. Die Poissonklammer zweier Erhaltungsgrössen ist daher wiederum eine Erhaltungsgrösse (Poissonscher Satz).

Beispiel: Der Drehimpuls \mathbf{L} ist durch $\mathbf{L} = \mathbf{x} \wedge \mathbf{p}$ definiert. Sind zwei Komponenten des Drehimpulses $\mathbf{e}_1 \cdot \mathbf{L}$ und $\mathbf{e}_2 \cdot \mathbf{L}$ erhalten, so gilt das auch für die dritte Komponente, da (**Übungsaufgabe**)

$$\{\mathbf{e}_1 \cdot \mathbf{L}, \mathbf{e}_2 \cdot \mathbf{L}\} = -(\mathbf{e}_1 \wedge \mathbf{e}_2) \cdot \mathbf{L}. \quad (6.7.3)$$

Entsprechend gilt

$$\{\mathbf{e}_1 \cdot \mathbf{L}, \mathbf{e}_2 \cdot \mathbf{p}\} = -(\mathbf{e}_1 \wedge \mathbf{e}_2) \cdot \mathbf{p}. \quad (6.7.4)$$

Ist eine Impulskomponente sowie der Drehimpuls erhalten, so sind in der Tat alle Impulskomponenten erhalten.

Wir bezeichnen die von F und H im Sinn von Kapitel 6.6 (siehe insbesondere (6.6.9)) erzeugten kanonischen Flüsse durch ψ^λ bzw. ϕ^t . Dann gilt

$$\left. \frac{d}{dt} F(\phi^t(x)) \right|_{t=0} = \{F, H\} = -\{H, F\} = -\left. \frac{d}{d\lambda} H(\psi^\lambda(x)) \right|_{\lambda=0}. \quad (6.7.5)$$

Die letzte Gleichung folgt daraus, dass (6.6.10) die gleiche Form wie die Bewegungsgleichung (6.4.1) hat, wobei λ den Platz der Zeit einnimmt und F die Hamiltonfunktion ersetzt.¹⁵ F ist also genau dann eine Erhaltungsgrösse, falls H unter dem von F erzeugten kanonischen Fluss invariant ist: $H(\psi^\lambda(x)) = H(x)$. Zu jeder Erhaltungsgrösse gehört daher eine kontinuierliche Symmetrie ψ^λ von H , und umgekehrt. In gewissem Sinn ist diese Aussage also die Umkehrung von Noether's Theorem.

Beispiel: Sei \mathbb{R}^6 der Phasenraum mit Koordinaten (\mathbf{x}, \mathbf{p}) und

$$F(\mathbf{x}, \mathbf{p}) = \mathbf{e} \cdot (\mathbf{x} \wedge \mathbf{p}), \quad (6.7.6)$$

wobei \mathbf{e} ein fester Einheitsvektor ist. Die Differentialgleichungen (6.6.10) des durch F erzeugten kanonischen Flusses sind (in der Form (6.2.1))

$$\frac{d\mathbf{x}}{d\lambda} = \frac{\partial F}{\partial \mathbf{p}} = \mathbf{e} \wedge \mathbf{x}, \quad \frac{d\mathbf{p}}{d\lambda} = -\frac{\partial F}{\partial \mathbf{x}} = \mathbf{e} \wedge \mathbf{p}, \quad (6.7.7)$$

d.h. der Fluss ist

$$\psi^\lambda : (\mathbf{x}, \mathbf{p}) \mapsto (R(\lambda)\mathbf{x}, R(\lambda)\mathbf{p}), \quad (6.7.8)$$

wobei $R(\lambda)$ die Drehung im \mathbb{R}^3 um die Achse \mathbf{e} mit Winkel λ ist. Die Komponenten (6.7.6) des Drehimpulses sind also die erzeugenden Funktionen der Drehungen! Wegen (6.7.5) ist daher die allgemeine Form des Drehimpulssatzes

$$\left. \frac{d}{dt} \mathbf{e} \cdot (\mathbf{x} \wedge \mathbf{p}) \right|_{t=0} = -\left. \frac{d}{d\lambda} H(R(\lambda)\mathbf{x}, R(\lambda)\mathbf{p}) \right|_{\lambda=0}. \quad (6.7.9)$$

Insbesondere ist also $\mathbf{e} \cdot (\mathbf{x} \wedge \mathbf{p})$ genau dann erhalten, wenn H unter den Rotationen um die Achse \mathbf{e} invariant ist. Analog ist die Impulskomponente $\mathbf{e} \cdot \mathbf{p}$ die Erzeugende der Translationen $\psi^\lambda : (\mathbf{x}, \mathbf{p}) \mapsto (\mathbf{x} + \lambda\mathbf{e}, \mathbf{p})$.

¹⁵ Damit folgt aus der Herleitung von (6.3.9) auch $\{G, F\} = \left. \frac{d}{d\lambda} G(\psi^\lambda(x)) \right|_{\lambda=0}$ für jede Funktion $G(x)$.

6.8 Das Hamiltonsche Prinzip im Phasenraum

Nach der Definition der Hamiltonfunktion (6.1.8) gilt

$$L(q, \dot{q}, t) = \sum_{\alpha} p_{\alpha} \dot{q}_{\alpha} - H(q, p, t).$$

Wir erwarten daher, dass sich die Bewegungsgleichungen aus dem Variationsprinzip

$$\delta \int_{(1)}^{(2)} \left(\sum_{\alpha} p_{\alpha} \dot{q}_{\alpha} - H \right) dt = 0. \quad (6.8.1)$$

ableiten lassen, wobei an den Endpunkten t_i , $q_{\alpha}(t_i)$ vorgegeben ist. Im Phasenraum betrachten wir jedoch q_{α} und p_{β} als unabhängige Variablen, während in der ursprünglichen Formulierung des Variationsprinzips, die Variation von \dot{q} durch jene von q vorgegeben war. Es ist daher *a priori* nicht offensichtlich, dass (6.8.1) die richtigen Bewegungsgleichungen ergibt. Tatsächlich ist dies aber der Fall:

$$\begin{aligned} \delta \int_{(1)}^{(2)} \left(\sum_{\alpha} p_{\alpha} \dot{q}_{\alpha} - H \right) dt &= \int_{(1)}^{(2)} \sum_{\alpha} \left(p_{\alpha} \delta \dot{q}_{\alpha} + \dot{q}_{\alpha} \delta p_{\alpha} - \frac{\partial H}{\partial q_{\alpha}} \delta q_{\alpha} - \frac{\partial H}{\partial p_{\alpha}} \delta p_{\alpha} \right) dt \\ &= \sum_{\alpha} p_{\alpha} \underbrace{\delta q_{\alpha}}_{=0} \Big|_{(1)}^{(2)} + \int_{(1)}^{(2)} \sum_{\alpha} \left[\left(\dot{q}_{\alpha} - \frac{\partial H}{\partial p_{\alpha}} \right) \delta p_{\alpha} - \left(\dot{p}_{\alpha} + \frac{\partial H}{\partial q_{\alpha}} \right) \delta q_{\alpha} \right] dt, \end{aligned}$$

wobei wir den ersten Term auf der linken Seite der ersten Zeile partiell integriert haben. Die Randterme verschwinden. Da dies für alle δq_{α} und δp_{α} gelten muss, erhalten wir gerade die Hamiltonschen Bewegungsgleichungen (6.2.1).

6.8.1 Erzeugung kanonischer Transformationen

Als Vorbereitung auf die Diskussion der Hamilton-Jacobi Gleichung im nächsten Kapitel wollen wir nun diskutieren, wie man systematisch kanonische Transformationen konstruieren kann. Wir betrachten ein Hamiltonsches System dessen Phasenraumkoordinaten wir mit $x = (q_1, p_1, \dots, q_f, p_f)$ bezeichnen und dessen Hamiltonfunktion $H(x, t)$ ist. Wir betrachten weiterhin eine möglicherweise zeitabhängige kanonische Transformation

$$x_i = x_i(\bar{x}_1, \dots, \bar{x}_{2f}, t), \quad \bar{x}_i = \bar{x}_i(x_1, \dots, x_{2f}, t). \quad (6.8.2)$$

Da die Transformation kanonisch ist, erwarten wir, dass die Bewegungsgleichungen in den neuen Koordinaten $\bar{x} = (Q_1, P_1, \dots, Q_f, P_f)$ wiederum die kanonischen Bewegungsgleichungen sind; da die kanonische Transformation aber gegebenenfalls von der Zeit abhängt, müssen wir im allgemeinen eine neue Hamiltonfunktion $K(\bar{x}, t)$ einführen. Die Bewegungsgleichungen der beiden Systeme hängen auf diese Weise zusammen, falls die beiden Variationsprobleme

$$\delta \int_{(1)}^{(2)} \left(\sum_{\alpha} p_{\alpha} \dot{q}_{\alpha} - H \right) dt = 0 \quad \text{und} \quad \delta \int_{(1)}^{(2)} \left(\sum_{\alpha} P_{\alpha} \dot{Q}_{\alpha} - K \right) dt = 0 \quad (6.8.3)$$

Lösungen $x(t)$ bzw. $\bar{x}(t)$ besitzen, die sich unter (6.8.2) entsprechen. Hinreichend dafür ist (vgl. (5.4.23)), dass sich die beiden Integranden in (6.8.3) um eine totale Ableitung einer Funktion $S_0(q_\alpha, Q_\alpha, t)$ unterscheiden. [Da wir in der Variationsrechnung die Endpunkte δq_α bzw. δQ_α festhalten, modifiziert dann S_0 nicht die Euler-Lagrange Gleichungen.] Dies bedeutet, dass für alle Kurven $\bar{x}(t)$ gilt

$$\sum_{\alpha} p_{\alpha}(t)\dot{q}_{\alpha}(t) - H(x(t), t) = \sum_{\alpha} P_{\alpha}(t)\dot{Q}_{\alpha}(t) - K(\bar{x}(t), t) + \frac{dS_0}{dt} \quad (6.8.4)$$

wobei

$$\frac{dS_0}{dt} \equiv \sum_{\alpha} \left(\frac{\partial S_0}{\partial Q_{\alpha}} \dot{Q}_{\alpha} + \frac{\partial S_0}{\partial q_{\alpha}} \dot{q}_{\alpha} \right) + \frac{\partial S_0}{\partial t}. \quad (6.8.5)$$

Als Identität von Differentialen kann man das auch als

$$\sum_{\alpha} p_{\alpha} dq_{\alpha} - H dt = \sum_{\alpha} P_{\alpha} dQ_{\alpha} - K dt + dS_0 \quad (6.8.6)$$

schreiben. Da $P_{\alpha} dQ_{\alpha} + Q_{\alpha} dP_{\alpha} = d(Q_{\alpha} P_{\alpha})$ selbst ein totales Differential ist, ist das gleichbedeutend damit, dass

$$dS = \sum_{\alpha} (p_{\alpha} dq_{\alpha} + Q_{\alpha} dP_{\alpha}) + (K - H) dt, \quad (6.8.7)$$

wobei $S = S_0 + \sum_{\alpha} P_{\alpha} Q_{\alpha}$. Wir nehmen an, dass $(q_1, \dots, q_f, P_1, \dots, P_f)$ lokale (nicht kanonische) Koordinaten sind und fassen S auf als $S = S(q_1, \dots, q_f, P_1, \dots, P_f, t)$. Dann lautet (6.8.7)

$$(i): \quad p_{\alpha} = \frac{\partial S}{\partial q_{\alpha}}(q, P, t), \quad (ii): \quad Q_{\alpha} = \frac{\partial S}{\partial P_{\alpha}}(q, P, t), \quad (6.8.8)$$

$$(iii): \quad K = H(q, p, t) + \frac{\partial S}{\partial t}(q, P, t). \quad (6.8.9)$$

Auflösen von (ii) nach q_{α} und Einsetzen in (i, iii) liefert

$$q_{\alpha} = q_{\alpha}(Q, P, t), \quad p_{\alpha} = p_{\alpha}(Q, P, t) \quad (6.8.10)$$

(also (6.8.2)). Die dritte Gleichung (iii) bestimmt nun die gesuchte Hamiltonfunktion $K(Q, P, t)$. Wir sollten noch darauf hinweisen, dass die Bedingung für die Auflösbarkeit von S nach $q(q_0, P_0)$ ist, dass dort

$$\det \left(\frac{\partial^2 S}{\partial q_{\alpha} \partial P_{\beta}} \right) \neq 0. \quad (6.8.11)$$

$S(q, P, t)$ heisst die *erzeugende Funktion* der (zu jeder Zeit t) kanonischen Transformation (6.8.10). Kanonisch ist sie, da sie für festes $t = t_0$ auch durch $\tilde{S}(q, P) = S(q, P, t_0)$ erzeugt wird; dann verschwindet die Zeitableitung von S und daher gilt $\tilde{K}(\bar{x}) = H(x)$ — solche Transformationen sind nach Definition kanonisch!

Durch freie Wahl der Funktion S lassen sich sehr elegant kanonische Transformationen konstruieren. Um dies zu demonstrieren betrachten wir zwei Beispiele.

Beispiel 1: Sei $S = \sum_{\beta} P_{\beta} Q_{\beta}(q_1, \dots, q_f)$ wobei die Funktionen $Q_{\alpha}(q)$ vorgegeben seien. Dies definiert eine beliebige Koordinatentransformation, welche wir nun so ergänzen wollen, dass die Gesamttransformation kanonisch ist. Dann lauten (ii, i)

$$Q_{\alpha} = Q_{\alpha}(q_1, \dots, q_f), \quad p_{\alpha} = \sum_{\beta} P_{\beta} \frac{\partial Q_{\beta}}{\partial q_{\alpha}}. \quad (6.8.12)$$

Nach Auflösen der ersten Gleichung, $q_{\alpha} = q_{\alpha}(Q)$, sind dies die im Zusammenhang mit (6.4.15) besprochenen Transformationen der Lagekoordinaten. Die Auflösbarkeitsbedingung (6.8.11) ist

$$\det \left(\frac{\partial Q_{\alpha}}{\partial q_{\beta}} \right) \neq 0. \quad (6.8.13)$$

Beispiel 2: In diesem Beispiel betrachten wir eine Transformation, welche Orts- und Impulskordinaten mischt. Im ein-dimensionalen Fall ($f = 1$) betrachten wir die erzeugende Funktion $S(q, P) = \frac{1}{2} q_2 P^2$. Aus

$$Q = q_2 P, \quad p = q P^2, \quad (6.8.14)$$

ergibt sich dann die kanonische Transformation

$$q = \sqrt{\frac{Q}{P}}, \quad p = \sqrt{Q P^3}. \quad (6.8.15)$$

Bemerkungen.

(1) In (6.8.6) kann man sich S_0 auch als Funktion $S_0(q, Q, t)$ denken, so dass

$$p_{\alpha} = \frac{\partial S_0}{\partial q_{\alpha}}, \quad P_{\alpha} = -\frac{\partial S_0}{\partial Q_{\alpha}}, \quad K = H + \frac{\partial S_0}{\partial t}, \quad (6.8.16)$$

was die selbe kanonische Transformation (6.8.10) erzeugt. Analog für $\tilde{S} = S_0 - \sum_{\alpha} p_{\alpha} q_{\alpha} \equiv \tilde{S}(p, Q, t)$ usw. Die verschiedenen Typen erzeugender Funktionen gehen (bis aufs Vorzeichen) durch Legendre-Transformationen auseinander hervor.

(2) In der Sprache von Differentialformen kann man auch leicht direkt zeigen, dass (6.8.10) für festes t kanonisch ist. Aus (6.8.7) folgt nämlich wegen $d^2 = 0$, dass für festes t

$$0 = d^2 S = \sum_{\alpha} (dp_{\alpha} \wedge dq_{\alpha} + dQ_{\alpha} \wedge dP_{\alpha}), \quad \text{d.h.} \quad \sum_{\alpha} dp_{\alpha} \wedge dq_{\alpha} = \sum_{\alpha} dP_{\alpha} \wedge dQ_{\alpha}. \quad (6.8.17)$$

(3) Die erzeugende Funktion $S(q, P, t)$ ist nicht dasselbe wie die erzeugende Funktion $F(q, p)$ eines kanonischen Flusses $(q, p) = \phi^t(Q, P)$ (6.6.10). Die Begriffe sind aber verwandt: für kleine t ist

$$S(q, P, t) = \sum_{\alpha} q_{\alpha} P_{\alpha} - tF(q, P) + O(t^2), \quad (6.8.18)$$

denn (6.8.8) lautet

$$p_{\alpha} = P_{\alpha} - t \left. \frac{\partial F}{\partial q_{\alpha}} \right|_{(q, P)} + O(t^2), \quad Q_{\alpha} = q_{\alpha} - t \left. \frac{\partial F}{\partial p_{\alpha}} \right|_{(q, P)} + O(t^2), \quad (6.8.19)$$

d.h.

$$q_{\alpha} = Q_{\alpha} + t \left. \frac{\partial F}{\partial p_{\alpha}} \right|_{(Q, P)} + O(t^2), \quad p_{\alpha} = P_{\alpha} - t \left. \frac{\partial F}{\partial q_{\alpha}} \right|_{(Q, P)} + O(t^2), \quad (6.8.20)$$

was der Lösung $(q(t), p(t))$ der kanonischen Gleichungen $\dot{q}_{\alpha} = \partial F / \partial p_{\alpha}$, $\dot{p}_{\alpha} = -\partial F / \partial q_{\alpha}$ zu den Anfangsbedingungen $(q(0), p(0)) = (Q, P)$ entspricht.

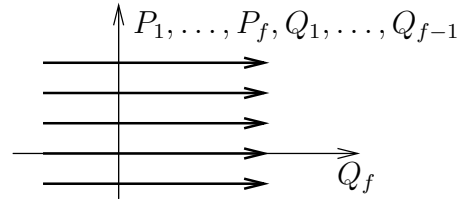
7 Die Hamilton-Jacobi Gleichung

Wir betrachten nun ein autonomes System, d.h. ein System bei dem die Hamiltonfunktion nicht explizit von t abhängt, $H = H(q_1, \dots, q_f, p_1, \dots, p_f)$. Wir wollen eine zeitunabhängige kanonische Transformation finden, so dass H gerade eine der neuen Impulskordinaten ist, z.B. $H(Q, P) = P_f$. Wenn dies gelingt, so sind die Bewegungsgleichungen in den neuen Koordinaten

$$\begin{aligned}\dot{P}_\alpha &= -\frac{\partial H}{\partial Q_\alpha} = 0, & (\alpha &= 1, \dots, f), \\ \dot{Q}_\alpha &= \frac{\partial H}{\partial P_\alpha} = 0, & (\alpha &= 1, \dots, f-1), & \dot{Q}_f &= \frac{\partial H}{\partial P_f} = 1\end{aligned}$$

trivial lösbar:

$$\begin{aligned}P_\alpha(t) &= P_\alpha(0), & (\alpha &= 1, \dots, f), \\ Q_\alpha(t) &= Q_\alpha(0), & (\alpha &= 1, \dots, f-1), \\ Q_f(t) &= Q_f(0) + t.\end{aligned}$$



7.1 Der zeitunabhängige Fall

Wie wir im vergangenen Kapitel gesehen haben muss die erzeugende Funktion $S = S(q, P)$ so bestimmt werden (siehe (6.8.9) wobei S nicht explizit von der Zeit abhängt), dass

$$H\left(q_1, \dots, q_f, \frac{\partial S}{\partial q_1}, \dots, \frac{\partial S}{\partial q_f}\right) = P_f. \quad (7.1.1)$$

Diese Gleichung wird die *zeitunabhängige Hamilton-Jacobi Gleichung* genannt. Sie ist eine (durch P_f parametrisierte) partielle Differentialgleichung 1. Ordnung für S . Die Lösung der Hamilton-Jacobi Gleichung ist also nicht eine Trajektorie, sondern die erzeugende Funktion einer Koordinatentransformation, so dass die Lösung in den neuen Koordinaten trivial ist. Auslösen der Hamilton-Jacobi-Gleichung nach S bestimmt allerdings noch nicht P_1, \dots, P_{f-1} , welche ben Damit S tatsächlich nach $Q_\alpha = \frac{\partial S}{\partial P_\alpha}$ aufgelöst werden kann, benötigt man eine Schar von Lösungen

$$S(q_1, \dots, q_f, P_1, \dots, P_f) \quad \text{mit} \quad \det\left(\frac{\partial^2 S}{\partial q_\alpha \partial P_\beta}\right) \neq 0, \quad (7.1.2)$$

eine sogenannte *vollständige Lösung*.¹⁶

Aus der vollständigen Lösung der Hamilton-Jacobi Gleichung ergibt sich die Bewegung in den ursprünglichen Koordinaten wie folgt. Zu gegebenen Werten der Erhaltungsgrößen P_1, \dots, P_f bestimmt jede der Gleichungen

$$\frac{\partial S}{\partial P_\beta}(q, P) = Q_\beta, \quad (\beta = 1, \dots, f-1) \quad (7.1.3)$$

¹⁶ $Q_\alpha(q, P) = \frac{\partial S}{\partial P_\alpha}$ ist genau dann nach q auflösbar, wenn die Jacobi-Determinante $\det\left(\frac{\partial^2 S}{\partial q_\alpha \partial P_\beta}\right)$ nicht verschwindet.

für $q = (q_1, \dots, q_f)$ eine Fläche im Konfigurationsraum, die nach (??) linear unabhängige Normalen haben. Ihr Durchschnitt ist die (1-dimensionale) Bahnkurve. Das zeitliche Durchlaufen dieser Bahnkurve ist durch

$$\frac{\partial S}{\partial P_f}(q, P) = Q_f(0) + t \quad (7.1.4)$$

bestimmt. Die $2f$ Konstanten $Q_1(0), \dots, P_f(0)$ ergeben sich aus den Anfangsbedingungen.

7.2 Separable Probleme

Falls man die Hamilton-Jacobi Gleichung (7.1.1) als

$$f\left(q_1, \frac{\partial S}{\partial q_1}\right) = F\left(q_2, \dots, q_f, \frac{\partial S}{\partial q_2}, \dots, \frac{\partial S}{\partial q_f}\right), \quad (7.2.1)$$

aufspalten kann, so heisst die Variable q_1 *separierbar*. Der Separationsansatz

$$S(q_1, \dots, q_f) = S_1(q_1) + \tilde{S}(q_2, \dots, q_f) \quad (7.2.2)$$

führt dann auf die beiden Gleichungen

$$f\left(q_1, \frac{dS_1}{dq_1}\right) = P_1, \quad F\left(q_2, \dots, q_f, \frac{\partial \tilde{S}}{\partial q_2}, \dots, \frac{\partial \tilde{S}}{\partial q_f}\right) = P_1, \quad (7.2.3)$$

wobei P_1 konstant ist, da die linke bzw. rechte Seite von (7.2.1) nicht von q_2, \dots, q_f bzw. q_1 abhängt.¹⁷ Aus (7.2.3) findet man durch Lösen der ersten Gleichung die Funktion $S_1(q_1, P_1)$. Das Problem ist vollständig separabel, falls man mit der zweiten Gleichung in gleicher Weise verfahren kann, usw. In diesem Fall ist das Resultat eine vollständige Lösung

$$S(q_1, \dots, q_f, P_1, \dots, P_{f-1}) = S_1(q_1, P_1) + S_2(q_2, P_1, P_2) + \dots + S_f(q_f, P_1 \dots P_{f-1}) \quad (7.2.4)$$

der Hamilton-Jacobi-Gleichung (7.1.1). (Jeder Term hängt noch von P_f ab.) Die Separierbarkeit eines Problems bezieht sich immer auf besondere Koordinaten $q = (q_1, \dots, q_f)$, deren Existenz eher die Ausnahme als die Regel ist. Dies soll nun an zwei Beispielen demonstriert werden.

7.2.1 Das ebene Zentralkraftproblem

Wir betrachten das ebene Zentralkraftproblem in Polarkoordinaten (siehe Beispiel (ii) in Kapitel 6.2), für das die Hamiltonfunktion durch

$$H = \frac{1}{2m} \left(p_r^2 + \frac{p_\varphi^2}{r^2} \right) + V(r) \quad (7.2.5)$$

¹⁷Die einzige Art $f(q_1) = F(q_2)$ zu erfüllen ist $f = F = konst$, da q_1 und q_2 unabhängige Variablen sind.

gegeben ist. Die Hamilton-Jacobi Gleichung ist nun

$$\frac{1}{2m} \left[\left(\frac{\partial S}{\partial r} \right)^2 + r^{-2} \left(\frac{\partial S}{\partial \varphi} \right)^2 \right] + V(r) = E, \quad (7.2.6)$$

wobei $E \equiv P_2$, bzw.

$$\left(\frac{\partial S}{\partial \varphi} \right)^2 = 2mr^2(E - V(r)) - r^2 \left(\frac{\partial S}{\partial r} \right)^2. \quad (7.2.7)$$

Diese Gleichung ist separabel: mit dem Separationsansatz $S(r, \varphi) = S_r(r) + S_\varphi(\varphi)$ müssen beide Seiten von (7.2.7) gleich einer Separationskonstanten $l^2 \equiv P_1^2$ sein. Es folgt daher, dass bis auf beliebige (irrelevante) additive Konstanten

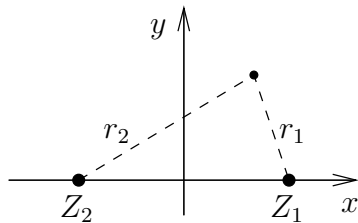
$$S_\varphi(\varphi) = l\varphi, \quad S_r(r) = \int^r ds \sqrt{2m(E - V(s)) - l^2 s^{-2}}. \quad (7.2.8)$$

Die Bahn $\varphi(r)$ sowie die zeitliche Abhängigkeit $t(r)$ ergeben sich aus (7.1.3, 7.1.4):

$$\begin{aligned} \frac{\partial S}{\partial l} &= \varphi - \int^r \frac{l}{s^2 \sqrt{2m(E - V(s)) - l^2 s^{-2}}} ds = \text{konst.}, \\ \frac{\partial S}{\partial E} &= \int^r \frac{m ds}{\sqrt{2m(E - V(s)) - l^2 s^{-2}}} = \text{konst.} + t, \end{aligned}$$

wobei die Konstanten durch die Anfangsbedingungen bestimmt sind. Dies stimmt mit (2.1.14, 2.1.12) überein.

7.2.2 Das ebene Zweizentren-Problem



Ein Massenpunkt bewegt sich in der xy -Ebene unter dem Einfluss von zwei festen Gravitationszentren bei $Z_1 = (1/2, 0)$ und $Z_2 = (-1/2, 0)$:

$$H = \frac{1}{4}(p_x^2 + p_y^2) - \frac{\mu_1}{r_1} - \frac{\mu_2}{r_2} \equiv T + V. \quad (7.2.9)$$

Dieses Problem wurde erstmals mit Hilfe der Hamilton-Jacobi Gleichung gelöst. Die Separation gelingt in elliptischen Koordinaten

$$\xi = r_1 + r_2, \quad \eta = r_1 - r_2. \quad (7.2.10)$$

Nach der Dreiecks-Ungleichung ist der Koordinatenbereich

$$\xi \geq 1, \quad -1 \leq \eta \leq +1. \quad (7.2.11)$$

Die Koordinatenlinien sind konfokale Ellipsen ($\xi = \text{konst}$) und Hyperbeln ($\eta = \text{konst}$) mit den Brennpunkten Z_1 und Z_2 . Für das Potential erhält man sofort

$$V = -\frac{\mu_1}{r_1} - \frac{\mu_2}{r_2} = -\frac{a\xi - b\eta}{\xi^2 - \eta^2}, \quad a = 2(\mu_1 + \mu_2), \quad b = 2(\mu_1 - \mu_2). \quad (7.2.12)$$

Zur Berechnung der kinetischen Energie sind weiterhin die Beziehungen

$$\xi\eta = -2x, \quad \xi^2 + \eta^2 = 4(x^2 + y^2) + 1, \quad \xi^2 - \eta^2 = 4r_1r_2 \quad (7.2.13)$$

nützlich, wobei (x, y) die ursprünglichen kartesischen Koordinaten beschreiben. Unter Benützung von

$$(\nabla r_1)^2 = (\nabla r_2)^2 = 1, \quad \nabla r_1 \cdot \nabla r_2 = \frac{4(x^2 + y^2) - 1}{4r_1r_2} \quad (7.2.14)$$

folgt

$$(\nabla \xi)^2 = 2(1 + \nabla r_1 \cdot \nabla r_2) = 4\frac{(\xi^2 - 1)}{\xi^2 - \eta^2}, \quad (\nabla \eta)^2 = 4\frac{(1 - \eta^2)}{\xi^2 - \eta^2}, \quad \nabla \xi \cdot \nabla \eta = 0. \quad (7.2.15)$$

(Die letzte Gleichung besagt, dass sich die Koordinatenlinien rechtwinklig schneiden.) Damit die durch $(x, y) \mapsto (\xi, \eta)$ definierte Transformation kanonisch ist, erhält man für die Impulse (siehe (6.4.15))

$$\mathbf{p} = (\nabla \xi)p_\xi + (\nabla \eta)p_\eta. \quad (7.2.16)$$

Damit wird also die Hamiltonfunktion

$$H(\xi, \eta, p_\xi, p_\eta) = \frac{1}{\xi^2 - \eta^2} [(\xi^2 - 1)p_\xi^2 + (1 - \eta^2)p_\eta^2 - a\xi + b\eta]. \quad (7.2.17)$$

Die Hamilton-Jacobi Gleichung

$$H\left(\xi, \eta, \frac{\partial S}{\partial \xi}, \frac{\partial S}{\partial \eta}\right) = E \quad (\equiv P_2) \quad (7.2.18)$$

lässt sich wie im ersten Beispiel mit dem Ansatz $S(\xi, \eta) = S_\xi(\xi) + S_\eta(\eta)$ separieren:

$$(\xi^2 - 1)(S'_\xi)^2 - a\xi - E\xi^2 = -(1 - \eta^2)(S'_\eta)^2 - b\eta - E\eta^2. \quad (7.2.19)$$

Beide Seiten müssen gleich einer Separationskonstante P_1 sein, und man erhält wieder eine (durch explizite Integrale gegebene) 2-parametrische Schar von Lösungen $S(\xi, \eta, P_1, E)$ der Hamilton-Jacobi Gleichung. Durch Ableitung nach P_1 und E gewinnt man die Bahnkurven und deren Zeitabhängigkeit. Durch Umschreiben auf kartesische Koordinaten der Erhaltungsgröße P_1 findet man explizit

$$-P_1 = (xp_y - yp_x)^2 + \frac{1}{4}p_x^2 - 2x\left(\frac{\mu_1}{r_1} - \frac{\mu_2}{r_2}\right). \quad (7.2.20)$$

Dass P_1 erhalten ist, lässt sich zwar auch anhand von (7.2.9) verifizieren, wohl aber kaum erraten.

7.3 Zeit-abhängige Hamilton-Jacobi Gleichung

Schliesslich betrachten wir ein (nicht autonomes) System dessen Hamiltonfunktion nun auch von der Zeit abhängt, $H \equiv H(q_1, \dots, q_f, p_1, \dots, p_f, t)$. Wir suchen nun eine *zeit-abhängige* kanonische Transformation, so dass in den neuen Koordinaten die Hamiltonfunktion $K(Q_1, \dots, Q_f, P_1, \dots, P_f, t) = 0$ ist. Dann sind Q_1, \dots, P_f konstant: die Bewegung soll auf Ruhe transformiert werden! Nach (6.8.9, 6.8.11) ist die entsprechende erzeugende Funktion $S(q, P, t)$ eine Lösung der *zeitabhängigen Hamilton-Jacobi Gleichung*

$$\frac{\partial S}{\partial t} + H\left(q, \frac{\partial S}{\partial q}, t\right) = 0, \quad (7.3.1)$$

wobei

$$\det\left(\frac{\partial^2 S}{\partial q_\alpha \partial P_\beta}\right) \neq 0. \quad (7.3.2)$$

Die Bewegung in den ursprünglichen Koordinaten ergibt sich aus den f Gleichungen

$$\frac{\partial S}{\partial P_\alpha}(q, P, t) = Q_\alpha, \quad \text{also} \quad q_\alpha = q_\alpha(Q, P, t), \quad (7.3.3)$$

wobei die Konstanten $(Q, P) = (Q_1, \dots, Q_f, P_1, \dots, P_f)$ durch die Anfangsbedingungen bestimmt sind. Einsetzen in $p_\alpha = (\partial S / \partial q_\alpha)(q, P, t)$ liefert noch $p_\alpha = p_\alpha(Q, P, t)$.

Falls das System tatsächlich autonom ist, so ist (7.3.1) äquivalent zu (7.1.1). In diesem Fall können wir den Separationsansatz

$$S(q, t) = S^*(q) + S_t(t) \quad (7.3.4)$$

machen, und erhalten

$$H\left(q, \frac{\partial S^*}{\partial q}\right) = -S'_t \equiv P_f, \quad (7.3.5)$$

wobei P_f die Separationskonstante ist (also $S_t(t) = -P_f t$). Da

$$\frac{\partial S}{\partial P_\beta} = \frac{\partial S^*}{\partial P_\beta} - \delta_{\beta f} t,$$

sind dann die Gleichungen (7.3.3) gerade (7.1.3) und (7.1.4).

8 Starre Körper

Ein starrer Körper ist ein System von Massepunkten, deren Relativabstände zwischen allen Punktepaaren während der Bewegung konstant bleiben. Obgleich das eine Idealisierung bedeutet, ist diese Konzeption sehr nützlich, und die Mechanik der Bewegung starrer Körper verdient eine ausführliche Darstellung. Das Beispiel, mit dem wir uns hauptsächlich beschäftigen werden, ist der Kreisel, bei dem ein Punkt des starren Körpers festgehalten wird.

Die Zwangsbedingungen, die den starren Körper charakterisieren sind holonom; wir können sie daher dadurch implementieren, dass wir direkt in dem tatsächlichen Konfigurationsraum des Systems, in dem dann keine Zwangsbedingungen mehr auftreten, arbeiten. Man kann sich leicht davon überzeugen, dass dieser Konfigurationsraum 6-dimensional ist: drei Koordinaten beschreiben die Position des Schwerpunktes, wohingegen die anderen drei Koordinaten die Ausrichtung des starren Körpers relativ zu einer Standardkonfiguration beschreiben — dies wird gerade durch eine Rotation im $SO(3)$ beschrieben. Eine geschickte Beschreibung ist die durch die sogenannten *Eulerwinkel*, die nun diskutiert werden soll.

8.1 Eulerwinkel

Wir verbinden mit dem starren Körper ein körperfestes (kartesisches) y -System, wobei $\mathbf{y} = 0$ einen bestimmten Punkt des starren Körpers (zum Beispiel den Schwerpunkt) bezeichnet. Wir nehmen an, dass der Punkt $\mathbf{y} = 0$ im raumfesten Koordinatensystem dem Punkt $\mathbf{x} = 0$ entspricht. Der Massenpunkt mit Koordinaten \mathbf{y} liegt dann im raumfesten \mathbf{x} -System bei

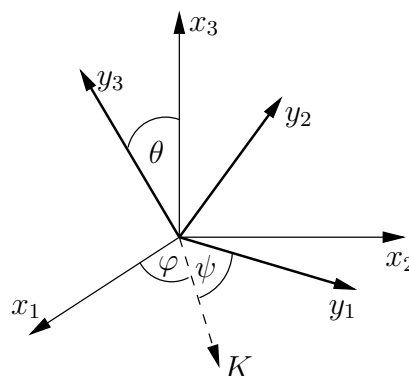
$$\mathbf{x} = R \mathbf{y}, \quad (8.1.1)$$

wobei $R \in SO(3)$ die Lage des starren Körpers angibt. Für $SO(3)$ kann man die Eulerschen Winkel (φ, θ, ψ) als Koordinaten verwenden: Das körperfeste y -System geht aus dem raumfesten x -System hervor durch die Folge von 3 Drehungen:

	Drehachse	Drehwinkel
1. Drehung	x_3	φ
2. "	K	θ
3. "	y_3	ψ

wobei die "Knotenlinie" K das Bild der x_1 -Achse unter der 1. Drehung ist. Die in (8.1.1) auftretenden Drehungen sind dann parametrisiert durch

$$\begin{aligned}
 R &= R(\varphi, \theta, \psi) \\
 &= \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \cos \psi & -\sin \psi & 0 \\ \sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (8.1.2)
 \end{aligned}$$



Die entsprechende Karte von $SO(3)$ ist das Gebiet

$$\{0 < \theta < \pi, 0 < \varphi, \psi < 2\pi\} \subset \mathbb{R}^3. \quad (8.1.3)$$

8.2 Der Trägheitstensor eines Kreisels

Wie wir oben gesehen haben, kann man eine allgemeine Auslenkung eines starren Körpers durch eine Translation und eine Drehung darstellen. Man kann daher die Bewegung eines starren Körpers in zwei getrennte Phasen aufzuspalten: die eine befasst sich mit der Translationsbewegung des Schwerpunktes, wohingegen die andere die Drehbewegung beschreibt. Oft interessiert einen jedoch die Situation, bei der ein Punkt des starren Körpers (nicht notwendigerweise der Schwerpunkt) festgehalten wird; in diesem Fall gibt es nur eine Drehbewegung um diesen festen Punkt und keine Translation. Ein starrer Körper, bei dem ein Punkt festgehalten wird, nennt man einen *Kreisel*.

Für den Fall eines Kreisels wählen wir den Ursprung des körperfesten Bezugssystems \mathbf{y} so, dass $\mathbf{y} = 0$ den festgehaltenen Punkt beschreibt. Dann gilt

$$\mathbf{x} = R(t) \mathbf{y}, \quad (8.2.1)$$

wobei für jede Zeit t , $R(t) \in SO(3)$. Im \mathbf{x} -System ist der Drehimpuls des Kreisels bezüglich des festgehaltenen Punktes $\mathbf{x} = \mathbf{0} = R(t)\mathbf{0}$

$$\mathbf{L} = \sum_i m_i \mathbf{x}_i \wedge \dot{\mathbf{x}}_i = \sum_i m_i R \mathbf{y}_i \wedge \dot{R} \mathbf{y}_i. \quad (8.2.2)$$

Wie zuvor (siehe Kapitel 1.7) schreiben wir

$$\dot{R} = R \Omega, \quad \Omega = R^t \dot{R}, \quad (8.2.3)$$

wobei $\Omega \mathbf{y} = \boldsymbol{\omega} \wedge \mathbf{y}$; $\boldsymbol{\omega}$ beschreibt also die Winkelgeschwindigkeit in \mathbf{y} -Komponenten. Daher ist der Drehimpuls im \mathbf{x} -Koordinatensystem

$$\mathbf{L} = R \sum_i m_i \mathbf{y}_i \wedge (\boldsymbol{\omega} \wedge \mathbf{y}_i) = R \sum_i m_i (\boldsymbol{\omega}(\mathbf{y}_i \cdot \mathbf{y}_i) - \mathbf{y}_i(\mathbf{y}_i \cdot \boldsymbol{\omega})). \quad (8.2.4)$$

Jede Komponente des Drehimpulses ist daher eine lineare Funktion von $\boldsymbol{\omega}$; wir definieren die Koeffizienten des Trägheitstensors durch

$$\Theta_{jk} = \sum_i m_i ((\mathbf{y}_i \cdot \mathbf{y}_i) \delta_{jk} - \mathbf{y}_i^j \mathbf{y}_i^k), \quad (8.2.5)$$

und dann gilt

$$\mathbf{L} = R \mathbf{S}, \quad \text{wobei} \quad S_j = \sum_{k=1}^3 \Theta_{jk} \omega_k. \quad (8.2.6)$$

Für kontinuierliche starre Körper ist die Summation durch eine Volumenintegration zu ersetzen; in diesem Fall ist

$$\Theta_{jk} = \int dm(\mathbf{y}) (\mathbf{y}^2 \delta_{jk} - y_j y_k). \quad (8.2.7)$$

Da die Masseverteilung im \mathbf{y} -System fest ist, hängen die Komponenten von Θ_{jk} nicht von der Zeit ab; sie charakterisieren vielmehr eine intrinsische Eigenschaft des starren Körpers. Weiterhin ist die Θ -Matrix symmetrisch, $\Theta_{jk} = \Theta_{kj}$. Θ besitzt daher drei ortho-normierte Eigenvektoren \mathbf{e}_i (Hauptträgheitsachsen) mit

$$\Theta \mathbf{e}_i = \Theta_i \mathbf{e}_i, \quad (i = 1, 2, 3). \quad (8.2.8)$$

Die zugehörigen Eigenwerte sind die Hauptträgheitsmomente Θ_i , $i = 1, 2, 3$.

Auch die kinetische Energie des Kreisels lässt sich durch den Trägheitstensor ausdrücken. Im \mathbf{x} -System ist die kinetische Energie natürlich einfach

$$T = \frac{1}{2} \int dm(\mathbf{x}) \dot{\mathbf{x}}^2. \quad (8.2.9)$$

Da $\dot{\mathbf{x}} = \dot{R}\mathbf{y}$ und das Skalarprodukt unter Rotationen invariant ist, folgt dann

$$\begin{aligned} T &= \frac{1}{2} \int dm(\mathbf{y}) (\dot{R}\mathbf{y}, \dot{R}\mathbf{y}) = \frac{1}{2} \int dm(\mathbf{y}) \underbrace{(\boldsymbol{\omega} \wedge \mathbf{y}, \boldsymbol{\omega} \wedge \mathbf{y})}_{(\boldsymbol{\omega}, \mathbf{y} \wedge (\boldsymbol{\omega} \wedge \mathbf{y}))} \\ &= \frac{1}{2} (\boldsymbol{\omega}, \mathbf{S}) = \frac{1}{2} \sum_{i,k=1}^3 \Theta_{ik} \omega_i \omega_k. \end{aligned} \quad (8.2.10)$$

Die zu Θ gehörende quadratische Form $2T \geq 0$ ist daher positiv-semidefinit, und tatsächlich positiv-definit, falls die Masseverteilung nicht entartet ist (d.h. nicht auf einer Geraden durch $\mathbf{y} = 0$ konzentriert ist). Im Folgenden betrachten wir den generischen Fall, für den Θ positiv-definit ist; dann sind die Hauptträgheitsmomente $\Theta_i > 0$ positiv. Wir werden meist in dem Hauptachsensystem arbeiten, in dem

$$\Theta = \begin{pmatrix} \Theta_1 & & \\ & \Theta_2 & \\ & & \Theta_3 \end{pmatrix}, \quad S_i = \Theta_i \omega_i, \quad T = \frac{1}{2} \sum_{i=1}^3 \Theta_i \omega_i^2. \quad (8.2.11)$$

8.3 Der freie Kiesel

Die Bewegungsgleichung des Kreisels ist der Drehimpulssatz, der besagt, dass die zeitliche Änderung des Drehimpulses das Drehmoment ist (1.6.2)

$$\dot{\mathbf{L}} = \mathbf{M}. \quad (8.3.1)$$

Im Fall des freien Kreisels ist nach Annahme $\mathbf{M} = 0$. Diese Situation ist gegeben, falls keine treibenden Kräfte einwirken (wie zum Beispiel im Weltall); auf der Erde, unter der

Einwirkung der Schwerkraft, ist der Kreisel frei, falls der feste Punkt $\mathbf{y} = 0$ gerade der Schwerpunkt ist.

Wegen (8.2.6) gilt nun

$$0 = \dot{\mathbf{L}} = \dot{R}\mathbf{S} + R\dot{\mathbf{S}}, \quad \text{also } \dot{\mathbf{S}} = -R^T \dot{R}\mathbf{S} = -\boldsymbol{\omega} \wedge \mathbf{S}. \quad (8.3.2)$$

In dem Hauptachsensystem (siehe (8.2.11)) ergeben sich daraus die *Eulerschen Gleichungen*

$$\begin{aligned} \Theta_1 \dot{\omega}_1 &= (\Theta_2 - \Theta_3) \omega_2 \omega_3, \\ \Theta_2 \dot{\omega}_2 &= (\Theta_3 - \Theta_1) \omega_3 \omega_1, \\ \Theta_3 \dot{\omega}_3 &= (\Theta_1 - \Theta_2) \omega_1 \omega_2. \end{aligned} \quad (8.3.3)$$

Diese Gleichungen definieren ein System nicht-linearer Differentialgleichungen erster Ordnung, deren Lösungen durch elliptische Funktionen darstellbar sind (wir gehen darauf hier nicht ein). Durch $\boldsymbol{\omega}(t)$ ist dann auch $\Omega(t) = R^T \dot{R}$ gegeben, und zur Bestimmung von $R(t)$ verbleibt die lineare Differentialgleichung

$$\dot{R} = R\Omega(t). \quad (8.3.4)$$

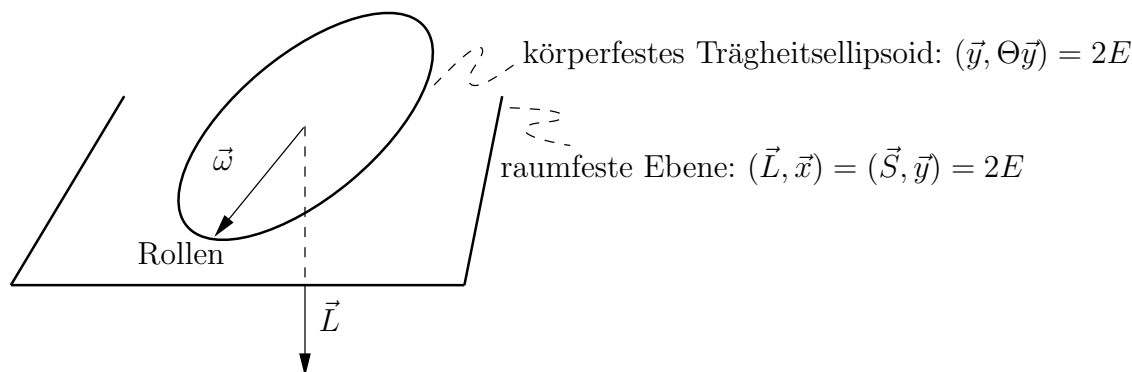
Da

$$T = \frac{1}{2} \sum_i \Theta_i \omega_i^2, \quad (8.3.5)$$

gilt

$$\begin{aligned} \dot{T} &= \sum_i \Theta_i \omega_i \dot{\omega}_i \\ &= \omega_1 \omega_2 \omega_3 [\Theta_1(\Theta_2 - \Theta_3) + \Theta_2(\Theta_3 - \Theta_1) + \Theta_3(\Theta_1 - \Theta_2)] = 0, \end{aligned}$$

und die Energie $T = E$ ist erhalten. Die Bewegungen des freien Kreisels zur Energie E kann man mit der *Konstruktion von Poinsot* geometrisch darstellen:



Für jedes R legen diese beiden Gleichungen $\boldsymbol{\omega}$ eindeutig fest: sei $\hat{\boldsymbol{\omega}}$ ein beliebiger Punkt der raumfesten Ebene. Dann zerlegen wir $\hat{\boldsymbol{\omega}} = \boldsymbol{\omega} + (\hat{\boldsymbol{\omega}} - \boldsymbol{\omega})$, und finden

$$(\hat{\boldsymbol{\omega}}, \Theta \hat{\boldsymbol{\omega}}) = \underbrace{(\boldsymbol{\omega}, \Theta \boldsymbol{\omega})}_{2E} + 2 \underbrace{(\hat{\boldsymbol{\omega}} - \boldsymbol{\omega}, \Theta \boldsymbol{\omega})}_{\mathbf{S}} + \underbrace{(\hat{\boldsymbol{\omega}} - \boldsymbol{\omega}, \Theta(\hat{\boldsymbol{\omega}} - \boldsymbol{\omega}))}_{\geq 0}. \quad (8.3.6)$$

$2E - 2E = 0$

Somit ist $(\hat{\boldsymbol{\omega}}, \Theta \hat{\boldsymbol{\omega}}) \geq 2E$ und ‘=’ nur für $\hat{\boldsymbol{\omega}} = \boldsymbol{\omega}$.

Die raumfeste Ebene ist daher die Tangentialebene am Trägheitsellipsoid im Punkt $\boldsymbol{\omega}$. Da $\boldsymbol{\omega}$ die instantane Drehachse ist, hat der Berührungspunkt des Ellipsoids die Geschwindigkeit Null, d.h. das Ellipsoid *rollt* auf der Ebene ab, ohne zu gleiten.

Eine einfache analytische Behandlung ist nur in Spezialfällen möglich; wir wollen im folgenden zwei davon beschreiben.

8.3.1 Permanente Rotationen

Eine einfache Lösung existiert, falls der Kreisel um eine seiner Hauptachsen rotiert, zum Beispiel, falls

$$\boldsymbol{\omega} = (\omega_1^0, 0, 0), \quad (8.3.7)$$

wobei ω_1^0 konstant ist. Das ist offensichtlich eine Lösung der Euler’schen Gleichungen. Um die Stabilität dieser Lösung zu untersuchen, betrachten wir in linearer Näherung $\boldsymbol{\omega} = (\omega_1^0 + \omega_1, \omega_2, \omega_3)$, d.h. wir behalten nur lineare Terme in den kleinen Grössen $\omega_1, \omega_2, \omega_3$. In dieser Approximation sind die Euler’schen Gleichungen

$$\begin{aligned} \Theta_1 \dot{\omega}_1 &= 0, \\ \Theta_2 \dot{\omega}_2 &= (\Theta_3 - \Theta_1) \omega_3 \omega_1^0, \\ \Theta_3 \dot{\omega}_3 &= (\Theta_1 - \Theta_2) \omega_1^0 \omega_2. \end{aligned} \quad (8.3.8)$$

Die erste Gleichung besagt, dass ω_1 konstant ist, und die beiden weiteren lauten in Matrixform

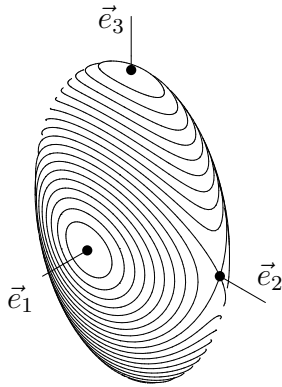
$$\begin{pmatrix} \dot{\omega}_2 \\ \dot{\omega}_3 \end{pmatrix} = \begin{pmatrix} 0 & \frac{\Theta_3 - \Theta_1}{\Theta_2} \omega_1^0 \\ \frac{\Theta_1 - \Theta_2}{\Theta_3} \omega_1^0 & 0 \end{pmatrix} \begin{pmatrix} \omega_2 \\ \omega_3 \end{pmatrix}. \quad (8.3.9)$$

Die Eigenwerte $\pm\lambda$ der 2×2 -Matrix sind gegeben durch

$$\lambda^2 = \frac{(\Theta_3 - \Theta_1)(\Theta_1 - \Theta_2)}{\Theta_2 \Theta_3} (\omega_1^0)^2 \quad (8.3.10)$$

und daher entweder beide reell, oder beide rein imaginär. Der erste Fall (mit $\lambda \neq 0$) tritt ein, falls Θ_1 zwischen Θ_2 und Θ_3 liegt, d.h. falls $\Theta_2 < \Theta_1 < \Theta_3$, oder falls $\Theta_3 < \Theta_1 < \Theta_2$. Dann hat (8.3.9) exponentiell wachsende Lösungen, d.h. die Rotation ist instabil. In den übrigen Fällen ist $\lambda^2 < 0$, die Eigenwerte rein imaginär und die Lösungen beschränkt. Dies bedeutet daher, dass von den drei Hauptachsen nur zwei stabile permanente Rotationen aufweisen, nämlich jene mit dem kleinsten und dem grössten Hauptträgheitsmoment.

Dasselbe Resultat folgt auch ausserhalb der linearen Näherung; zum Beispiel kann man das geometrisch verstehen. Wie wir oben gesehen haben, sind die kinetische Energie T und der Drehimpuls \mathbf{L} erhalten, und daher gilt



$$T = \frac{1}{2} \sum_i \Theta_i \omega_i^2 = \frac{1}{2} \sum_i \frac{S_i^2}{\Theta_i} \equiv E,$$

$$(\mathbf{L}, \mathbf{L}) = (\mathbf{S}, \mathbf{S}) = \sum_i S_i^2 \equiv l^2.$$

Damit liegt \mathbf{S} stets auf dem Schnitt eines Ellipsoids (mit Halbachsen $\sqrt{2E\Theta_i}$) mit einer Kugel vom Radius l :

Die Schnittkurven in der Nähe der Hauptachsen \mathbf{e}_1 und \mathbf{e}_3 sind kleine geschlossene Kurven, d.h. \mathbf{S} (und damit $\boldsymbol{\omega}$) bleibt in deren Nähe. Nicht so bei \mathbf{e}_2 .

[In dem Diagramm sind die Schnittkurven für den Fall $\Theta_1 < \Theta_2 < \Theta_3$ dargestellt.]

8.3.2 Der symmetrische freie Kreisel

In dem besonderen Fall eines symmetrischen Körpers ist das Trägheitsellipsoid ein Rotationsellipsoid, d.h. zwei der drei Hauptträgheitsmomente sind gleich. Ohne Beschränkung der Allgemeinheit sei $\Theta_1 = \Theta_2$. Die Euler'schen Gleichungen vereinfachen sich dann zu

$$\begin{aligned} \dot{\omega}_1 &= -\alpha \omega_2, \\ \dot{\omega}_2 &= \alpha \omega_1, \\ \dot{\omega}_3 &= 0. \end{aligned} \quad \left(\alpha = \frac{\Theta_3 - \Theta_1}{\Theta_1} \omega_3 \right), \quad (8.3.11)$$

Die letzte Gleichung impliziert, dass ω_3 konstant ist, und daher ist auch α konstant. Die Lösung der Differentialgleichung für ω_1 und ω_2 ist dann

$$\omega_1(t) + i\omega_2(t) = (\omega_1(0) + i\omega_2(0))e^{i\alpha t}. \quad (8.3.12)$$

Der Vektor $\boldsymbol{\omega}$ dreht sich also um die Rotationsachse (hier \mathbf{e}_3 , da wir angenommen haben, dass $\Theta_1 = \Theta_2$) mit konstanter Winkelgeschwindigkeit α .

Die gesamte Winkelgeschwindigkeit $\boldsymbol{\omega} = (\omega_1, \omega_2, \omega_3)$ ist daher dem Betrag nach konstant, und präzessiert um die 3-Achse — das passt natürlich gerade mit der geometrischen Beschreibung gemäss der Poinso'schen Konstruktion zusammen! Die durch α parametrisierte Präzession erfolgt relativ zu den Körperachsen, die sich selbst mit der grösseren Frequenz $\boldsymbol{\omega}$ im Raum drehen. Das Verhältnis von α zu $|\boldsymbol{\omega}|$ ist um so kleiner, je kleiner die Differenz zwischen Θ_1 und Θ_3 ist.

Man sollte erwarten, dass die Drehachse der Erde diese Präzession ausführt, denn die äusseren auf die Erde wirkenden Drehmomente sind so schwach, dass die Rotationsbewegung als kräftefrei angesehen werden kann. Die Erde ist symmetrisch um die Polachse und

an den Polen leicht abgeplattet, so dass Θ_1 kleiner als Θ_3 ist. Das numerische Verhältnis der Trägheitsmomente ist so, dass

$$\frac{\Theta_1 - \Theta_3}{\Theta_1} = -0.0033. \quad (8.3.13)$$

Die Periode der Präzessionsbewegung ist daher

$$T = \frac{2\pi}{\alpha} = \frac{2\pi}{\underbrace{\omega_3}_{=1\text{Tag}}} \frac{\Theta_1}{\underbrace{\Theta_3 - \Theta_1}_{\cong 300}} \cong 300 \text{ Tage}. \quad (8.3.14)$$

Ein Beobachter auf der Erde sollte deshalb feststellen, dass die Rotationsachse im Lauf von 10 Monaten einen Kreis um den Nordpol beschreibt. Etwas, das einem solchen Phänomen ungefähr ähnelt, ist tatsächlich beobachtet worden. Die Amplitude der Präzession ist sehr klein, die Drehachse wandert niemals mehr als etwa 4.5 m vom Nordpol weg. Die Bahn ist aber völlig unregelmässig, und die Grundperiode scheint annähernd 427 Tage und nicht 300 Tage zu sein. Die Fluktuationen werden kleinen Verschiebungen der Masseverteilung auf der Erde zugeschrieben, wie sie etwa durch atmosphärische Bewegungen hervorgerufen werden, während der Unterschied in der Periode von der Tatsache herrührt, dass die Erde nicht vollständig starr ist, sondern die elastischen Eigenschaften eines Materials wie Stahl hat. [Die kräftefreie Präzession der Erdachse darf nicht mit ihrer langsamen Präzession um die Normale zur Ekliptik verwechselt werden. Diese astronomische Präzession rührt von Graviationskräften der Sonne und des Mondes her, die in der obigen Diskussion als vernachlässigbar angesehen wurden. Diese Annahme ist berechtigt, da die Periode dieser astronomischen Präzession sehr gross (26'000 Jahre) ist.]

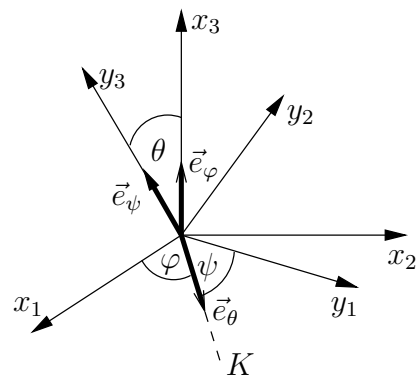
8.4 Der schwere symmetrische Kreisel

Das andere wichtige Beispiel ist der Kreisel in einem homogenen Schwerfeld. Im folgenden betrachten wir den Fall des symmetrischen Kreisels, d.h. wir nehmen wiederum an, dass $\Theta_1 = \Theta_2$. Aus Symmetriegründen liegt dann der Schwerpunkt auf der Symmetrieachse (3-Achse); in dem körperfesten Koordinatensystem hat er deshalb die Koordinaten

$$(0, 0, l), \quad l > 0. \quad (8.4.1)$$

Wir wählen das raumfeste \mathbf{x} -System so, dass die x_3 -Achse vertikal nach oben zeigt. Als Lagekoordinaten benutzen wir die Eulerschen Winkel (φ, θ, ψ) . Diese beschreiben Drehungen um die Achsen \mathbf{e}_φ , \mathbf{e}_θ , \mathbf{e}_ψ . Im körperfesten Koordinatensystem ist \mathbf{e}_ψ gerade der Einheitsvektor in der y^3 -Richtung, $\mathbf{e}_\psi = \mathbf{e}_3$; die anderen beiden Rotationsachsen sind komplizierter, aber man findet leicht geometrisch, dass

$$\begin{aligned} \mathbf{e}_\theta &= \cos \psi \mathbf{e}_1 - \sin \psi \mathbf{e}_2, \\ \mathbf{e}_\varphi &= \sin \theta \sin \psi \mathbf{e}_1 + \sin \theta \cos \psi \mathbf{e}_2 + \cos \theta \mathbf{e}_3. \end{aligned}$$



Die gesamte Winkelgeschwindigkeit (in körperfesten Komponenten) ist dann

$$\begin{aligned}\boldsymbol{\omega} &= \dot{\varphi} \mathbf{e}_\varphi + \dot{\theta} \mathbf{e}_\theta + \dot{\psi} \mathbf{e}_\psi, \\ &= \left(\dot{\varphi} \sin \theta \sin \psi + \dot{\theta} \cos \psi, \dot{\varphi} \sin \theta \cos \psi - \dot{\theta} \sin \psi, \dot{\varphi} \cos \theta + \dot{\psi} \right).\end{aligned}\quad (8.4.2)$$

Die kinetische Energie

$$T = \frac{\Theta_1}{2}(\omega_1^2 + \omega_2^2) + \frac{\Theta_3}{2}\omega_3^2 \quad (8.4.3)$$

berechnet sich daher zu

$$T = \frac{\Theta_1}{2} \left(\dot{\varphi}^2 \sin^2 \theta + \dot{\theta}^2 \right) + \frac{\Theta_3}{2} \left(\dot{\psi} + \dot{\varphi} \cos \theta \right)^2. \quad (8.4.4)$$

Dies führt schliesslich zu der Lagrangefunktion

$$L = T - V, \quad (8.4.5)$$

wobei

$$V = mgl \cos \theta. \quad (8.4.6)$$

8.4.1 Erhaltungssätze

Da L nicht von t , φ und ψ abhängt, besitzt das System die drei Erhaltungsgrössen

$$T + V = E, \quad (8.4.7)$$

$$p_\varphi = \frac{\partial L}{\partial \dot{\varphi}} = \dot{\varphi} \Theta_1 \sin^2 \theta + (\dot{\psi} + \dot{\varphi} \cos \theta) \Theta_3 \cos \theta = M_z, \quad (8.4.8)$$

$$p_\psi = \frac{\partial L}{\partial \dot{\psi}} = (\dot{\psi} + \dot{\varphi} \cos \theta) \Theta_3 = M_3. \quad (8.4.9)$$

Wir bezeichnen die entsprechenden konstanten Werte mit E , M_z und M_3 . p_φ und p_ψ sind die Projektionen des Drehimpulses auf \mathbf{e}_φ (Vertikale), bzw. auf \mathbf{e}_ψ (Symmetrieachse),

$$p_\varphi = \mathbf{S} \cdot \mathbf{e}_\varphi, \quad p_\psi = \mathbf{S} \cdot \mathbf{e}_\psi, \quad (8.4.10)$$

was leicht aus $\mathbf{S} = \Theta \boldsymbol{\omega}$ folgt,

$$\mathbf{S} = \left(\Theta_1(\dot{\varphi} \sin \theta \sin \psi + \dot{\theta} \cos \psi), \Theta_1(\dot{\varphi} \sin \theta \cos \psi - \dot{\theta} \sin \psi), \Theta_3(\dot{\varphi} \cos \theta + \dot{\psi}) \right). \quad (8.4.11)$$

Die Erhaltung von $T + V$ und p_φ gilt auch allgemeiner, falls der Kreisel nicht symmetrisch ist; p_ψ ist hingegen nur für symmetrische Kreisel erhalten.

8.4.2 Integration der Bewegungsgleichungen

Aus (8.4.8, 8.4.9) folgt

$$\dot{\psi} + \dot{\varphi} \cos \theta = \frac{M_3}{\Theta_3}, \quad \dot{\varphi} = \frac{M_z - M_3 \cos \theta}{\Theta_1 \sin^2 \theta}. \quad (8.4.12)$$

Einsetzen in (8.4.7) liefert

$$E' \equiv E - \frac{M_3^2}{2\Theta_3} = \frac{\Theta_1}{2} \dot{\theta}^2 + \frac{(M_z - M_3 \cos \theta)^2}{2\Theta_1 \sin^2 \theta} + mgl \cos \theta. \quad (8.4.13)$$

Dies ist eine Differentialgleichung 1. Ordnung für die *Nutationsbewegung* $\theta(t)$. Um sie zu lösen, führen wir die Variable $u = \cos \theta$ ein. Mit $\dot{u} = -\sin \theta \dot{\theta}$ findet man dann

$$\dot{u}^2 = (\alpha - \beta u)(1 - u^2) - (a - bu)^2 \equiv f(u), \quad (8.4.14)$$

wobei die Konstanten α , β , a und b durch

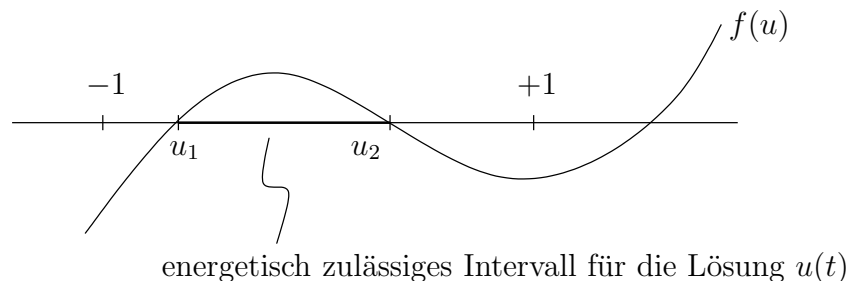
$$\alpha = \frac{2E'}{\Theta_1}, \quad \beta = \frac{2mgl}{\Theta_1} > 0, \quad a = \frac{M_z}{\Theta_1}, \quad b = \frac{M_3}{\Theta_1} \quad (8.4.15)$$

gegeben sind. Man beachte, dass u nur für $-1 \leq u \leq 1$ definiert ist. Da die linke Seite nicht-negativ ist, $\dot{u}^2 \geq 0$, ist die allgemeine Lösung beschränkt auf das Gebiet, für das $f(u) \geq 0$; dort lautet sie

$$t(u) - t(u_0) = \int_{u_0}^u \frac{dx}{\sqrt{f(x)}}. \quad (8.4.16)$$

Die Lösung ist ein elliptisches Integral, dessen Lösungen sich durch elliptische Funktionen beschreiben lassen. Man kann jedoch die allgemeine Natur der Bewegung verstehen, ohne das Integral tatsächlich ausführen zu müssen.

Die Funktion $f(u)$ ist ein Polynom 3. Grades. Die Wurzeln dieses kubischen Polynoms entsprechen den Winkeln, für die $\dot{\theta}$ sein Vorzeichen ändert, d.h. den Umkehrwinkeln von θ . Für grosse u ist βu^3 der dominierende Term; da $\beta > 0$, ist daher $f(u)$ positiv für grosse positive u , und negativ für grosse negative u .



An den Punkte $u = \pm 1$ wird $f(u) = -(a \mp b)^2$, und ist deshalb immer negativ, es sei denn, dass $u = \pm 1$ eine Wurzel ist (was wir zunächst ausschliessen wollen). Demnach

muss mindestens eine der Wurzeln in dem Gebiet $u > 1$ liegen, das keinen reellen Winkeln entspricht. Wie wir oben gesehen haben, kann die physikalische Bewegung nur in dem Bereich stattfinden, wo $f(u)$ nicht-negativ ist, und dieses Intervall muss in dem Bereich reeller Winkel, also zwischen -1 und 1 liegen. Für einen wirklichen Kreisel muss also $f(u)$ zwei Wurzeln $u_1 \leq u_2$ haben, die im Intervall $[-1, 1]$ liegen — siehe Skizze.

Das energetisch zulässige Intervall liegt zwischen u_1 und u_2 . u_1 und u_2 sind Nullstellen erster Ordnung von $f(u)$, also Umkehrpunkte der Nutationsbewegung. Diese ist periodisch mit der Periode

$$T = 2 \int_{u_1}^{u_2} \frac{dx}{\sqrt{f(x)}}. \quad (8.4.17)$$

Mit Hilfe von (8.4.12) kann man aus einer Lösung für $u(t) = \cos \theta(t)$ dann auch $\varphi(t)$ und $\psi(t)$ bestimmen. Insbesondere gilt

$$\dot{\varphi} = \frac{a - bu}{1 - u^2}, \quad (8.4.18)$$

und daher hat $\dot{\varphi}(t)$ dieselbe Periode T . Andererseits beschreibt $\varphi(t)$ die Präzession der Figurenachse um die Vertikale; durch Integration von (8.4.18) finden wir daher, dass diese Präzession die Summe einer linearen Funktion in t (mittlere Präzession) und einer periodischen Funktion der Periode T ist.

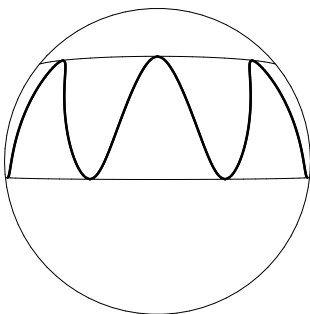
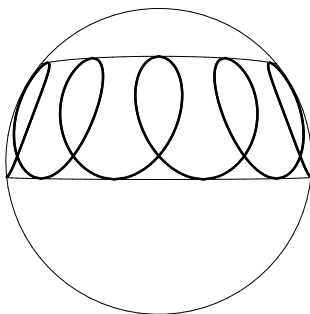
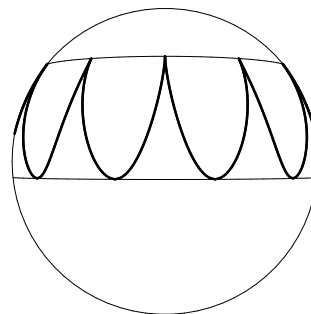
Es ist üblich, die Bewegung des Kreisels dadurch zu beschreiben, dass man die Schnittkurve der Figurenachse auf einer Kugel mit Einheitsradius um den festgehaltenen Punkt aufträgt. Diese Kurve wird der *Locus* der Figurenachse genannt. Die Polarkoordinaten eines Punktes auf dem Locus sind identisch mit den Eulerschen Winkeln θ und φ für das körperfeste System. (Der dritte Winkel, ψ , beschreibt ja nur die Rotation des Kreisels um seine eigene Symmetrieachse!) Wie wir oben gesehen haben, verläuft der Locus zwischen den zwei begrenzenden Kolatitudenkreisen $\theta_1 = \arccos u_1$ und $\theta_2 = \arccos u_2$, wobei $\dot{\theta}$ auf beiden Kreisen verschwindet. Die Gestalt der Locuskurve ist in grossem Masse durch den Wert der Wurzel von $a - bu$ bestimmt, die wir mit u' bezeichnen wollen,

$$u' \equiv \frac{b}{a}. \quad (8.4.19)$$

Nehmen wir zum Beispiel an, dass die Anfangsbedingungen so sind, dass $u' > u_2$. Dann wird $\dot{\varphi}$ wegen (8.4.18) immer das gleiche Vorzeichen für die erlaubten Neigungswinkel zwischen θ_1 und θ_2 haben. Deshalb muss der Locus der Figurenachse die begrenzenden Kreise so berühren, dass $\dot{\varphi}$ die gleiche Richtung für θ_1 und θ_2 hat — siehe Fall (a). Da φ deshalb in der einen oder anderen Richtung anwächst, sagt man, die Achse des Kreisels *präzessiert* um die vertikale Achse. Es handelt sich dabei aber nicht um die reguläre Präzession, die man im kräftefreien Fall antrifft, denn wenn die Figurenachse umläuft schwankt sie gleichzeitig zwischen den Grenzwinkeln θ_1 und θ_2 hin und her — der Kreisel führt während der Präzession Nutationen (Nickbewegungen) aus.

Hat andererseits $u' \equiv a/b$ einen Wert, der zwischen u_1 und u_2 liegt, dann ist die Richtung der Präzession an den zwei Begrenzungskreisen verschieden, und der Locus der

Figurenachse zeigt Schleifen — siehe Fall (b). Im Mittel wird $\dot{\varphi}$ jedoch im allgemeinen nicht verschwinden, so dass stets eine reine Präzession in der einen oder anderen Richtung bleibt. Es kann auch vorkommen, dass u' mit einer der Wurzeln von $f(u)$ zusammenfällt. An den entsprechenden Bindungskreisen muss dann sowohl $\dot{\theta}$ als auch $\dot{\varphi}$ verschwinden. Dies bedeutet, dass der Locus Spitzen hat, die den Kreis berrühren — siehe Fall (c).

(a) $a/b > u_2$ (b) $u_1 < a/b < u_2$ (c) $a/b = u_2$

Dieser letzte Fall ist nicht so selten, wie man vielleicht meinen möchte; er entspricht nämlich tatsächlich den Anfangsbedingungen, die man gemeinhin bei elementaren Diskussionen der Kreisel ansetzt. Dabei nimmt man an, dass sich der symmetrische Kreisel anfangs um seine Figurenachse dreht, die in einer Richtung θ_0 festgelegt ist. Zur Zeit $t = 0$ wird die Figurenachse freigelassen. Das Problem ist nun die nachfolgende Bewegung zu beschreiben. Explizit lauten also diese Anfangsbedingungen für $t = 0$ $\theta = \theta_0$, sowie $\dot{\theta} = \dot{\varphi} = 0$. Der Winkel θ_0 muss daher eine der Wurzeln von $f(u)$ sein. Tatsächlich muss er dem oberen Kreis entsprechen. Dies folgt daraus, dass E' am Anfang gleich $mgl \cos \theta_0$ ist, da $M_z = \dot{\psi} \Theta_3 \cos \theta_0 = M_3 \cos \theta_0$ ist. Andererseits sind die Terme, die $\dot{\varphi}$ und $\dot{\theta}$ enthalten immer positiv, und daher muss $mgl \cos \theta$ abnehmen, falls $\dot{\varphi}$ und/oder $\dot{\theta}$ nicht verschwinden. Der Anfangswert θ_0 ist daher gleich $\theta_0 = \theta_2$, dem kleinsten Wert, den θ annehmen kann! Wenn der Kreisel auf diese Weise freigelassen wird, beginnt er immer zu fallen und fällt weiter, bis der andere Grenzwinkel θ_1 erreicht ist. Zwischendurch präzessiert er. Die Figurenachse beginnt dann wieder gegen θ_2 anzusteigen — die vollständige Bewegung ist wie im Fall (c) angedeutet.

8.4.3 Der schnelle Kreisel

Für den Fall des ‘schnellen Kreisels’ kann man auch quantitativere Aussagen machen. Von einem schnellen Kreisel spricht man, falls die kinetische Energie der Rotation (bei der Anfangsbedingung $\dot{\varphi} = \dot{\theta} = 0$) gross ist relativ zur maximalen Änderung der potentiellen Energie,

$$\frac{1}{2} \Theta_3 \omega_3^2(0) = \frac{1}{2} \Theta_3 \dot{\psi}^2(0) \gg 2mgl. \quad (8.4.20)$$

Die Wirkung der Gravitationsmomente, nämlich die Präzession und die sie begleitende Nutation, sind dann nur kleine Störungen der vorherrschenden Rotation des Kreisels um

seine Figurenachse.

Unter diesen Anfangsbedingungen entspricht θ_0 dem oberen Begrenzungskreis, $\theta_0 = \theta_2$, bzw. $u_0 = u_2$. Das Ausmass der Nutation hängt deshalb von der Lage des anderen reellen Winkels ab, der eine Wurzel von $f(u)$ ist. Aus der Forderung $\dot{\varphi} = 0$ für $u = u_0$ folgt aus (8.4.18), dass

$$a = b u_0. \quad (8.4.21)$$

Ausserdem wissen wir, dass $f(u_0) = 0$, und daher ist auch

$$\alpha = \beta u_0. \quad (8.4.22)$$

Diese letzte Relation drückt lediglich aus, dass

$$E' = mgl \cos \theta_0, \quad (8.4.23)$$

wie wir schon zuvor gesehen hatten. Mit diesen Beziehungen können wir nun $f(u)$ einfach als

$$f(u) = (u_0 - u) [\beta(1 - u^2) - b^2(u_0 - u)] \quad (8.4.24)$$

schreiben. Die anderen beiden Wurzeln (abgesehen von u_0) sind also gerade die Wurzeln des quadratischen Polynoms in der eckigen Klammer; die gesuchte Wurzel u_1 erfüllt deshalb die Gleichung

$$(1 - u_1^2) - \frac{b^2}{\beta}(u_0 - u_1) = 0. \quad (8.4.25)$$

Bezeichnet man $x_1 \equiv u_0 - u_1$, dann hat die obige Gleichung die Form

$$x_1^2 + p x_1 - q = 0 \quad \text{mit} \quad p = \left[\frac{b^2}{\beta} - 2 \cos \theta_0 \right], \quad q = \sin^2 \theta_0. \quad (8.4.26)$$

Für den schnellen Kreisel ist der Ausdruck p in eckigen Klammern

$$\begin{aligned} \left[\frac{b^2}{\beta} - 2 \cos \theta_0 \right] &= \frac{M_3^2}{\Theta_1^2} \frac{\Theta_1}{2mgl} - 2 \cos \theta_0 \\ &= \left(\frac{\Theta_3}{\Theta_1} \right) \frac{\Theta_3 \omega_3^2}{2mgl} - 2 \cos \theta_0 \simeq \left(\frac{\Theta_3}{\Theta_1} \right) \frac{\Theta_3 \omega_3^2}{2mgl}, \end{aligned}$$

wobei wir angenommen haben, dass Θ_3/Θ_1 nicht sehr klein ist, so dass wir den zweiten Term weglassen können (unter der Annahme, dass der Kreisel ‘schnell’ ist, d.h. dass (8.4.20) gilt). Unter derselben Annahme ist dann p^2 viel grösser als $4q$, so dass die Lösungen von (8.4.26) näherungsweise lauten

$$x_1 = \frac{q}{p}, \quad -p - \frac{q}{p}. \quad (8.4.27)$$

Die zweite Lösung entspricht $u > 1$, da $x_1 = u_0 - u_1$, und $u_0 = u_2$ die grössere der beiden Lösungen mit $u < 1$ ist. Die gesuchte Lösung ist also

$$x_1 = u_0 - u_1 = \left(\frac{\Theta_1}{\Theta_3} \right) \frac{2mgl}{\Theta_3 \omega_3^2} \sin^2 \theta_0. \quad (8.4.28)$$

Diese Grösse ist daher ein Mass der Nutation: insbesondere sehen wir, dass die Nutation umso kleiner ist, je schneller sich der Kreisel dreht.

Entsprechend kann man auch die *Frequenz der Nutation* für den Fall des schnellen Kreisels bestimmen. Für einen schnellen Kreisel ist das Mass der Nutation klein, und man kann daher in (8.4.24) den Term $(1 - u^2)$ durch seinen Anfangswert $\sin^2 \theta_0$ ersetzen. Damit wird (8.4.14)

$$\dot{x}^2 = f(u) = x(\beta \sin^2 \theta_0 - b^2 x), \quad (8.4.29)$$

wobei $x = u_0 - u$. Diese Differentialgleichung kann leicht integriert werden, und die Lösung zur Anfangsbedingung $x(t = 0) = 0$ ist

$$x = \frac{x_1}{2}(1 - \cos bt), \quad (8.4.30)$$

wobei x_1 durch (8.4.28) gegeben ist. Die Winkelfrequenz der Nutation der Figurenachse zwischen $\theta_0 = \theta_2$ und θ_1 ist deshalb

$$b = \frac{\Theta_3}{\Theta_1} \omega_3. \quad (8.4.31)$$

Die Winkelgeschwindigkeit der Nutation wächst daher um so stärker, je schneller sich der Kreisel anfangs dreht.

Schliesslich ist die Winkelgeschwindigkeit der Präzession gemäss (8.4.18)

$$\dot{\varphi} = \frac{b(u_0 - u)}{1 - u^2} \simeq \frac{bx}{\sin^2 \theta_0}. \quad (8.4.32)$$

Nach Substitution von (8.4.30) und (8.4.28) führt dies zu

$$\dot{\varphi} = \frac{\beta}{2b}(1 - \cos bt). \quad (8.4.33)$$

Die Präzessionsgeschwindigkeit ist deshalb nicht gleichförmig, sondern ändert sich harmonisch mit der Zeit, und zwar mit der gleichen Frequenz wie die Nutation. (Das hatten wir bereits allgemein nach (8.4.18) bemerkt.) Nun können wir jedoch auch die mittlere Präzessionsfrequenz berechnen: sie ist

$$\bar{\varphi} = \frac{\beta}{2b} = \frac{mgl}{\Theta_3 \omega_3}. \quad (8.4.34)$$

Die Präzessionsgeschwindigkeit ist also antiproportional zur anfänglichen Rotationsgeschwindigkeit!

Je rascher sich der Kreisel am Anfang dreht, um so stärker nimmt die Nutation ab, um so grösser wird jedoch die Nutationsfrequenz, während die Präzession um die Vertikale gleichzeitig langsamer wird. In der Praxis wird die Nutation eines hinreichend schnellen Kreisels durch die Reibung am Auflagepunkt gedämpft, und man kann sie nicht beobachten. Der Kreisel scheint dann gleichförmig um die vertikale Achse zu präzessieren. [Tatsächlich ist die Präzession nur scheinbar regulär — sie wird deshalb manchmal pseudoregulär genannt.]

8.4.4 Reine Präzession und andere Grenzfälle

Es ist interessant, genau zu bestimmen, welche Anfangsbedingungen zu einer wirklich regulären Präzession führen. Dies ist genau dann der Fall, falls $\dot{\varphi}$ konstant ist, d.h. falls der Winkel θ konstant gleich seinem Anfangswert θ_0 bleibt. Das ist natürlich genau dann der Fall, falls die beiden Wurzeln u_1 und u_2 zusammenfallen; dann ist $u = u_0$ auch eine Nullstelle von f' ,

$$\beta(1 - u_0^2) + (\alpha - \beta u_0)2u_0 - 2b(a - bu_0) = 0. \quad (8.4.35)$$

Ausserdem gilt natürlich auch $f(u_0) = 0$, und daher folgt

$$\frac{\beta}{2} = b \frac{a - bu_0}{1 - u_0^2} - u_0 \left(\frac{a - bu_0}{1 - u_0^2} \right)^2 = b \dot{\varphi} - u_0 \dot{\varphi}^2. \quad (8.4.36)$$

Für gegebene β , b und u_0 (bzw. l , ω_3 und u_0) gibt es dann (höchstens) zwei Lösungen für $\dot{\varphi}$, die man die ‘schnelle’ und die ‘langsame’ Präzession nennt. Falls $\dot{\varphi}$ endlich ist, ist $\dot{\varphi} = 0$ keine Lösung; um eine gleichförmige Präzession zu erhalten, müssen wir also stets dem Kreisel einen Stoss geben, damit er auf seine Bahn kommt. Ohne diese richtige anfängliche Präzessionsgeschwindigkeit kann man bestenfalls nur eine pseudoreguläre Präzession erhalten.

Ein weiterer interessanter Spezialfall tritt auf, falls $u = 1$ eine der Wurzeln von $f(u)$ ist. Nehmen wir zum Beispiel an, der sich drehende Kreisel sei am Anfang so gestellt, dass seine Figurenachse vertikal ist, $\theta_0 = 0$. Dann ist natürlich $M_z = M_3$ (da am Anfang die vertikale und die Figurenachse übereinstimmen), und daher $a = b$. In diesem Fall ist $u = 1$ eine Wurzel von $f(u)$. Da weiterhin am Anfang $\dot{\theta} = 0$, gilt dann ausserdem $E' = mgl$, und daher auch $\alpha = \beta$,

$$\alpha = \beta = \frac{2mgl}{\Theta_1}, \quad a = b = \frac{M_3}{\Theta_1}. \quad (8.4.37)$$

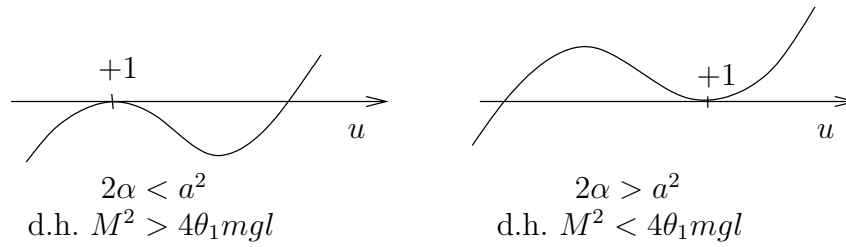
In diesem Fall hat daher $f(u)$ die Form

$$f(u) = (1 - u)^2[\alpha(1 + u) - a^2], \quad (8.4.38)$$

und $u = 1$ ist eine doppelte Wurzel. Die dritte Wurzel ist durch

$$u_3 = \frac{a^2}{\alpha} - 1 \quad (8.4.39)$$

gegeben. Falls $a^2/\alpha > 2$ — dies entspricht der Bedingung für einen ‘schnellen’ Kreisel — ist $u_3 > 1$, und die einzig mögliche Bewegung ist die für $u = 1$: in diesem Fall dreht sich der Kreisel lediglich um seine Vertikale. Wenn andererseits $a^2/\alpha < 2$ ist, dann ist $u_3 < 1$, und $f(u)$ ist nicht-negativ in dem Bereich $u_3 \leq u \leq 1$. Der Kreisel nutiert dann zwischen $\theta = 0$ und $\theta = \theta_3$.



Es gibt daher also eine kritische Winkelgeschwindigkeit ω' oberhalb der nur eine vertikale Bewegung möglich ist. Der Wert von ω' ist durch

$$\omega'^2 = 4 \frac{mgl\Theta_1}{\Theta_3^2} \quad (8.4.40)$$

gegeben. Wenn die Drehachse des rotierenden Kreisels anfangs vertikal steht und er sich mit der Winkelgeschwindigkeit ω_3 dreht, die grösser als ω' ist, so wird er sich in der Praxis für eine Weile weiter nur um die Vertikale drehen (schlafender Kreisel). Durch die Reibung wird die Frequenz der Rotation jedoch allmählich unter den kritischen Wert gebracht, und der Kreisel beginnt dann in immer grösserer Masse zu taumeln, so wie er langsamer wird.

9 Die spezielle Relativitätstheorie

Die spezielle Relativitätstheorie (SRT) ist eine physikalische Theorie über die Bewegung von Körpern und Feldern in Raum und Zeit. Sie wurde durch Einstein im Jahr 1905 veröffentlicht und erweitert das galileische Relativitätsprinzip zum speziellen Relativitätsprinzip.

Die Ursprünge der SRT liegen in der Elektrodynamik. Man kann sogar sagen, dass die Entwicklung der Maxwell-Gleichungen mit ihrer Vereinigung von Elektrizität, Magnetismus und Optik uns gleichermassen die spezielle Relativitätstheorie aufgezwungen hat. Wichtigen Pioniere dieser Entwicklung waren Lorentz und Poincaré, aber es war Einstein, der die Verallgemeinerung des zu Grunde liegenden Prinzips auf alle Phänomene der Physik erkannte und die weitreichenden Konsequenzen des zweiten Postulats verstand.

9.1 Das Gesetz der Lichtausbreitung

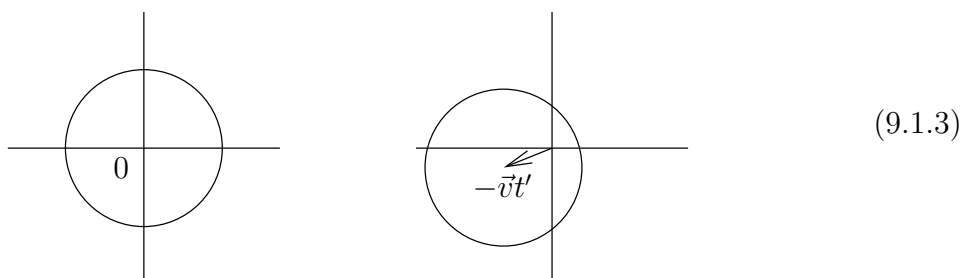
Das Gesetz der Lichtausbreitung

$$c^2(t_1 - t_2)^2 - (\mathbf{x}_1 - \mathbf{x}_2)^2 = 0 \quad (9.1.1)$$

besagt, dass die Ereignisse (t_1, \mathbf{x}_1) und (t_2, \mathbf{x}_2) durch ein Lichtsignal der festen Geschwindigkeit c verbunden werden können, und ist eine Folge der Maxwell-Gleichungen. Es ist aber nur invariant unter Galilei-Transformationen (1.2.3) mit $\mathbf{v} = 0$. Zum Beispiel beschreibt $|\mathbf{x}| = ct$ die Front einer vom Ereignis $(0, \mathbf{0})$ ausgehenden Lichtwelle. Unter der Galilei-Transformation

$$t' = t, \quad \mathbf{x}' = \mathbf{x} - \mathbf{v}t \quad (9.1.2)$$

behält das auslösende Ereignis die Koordinaten $(0, 0)$, aber die Wellenfront zur Zeit $t = t'$ wird zur Kugel $|\mathbf{x}' + \mathbf{v}t'| = ct'$ mit Zentrum $-\mathbf{v}t'$:



Diese Beobachtung gilt natürlich nicht nur für Licht, sondern für jede Welle. Normalerweise stört uns jedoch der Umstand, dass die Wellenausbreitung nicht Galilei-invariant ist, nicht, denn Wellen breiten sich typischerweise in einem Medium aus — zum Beispiel propagieren Wasserwellen im Wasser — und das einfache Gesetz der Wellenausbreitung gilt nur in dem Bezugssystem, in dem das Medium ruht. Es war daher natürlich zu postulieren, dass auch das Licht in einem Medium (das man den 'Äther' nannte) propagiert. Nach allem, was man über Licht wusste, musste der Äther überall sein; er musste vernachlässigbare Dichte haben und vernachlässigbare Wechselwirkungen mit der übrigen Materie besitzen.

Diese Ätherhypothese impliziert insbesondere, dass sich die Elektrodynamik in wesentlicher Weise von der Mechanik unterscheidet, nämlich, dass es ein bevorzugtes Bezugssystem gibt, in dem der Äther ruht, und daher das Gesetz der Lichtausbreitung (9.1.1) gilt. Diese Hypothese ist jedoch ein wenig unbefriedigend, da der Äther wenig fassbar ist und seine einzige Rolle darin zu bestehen scheint, als Medium für elektromagnetische Wellen zu fungieren.

9.2 Die Postulate von Einstein

Versuche, die Bewegung der Erde oder bewegter Bezugssysteme relativ zum Äther zu messen (insbesondere das Michelson-Morley Experiment) schlugen fehl. Lorentz erklärte dieses ‘Nullexperiment’, in dem er postulierte, dass Objekte, die sich mit der Geschwindigkeit \mathbf{v} relativ zum Äther bewegen, in der Richtung ihrer Bewegung kürzer erscheinen, und zwar gerade

$$L(v) = L_0 \sqrt{1 - \frac{v^2}{c^2}}. \quad (9.2.1)$$

Weiterhin zeigt er und Poincaré, dass die Maxwell-Gleichungen unter den Transformationen, die wir heute die *Lorentz-Transformationen* nennen, invariant sind. [Wir werden das in Kürze im Detail erklären.]

Einstein erkannte, dass die Ätherhypothese grundsätzlich unbefriedigend war, und dass das Problem darin lag, die Forminvarianz der Gleichungen der Physik unter Galilei-Transformationen zu fordern. Er schlug vor, dass alle physikalischen Gesetze den folgenden zwei Postulaten der SRT genügen müssen.

1. Relativitätsprinzip: Die Naturgesetze sind unabhängig vom Koordinatensystem. Insbesondere haben alle Naturgesetze die gleiche Form in Koordinatensystemen, die sich mit konstanter Geschwindigkeit relativ zueinander bewegen.

Bemerkung. Bezugssysteme, die sich mit konstanter Geschwindigkeit relativ zueinander bewegen, werden auch *Inertialsysteme* (IS) genannt. Die Eigenschaft ein IS zu sein, ist also eine relative Eigenschaft: Man muss zuerst ein System als IS auszeichnen und kann dann bestimmen, ob ein anderes auch ein IS ist. Erst später werden wir IS dadurch charakterisieren, dass sich freie Teilchen in ihnen geradlinig bewegen.

2. Konstanz der Lichtgeschwindigkeit: Die Lichtgeschwindigkeit ist unabhängig von der Geschwindigkeit ihrer Quelle, d.h. Licht hat dieselbe Geschwindigkeit in allen Inertialsystemen.

Wie wir später sehen werden implizieren diese Postulate auch, dass die Gesetze der klassischen Mechanik modifiziert werden müssen!

Als Einstein diese Postulate aufstellte, gab es dafür noch keine experimentellen Beweise; inzwischen sind diese Postulate jedoch in vielfältiger Weise experimentell überprüft worden, und es gibt keine Evidenz dafür, dass sie falsch sein könnten. Im Gegensatz dazu sind zum Beispiel die Postulate der Quantenmechanik nicht experimentiell überprüfbar.

9.3 Lorentzgruppe und Poincarégruppe

Das zweite Postulat von Einstein behauptet, dass die Lichtgeschwindigkeit in allen Inertialsystemen gleich ist. Insbesondere impliziert dies, dass die relevanten Transformationsgleichungen, unter denen die Naturgesetze invariant sind, nicht die Galileitransformationen sein können. Wir wollen nun die Struktur dieser Transformationen ableiten; diese sind im wesentlichen dadurch bestimmt, dass die Lichtgeschwindigkeit in allen Inertialsystemen gleich ist.

Betrachte dazu wiederum zwei Koordinatensysteme S und \hat{S} , die sich mit der konstanten Geschwindigkeit \mathbf{v} relativ zueinander bewegen. Wir bezeichnen die Koordinaten von S und \hat{S} durch (t, x, y, z) bzw. $(\hat{t}, \hat{x}, \hat{y}, \hat{z})$. Für das folgende ist es bequem, ausserdem die sogenannte 4-er Schreibweise einzuführen: wir definieren also

$$x = (ct, x, y, z) \equiv (x^0, x^1, x^2, x^3), \quad (9.3.1)$$

und schreiben die Komponenten von x als x^μ , wobei $\mu = 0, 1, 2, 3$.

Ohne Beschränkung der Allgemeinheit können wir annehmen, dass der Ursprung der beiden Koordinatensysteme für $t = \hat{t} = 0$ gerade übereinstimmt. Diese zu den Postulaten zusätzliche Einschränkung wird von den Elementen der sogenannten *Lorentz-Gruppe* erfüllt. Lässt man sie fallen, erhält man die *Poincaré-Gruppe*.

Wir nehmen an, dass eine Lichtquelle, die im System S bei $x = y = z = 0$ ruht zur Zeit $t = 0$ einen Lichtblitz aussendet. Nach Einstein's zweitem Postulat wird sich dieser Lichtblitz in beiden Inertialsystemen mit der selben Geschwindigkeit c ausbreiten. Im System S erreicht der Lichtblitz daher den Punkt (x, y, z) zur Zeit t , wobei t durch

$$(x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = c^2 t^2 - (x^2 + y^2 + z^2) = 0 \quad (9.3.2)$$

gegeben ist. (Die Punkte, die diese Gleichung erfüllen werden üblicherweise der *Lichtkegel* genannt.) Entsprechend erreicht der Lichtblitz den Punkt $(\hat{x}, \hat{y}, \hat{z})$ zur Zeit \hat{t} , wobei

$$(\hat{x}^0)^2 - (\hat{x}^1)^2 - (\hat{x}^2)^2 - (\hat{x}^3)^2 = c^2 \hat{t}^2 - (\hat{x}^2 + \hat{y}^2 + \hat{z}^2) = 0. \quad (9.3.3)$$

Die Koordinatentransformation, die die Koordinaten (t, x, y, z) auf $(\hat{t}, \hat{x}, \hat{y}, \hat{z})$ abbildet muss daher den Lichtkegel (9.3.2) auf den Lichtkegel (9.3.3) abbilden.

Um den Lichtkegel einfach zu beschreiben, führen wir nun die *Metrik* $g_{\mu\nu}$ durch

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (9.3.4)$$

ein. Die Punkte, die auf dem Lichtkegel (9.3.2) liegen, erfüllen dann die Bedingung

$$g_{\mu\nu} x^\mu x^\nu \equiv \sum_{\mu=0}^3 \sum_{\nu=0}^3 g_{\mu\nu} x^\mu x^\nu \equiv x^t g x = 0. \quad (9.3.5)$$

Solche x werden manchmal *licht-artig* genannt. Entsprechend ist der Lichtkegel im System \hat{S} durch die Gleichung

$$g_{\mu\nu} \hat{x}^\mu \hat{x}^\nu = \hat{x}^t g \hat{x} = 0 \quad (9.3.6)$$

bestimmt. Der 4-dimensionale reelle Vektorraum mit der Metrik g wird üblicherweise *Minkowski Raum* genannt. Das 4-er Skalarprodukt ist durch

$$(x, y) = g_{\mu\nu} x^\mu y^\nu \quad (9.3.7)$$

definiert.

Ein kraftfreies Teilchen bewegt sich im Inertialsystem S mit konstanter Geschwindigkeit, d.h. seine Koordinaten $x = (ct, x, y, z,)$ werden durch eine Gerade dargestellt. (Diese Gerade nennt man manchmal auch die ‘Weltlinie’ des Teilchens.) Wegen des ersten Postulates muss es sich daher auch im Inertialsystem \hat{S} auf einer Gerade bewegen.¹⁸ Dies impliziert, dass die Koordinatentransformation von S nach \hat{S} linear sein muss, d.h.

$$\hat{x}^\mu = A^\mu_\nu x^\nu + a^\mu, \quad (9.3.8)$$

wobei die Matrix $A \equiv A(\mathbf{v})$ von der Relativgeschwindigkeit zwischen den beiden Inertialsystemen abhängt. In unserem Fall haben wir den Ursprung der beiden Inertialsysteme so gewählt, dass $a^\mu = 0$. Wir können dann (9.3.8) als $\hat{x} = Ax$ schreiben. Wegen des zweiten Postulats, hat die Transformation A dann die Eigenschaft, dass $x^t g x = 0$ impliziert, dass

$$x^t (A^t g A) x = 0. \quad (9.3.9)$$

Da dies für alle licht-artigen x gelten muss, kann man sehen, dass dies nur der Fall sein kann, falls

$$A^t g A = \alpha g, \quad (9.3.10)$$

wobei $\alpha \neq 0$. Um dies zu sehen, beobachten wir zuerst, dass $A^t g A$ symmetrisch und daher von folgender Form ist

$$A^t g A = \begin{pmatrix} a & b & * & * \\ b & c & * & * \\ * & * & * & * \\ * & * & * & * \end{pmatrix}$$

Nun betrachten wir die Vektoren $\bar{x}^t = (1, 1, 0, 0)$ und $\bar{\bar{x}}^t = (1, -1, 0, 0)$. Beide erfüllen $x^t g x = 0$ und sind daher auf dem Lichtkegel. Es muss also $\bar{x}^t (A^t g A) \bar{x} = 0$ und $\bar{\bar{x}}^t (A^t g A) \bar{\bar{x}} = 0$ gelten. Damit erhalten wir die Bedingungen $a + c = 0$ und $b = 0$. Aus Symmetrie bezüglich Vertauschen der zweiten und dritten Zeile (oder alternativ dem obigen Argument für $\bar{x}^t = (1, 0, 1, 0)$ und $\bar{\bar{x}}^t = (1, 0, -1, 0)$) folgt, dass $A^t g A = \alpha g$ gilt.

¹⁸Aus der Formulierung der Postulate ist nicht direkt klar, was ein Naturgesetz genau ist. Genauer wird nicht zwischen Anfangsbedingungen und Naturgesetzen unterschieden. Zum Beispiel könnte der genaue Wert der Gravitationskonstante von den Anfangsbedingungen abhängen oder als Naturgesetz formuliert werden. In diesem Fall ist das erste Postulat so zu verstehen, dass sich das Teilchen in allen IS auf einer Geraden bewegt. Das muss aber nicht unbedingt die gleiche Gerade sein.

Wichtig ist, dass der genaue Wert von α nicht beobachtbar ist.¹⁹ Obwohl es durch die Postulate nicht direkt ausgeschlossen ist, macht es Sinn, den Fall $\alpha = 0$ auszuschliessen.²⁰ Ausserdem sollten die Transformationen auch stetig sein. Weiter haben wir $A = id$ für $v = 0$, in diesem Fall ist also $\alpha = 1 > 0$. Wenn die Transformation stetig von v abhängt, dann muss auch α stetig von v abhängen. Aus dem Zwischenwertsatz folgt dann, dass $\alpha > 0$ gilt.

Wir können dann jedes A eindeutig als $A = \lambda\Lambda$ schreiben, wobei $\alpha = \lambda^2$ mit $\lambda > 0$. Die durch Λ definierte Transformation erfüllt dann gerade

$$\Lambda^t g \Lambda = g. \quad (9.3.11)$$

In Komponenten lautet diese Gleichung

$$g_{\mu\nu} = \Lambda^\rho{}_\mu \Lambda^\sigma{}_\nu g_{\rho\sigma}. \quad (9.3.12)$$

Die linearen Transformationen, die diese Gleichung erfüllen, werden *Lorentztransformationen* genannt. Diese Transformationen generieren offensichtlich eine Gruppe, die sogenannte Lorentzgruppe L . Wenn wir die inhomogenen Translationen (die oben durch a^μ beschrieben wurden) mithinzunehmen, erhalten wir die sogenannte *Poincarégruppe*; dies ist also die Gruppe der Transformationen

$$x^\mu \mapsto \Lambda^\mu{}_\nu x^\nu + a^\mu, \quad (9.3.13)$$

wobei Λ (9.3.11) erfüllt.

9.4 Eigenschaften der Lorentztransformationen

Da jede Lorentztransformation (9.3.11) erfüllt, gilt insbesondere, dass

$$\begin{aligned} (\det \Lambda)^2 &= 1, \\ (\Lambda^0{}_0)^2 - \sum_{k=1}^3 (\Lambda^k{}_0)^2 &= 1. \end{aligned} \quad (9.4.1)$$

Die erste Gleichung impliziert, dass $\det \Lambda = \pm 1$, und es folgt aus der zweiten, dass $\pm \Lambda^0{}_0 \geq 1$. Die Lorentzgruppe L zerfällt deshalb in vier Komponenten, zwischen denen es keinen stetigen Übergang gibt:

Die Zusammenhangskomponente, die die Identität enthält, ist die Untergruppe

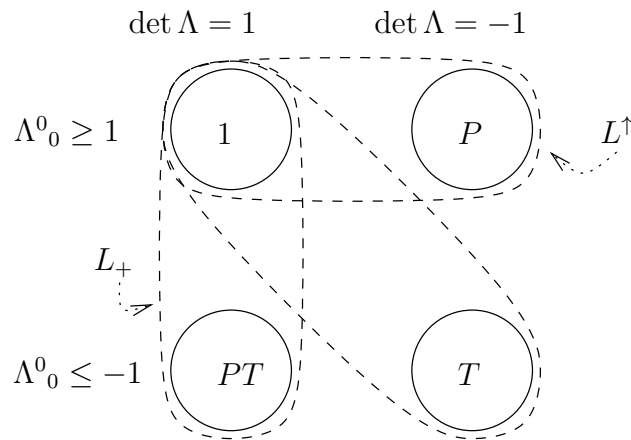
$$L_+^\uparrow = \{\Lambda \in L \mid \det \Lambda = 1, \Lambda^0{}_0 \geq 1\}, \quad (9.4.2)$$

die auch als *eigentliche orthochrone Lorentzgruppe* bezeichnet wird. Sie ist die Schnittmenge der Untergruppen

$$\begin{aligned} L^\uparrow &= \{\Lambda \in L \mid \Lambda^0{}_0 \geq 1\}, \\ L_+ &= \{\Lambda \in L \mid \det \Lambda = +1\}. \end{aligned}$$

¹⁹Zum Beispiel messen wir Längen als die Zeit, die ein Lichtsignal braucht um von A nach B zu kommen. Wenn wir alle Längen und Zeiten mit einem Faktor skalieren, ändert sich nichts.

²⁰Es ist nicht klar, was es überhaupt bedeuten soll, dass wir auf den Nullpunkt transformieren.



Bemerkung. Um zu sehen, dass L^\uparrow_+ eine Untergruppe ist, betrachten wir zwei Lorentztransformationen $\Lambda, \Lambda' \in L^\uparrow_+$ und zeigen die Gruppeneigenschaft $\Lambda \cdot \Lambda' \in L^\uparrow_+$. Die erste Eigenschaft $\det \Lambda \cdot \Lambda' = 1$ folgt direkt aus der Produktregel für die Determinante. Weiter müssen wir $(\Lambda \Lambda')^0_0 \geq 1$ zeigen. Da L eine Gruppe ist, folgt schon mal $\Lambda \cdot \Lambda' \in L$ und damit $(\Lambda \Lambda')^0_0 \leq -1$ oder $(\Lambda \Lambda')^0_0 \geq 1$. Also reicht es zu zeigen, dass $(\Lambda \Lambda')^0_0 \geq 0$ gilt. Um dies zu sehen, betrachten wir

$$(\Lambda \cdot \Lambda')^0_0 = \Lambda^0_0 (\Lambda')^0_0 + \sum_{k=1}^3 \Lambda^0_k (\Lambda')^k_0.$$

Wir definieren die Vektoren $\mathbf{a} = (\Lambda^0_1, \Lambda^0_2, \Lambda^0_3)$ und $\mathbf{b} = ((\Lambda')^1_0, (\Lambda')^2_0, (\Lambda')^3_0)$. Damit können wir schreiben

$$(\Lambda \cdot \Lambda')^0_0 = \Lambda^0_0 (\Lambda')^0_0 + \mathbf{a} \cdot \mathbf{b} \geq \Lambda^0_0 (\Lambda')^0_0 - |\mathbf{a} \cdot \mathbf{b}|.$$

Wegen (9.3.11) haben wir $|\mathbf{a}| \leq |\Lambda^0_0|$ und $|\mathbf{b}| \leq |(\Lambda')^0_0|$, womit die Behauptung folgt.

Typische Vertreter der anderen drei Zusammenhangskomponenten sind die Matrizen

$$T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 \\ 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & +1 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad (9.4.3)$$

sowie

$$PT = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (9.4.4)$$

T beschreibt Zeitumkehr, wohingegen P eine Raumspiegelung an einem Punkt ist; schliesslich ist PT die Kombination dieser beiden Lorentztransformationen. Die allgemeinsten

Elemente der entsprechenden Komponenten können durch Komposition mit der Untergruppe L_+^\uparrow erzeugt werden. Im weiteren werden wir uns deshalb auf L_+^\uparrow beschränken. Im folgenden wollen wir verschiedene Untergruppen von L_+^\uparrow beschreiben.

Definition. L_+^\uparrow enthält die Untergruppe der *Rotationen*. Die zugehörigen Matrizen sind von der Form

$$\Lambda(R) := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & R & & \\ 0 & & & \end{pmatrix}. \quad (9.4.5)$$

Wegen der Eigenschaft (9.3.11) folgt für jede Matrix dieser Form, dass $R^t R = id$. Also ist $R \in \text{SO}(3, \mathbb{R})$ und beschreibt eine Rotation des \mathbb{R}^3 .

Definition. Eine interessantere Untergruppe der Lorentzgruppe ist die Gruppe der *speziellen Lorentz-Transformationen*. Es handelt sich dabei um die Lorentztransformationen der Form

$$\Lambda(u) := \begin{pmatrix} \cosh u & -\sinh u & 0 & 0 \\ -\sinh u & \cosh u & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (9.4.6)$$

wobei $u \in \mathbb{R}$.

Es ist leicht zu sehen, dass $\det \Lambda(u) = 1$, und dass $\Lambda(u)^0_0 = \cosh u \geq 1$. Diese Matrizen erfüllen die Bedingung (9.3.11), da (wegen der Blockdiagonalform genügt es den oberen 2×2 Block zu betrachten)

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 - c^2 & ab - cd \\ ab - cd & b^2 - d^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (9.4.7)$$

wobei $a = d = \cosh u$ und $b = c = -\sinh u$. Die speziellen Lorentztransformationen bilden eine einparametrische Untergruppe mit dem Multiplikationsgesetz

$$\Lambda(u_1) \Lambda(u_2) = \Lambda(u_1 + u_2). \quad (9.4.8)$$

Bemerkung. Es gilt, dass wenn eine Lorentz-Transformation von der Form

$$\Lambda(u) := \begin{pmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

ist, dann lässt sie sich als Boost $\Lambda(u)$ für ein geeignet gewähltes u schreiben. Um dies zu sehen, betrachten wir wiederum den oberen 2×2 Block und das Produkt (9.4.7). Ohne

Beschränkung der Allgemeinheit können wir $a = \cosh u$ wählen. Aus $a^2 - c^2 = 1$ folgt, $c^2 = a^2 - 1 = \sinh^2 u$. Weiter impliziert $ab - cd = 0$, dass $\frac{a}{c} = \frac{d}{b}$ und damit $d = k \cdot \cosh u$ und $b = k \cdot \sinh u$ für eine Konstante k . Der Wert von k ergibt sich aus $b^2 - d^2 = 1$, also ist $k = 1$.

Satz. Jedes $\Lambda \in L_+^\uparrow$ lässt sich nämlich als

$$\Lambda = \Lambda(R_1)\Lambda(u)\Lambda(R_2) \quad (9.4.9)$$

scheiben.

Beweis. Es sei Λ ein beliebiges Element in L_+^\uparrow . Wir betrachten den Unterraum

$$M = \{x \mid x^0 = (\Lambda x)^0 = 0\}. \quad (9.4.10)$$

Es gibt dann zwei Fälle:

(a) $\dim M = 3$

Also ist

$$M = \left\{ \begin{pmatrix} 0 \\ u \\ v \\ w \end{pmatrix} ; u, v, w \in \mathbb{R} \right\}.$$

Da

$$\Lambda \begin{pmatrix} 0 \\ u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ u' \\ v' \\ w' \end{pmatrix} \quad (9.4.11)$$

gelten muss, folgt, dass Λ von der Form

$$\Lambda(u) := \begin{pmatrix} \cdot & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix},$$

sein muss, wobei \cdot heisst, dass wir noch nichts über den betreffenden Eintrag aussagen können. Aus (9.4.11) folgt, dass

$$\Lambda(u)^{-1} := \begin{pmatrix} \cdot & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix},$$

und mit $\Lambda^t = g\Lambda^{-1}g$ folgt, dass

$$\Lambda(u) := \begin{pmatrix} \cdot & 0 & 0 & 0 \\ 0 & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot \end{pmatrix}.$$

Wie wir oben gesehen haben, ist jede Lorentztransformation dieser Form gleich $\Lambda(R)$ für ein geeignet gewähltes R . In diesem Fall ist die Zerlegung also trivial.

(b) $\dim M = 2$. Durch eine Drehung (von rechts) können wir dann erreichen, dass M gerade mit der y, z -Ebene übereinstimmt. Nach geeigneter Drehung von links hat dann Λ die Blockform,

$$\Lambda = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}, \quad (9.4.12)$$

wobei A und B 2×2 Matrizen sind. Da $L \in L_+^\uparrow$ gibt es dann nur die beiden Möglichkeiten $\det A = \det B = \pm 1$. Im ersten Fall ist dann

$$\Lambda = \begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & B \end{pmatrix}, \quad (9.4.13)$$

und damit das Produkt einer speziellen Lorentztransformation und einer Drehung. Im zweiten Fall schreiben wir

$$\Lambda = \begin{pmatrix} AC & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} C & 0 \\ 0 & B \end{pmatrix}, \quad (9.4.14)$$

wobei C die 2×2 Matrix

$$C = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (9.4.15)$$

ist. Damit ist Λ wiederum das Produkt einer speziellen Lorentztransformation und einer Drehung, und wir haben die obige Zerlegung bewiesen. \square

Die speziellen Lorentztransformationen werden manchmal auch als *boosts* bezeichnet. Um die diesem Begriff zu Grunde liegende Interpretation zu verstehen, schreiben wir die zu (9.4.6) gehörende Transformation $\hat{x} = \Lambda(u)x$ in Komponenten

$$\begin{aligned} \hat{t} &= ct \cosh u - x^1 \sinh u, & \hat{x}^2 &= x^2 \\ \hat{x}^1 &= -ct \sinh u + x^1 \cosh u, & \hat{x}^3 &= x^3. \end{aligned}$$

Der Ursprung des Inertialsystems \hat{S} (d.h. der Punkt mit Koordinaten $\hat{x}^1 = \hat{x}^2 = \hat{x}^3 = 0$) hat im Inertialsystem S die Koordinaten

$$x^1 = ct \tanh u, \quad x^2 = 0, \quad x^3 = 0. \quad (9.4.16)$$

Dies bedeutet, dass $\Lambda(u)$ das Inertialsystem S auf ein sich mit der Relativgeschwindigkeit

$$v = c \tanh u \quad (9.4.17)$$

in der x^1 -Richtung bewegendes Inertialsystem \hat{S} abbildet. Da $-1 < \tanh u < 1$ gilt insbesondere, dass $-c < v < c$. Der andere Hauptunterschied zu den üblichen Galileitransformationen besteht ausserdem darin, dass nun auch die Zeitkomponente nicht-trivial transformiert wird.

Wir können natürlich auch die speziellen Lorentztransformationen statt durch u durch die Relativgeschwindigkeit v parametrisieren. Da

$$\frac{1}{\cosh^2 u} = 1 - \tanh^2 u \quad (9.4.18)$$

folgt aus (9.4.17), dass

$$\cosh u = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad \sinh u = \frac{v}{c} \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad (9.4.19)$$

und die obige Koordinatentransformation lautet

$$\begin{aligned} \hat{t} &= \frac{t}{\sqrt{1 - \frac{v^2}{c^2}}} - \frac{x^1 v}{c^2} \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}, & \hat{x}^2 &= x^2 \\ \hat{x}^1 &= \frac{x^1}{\sqrt{1 - \frac{v^2}{c^2}}} - vt \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}, & \hat{x}^3 &= x^3. \end{aligned}$$

Im Limes $c \rightarrow \infty$ gehen die speziellen Lorentztransformationen daher in die entsprechende Galilei-Transformation

$$\hat{t} = t, \quad \hat{x}^1 = x^1 - vt, \quad \hat{x}^2 = x^2, \quad \hat{x}^3 = x^3 \quad (9.4.20)$$

über. Im Gegensatz zu diesen Galileitransformationen (bei denen die Relativgeschwindigkeiten additiv sind) ist das bei den Lorentztransformationen nicht der Fall. [Dort ist statt dessen der Parameter u additiv!] Führt man zwei Transformationen mit Relativgeschwindigkeiten v_1 und v_2 (beide in der x^1 Richtung) hintereinander aus, dann resultiert daraus eine Transformation mit Relativgeschwindigkeit

$$v = c \tanh(u_1 + u_2) = c \frac{\tanh u_1 + \tanh u_2}{1 + \tanh u_1 \tanh u_2} = \frac{v_1 + v_2}{1 + \frac{v_1 v_2}{c^2}}. \quad (9.4.21)$$

Dies ist das relativistische Additionsgesetz für Geschwindigkeiten. Es sorgt insbesondere dafür, dass die Relativgeschwindigkeit v immer im Betrag kleiner als die Lichtgeschwindigkeit c bleibt!

9.4.1 Kontravariante und kovariante Tensoren

In der obigen Analyse haben wir die Koordinaten x^μ als einen kontravarianten Tensor aufgefasst, d.h. seine Indizes stehen oben. Die Lorentztransformation ist dann durch

$$\hat{x}^\mu = \Lambda^\mu{}_\nu x^\nu \quad (9.4.22)$$

beschrieben. Indizes werden mit Hilfe der Metrik $g_{\mu\nu}$, bzw. ihrer Inversen $g^{\mu\nu}$ gesenkt und gehoben. [Die Inverse Metrik erfüllt also

$$g^{\mu\nu} g_{\nu\rho} = \delta_\rho^\mu. \quad (9.4.23)$$

In unserem Fall ist natürlich einfach, als Matrix, $g^{\mu\nu} = g_{\mu\nu}$.] Also ist zum Beispiel der kovariante Tensor, der zu x^μ gehört, durch

$$x_\nu = g_{\nu\mu} x^\mu \quad (9.4.24)$$

definiert. Er transformiert sich dann als *kovarianter Tensor*, d.h.

$$\hat{x}_\nu = \Lambda_\nu^\mu x_\mu, \quad (9.4.25)$$

denn es gilt

$$\hat{x}_\nu = g_{\nu\rho} \hat{x}^\rho = g_{\nu\rho} \Lambda^\rho_\sigma x^\sigma = g_{\nu\rho} \Lambda^\rho_\sigma g^{\sigma\tau} x_\tau, \quad (9.4.26)$$

und

$$g_{\nu\rho} \Lambda^\rho_\sigma g^{\sigma\tau} = \Lambda_\nu^\tau \quad (9.4.27)$$

nach der Konvention bezüglich des Hebens und Senkens von Indizes.

Die Matrix Λ_ν^μ hat auch noch eine andere Interpretation: sie beschreibt die inverse Transformation zu Λ^ρ_σ . Konkret: sei $\hat{x}^\mu = \Lambda^\mu_\nu x^\nu$, dann gilt

$$x^\nu = (\Lambda^{-1})^\nu_\mu \hat{x}^\mu \equiv \Lambda_\mu^\nu \hat{x}^\mu, \quad (9.4.28)$$

d.h. die beiden Matrizen Λ^μ_ν und Λ_μ^ν sind invers zueinander,

$$\Lambda^\mu_\nu \Lambda_\rho^\nu = \delta_\rho^\mu, \quad \Lambda_\mu^\nu \Lambda^\mu_\rho = \delta_\rho^\nu. \quad (9.4.29)$$

Wegen (9.4.27) gilt

$$\Lambda_\mu^\nu = g_{\mu\rho} g^{\nu\sigma} \Lambda^\rho_\sigma. \quad (9.4.30)$$

Da Lorentztransformationen die Minkowski Metrik invariant lassen, gilt dann andererseits

$$\begin{aligned} \Lambda^\mu_\nu \Lambda_\rho^\nu &= \Lambda^\mu_\nu g_{\rho\sigma} g^{\nu\tau} \Lambda^\sigma_\tau \\ &= g_{\rho\sigma} g^{\mu\sigma} = \delta_\rho^\mu, \end{aligned} \quad (9.4.31)$$

wobei wir benützt haben, dass

$$g^{\nu\tau} \Lambda^\sigma_\tau \Lambda^\mu_\nu = g^{\mu\sigma}. \quad (9.4.32)$$

[Letztere Identität folgt aus

$$\begin{aligned} g_{\nu\mu} g^{\mu\sigma} &= g_{\nu\mu} g^{\rho\tau} \Lambda^\sigma_\tau \Lambda^\mu_\rho \\ &= g_{\alpha\mu} \Lambda^\sigma_\tau \Lambda^\mu_\rho \Lambda^\alpha_\beta g^{\rho\tau} (\Lambda^{-1})^\beta_\nu \\ &= g_{\beta\rho} g^{\rho\tau} \Lambda^\sigma_\tau (\Lambda^{-1})^\beta_\nu \\ &= \delta_\beta^\tau \Lambda^\sigma_\tau (\Lambda^{-1})^\beta_\nu = \delta_\nu^\sigma, \end{aligned} \quad (9.4.33)$$

wobei wir in der dritten Zeile (9.3.12) verwendet haben.]

Schliesslich sei noch bemerkt, dass die Ableitung nach x^μ einen kovarianten Tensor definiert: es gilt nämlich, dass

$$\frac{\partial f}{\partial \hat{x}^\mu} = \frac{\partial f}{\partial x^\nu} \frac{\partial x^\nu}{\partial \hat{x}^\mu} = \frac{\partial f}{\partial x^\nu} \Lambda_\mu^\nu . \quad (9.4.34)$$

Daher schreibt man oft auch $\partial_\mu = \frac{\partial}{\partial x^\mu}$. Dies erklärt die Indexstruktur, die wir in der Diskussion des Phasenraumes benützt haben: die Ortskoordinaten sind kontravariante Tensoren, q^i , aber die Impulse sind Ableitungen nach den kontravarianten Tensoren \dot{q}^j , und daher kovariante Tensoren, p_j .

9.5 Relativistische Mechanik

Wie wir oben erwähnt haben, war es Einstein der realisierte, dass die der Lorentzsymmetrie der Elektrodynamik zu Grunde liegende Prinzipien weit über die Elektrodynamik hinausgehen. Seine Postulate der SRT betreffen die ganze Physik und daher insbesondere auch die klassische Mechanik. In diesem Kapitel wollen wir die Auswirkung dieser Postulate auf die Mechanik beschreiben.

Die Bewegung eines Teilchens in der Raum-Zeit \mathbb{R}^4 wird durch seine *Weltlinie* dargestellt,

$$x(\lambda) = (x^0(\lambda), \mathbf{x}(\lambda)) , \quad (9.5.1)$$

wobei λ ein beliebiger Kurvenparameter ist, z.B. die Zeitkoordinate,

$$x(t) = (ct, \mathbf{x}(t)) . \quad (9.5.2)$$

Wie wir jedoch zuvor gesehen haben, hängt die Zeitkoordinate von dem benutzten Koordinatensystem ab. Eine Lorentz-invariante Grösse ist jedoch die Bogenlänge,

$$\int_{\lambda_1}^{\lambda_2} d\lambda \sqrt{\left(\frac{dx}{d\lambda}, \frac{dx}{d\lambda} \right)} = \int_{s_1}^{s_2} ds , \quad (9.5.3)$$

wobei (\cdot, \cdot) das Minkowski-Skalarprodukt ist. Die Bogenlänge s ist hier dadurch charakterisiert, dass

$$\left(\frac{dx}{ds}, \frac{dx}{ds} \right) = 1 , \quad (9.5.4)$$

d.h.

$$ds^2 = (dx, dx) = g_{\mu\nu} dx^\mu dx^\nu . \quad (9.5.5)$$

Die Bogenlänge ist dadurch bis auf die Transformationen $s \mapsto s' = \pm s + a$ eindeutig festgelegt. Statt s benutzen wir üblicherweise die sogenannte *Eigenzeit*

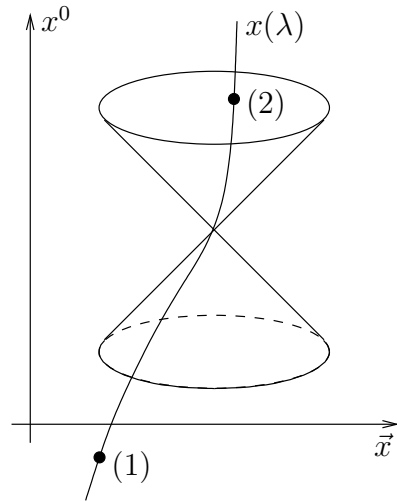
$$\tau = \frac{s}{c} . \quad (9.5.6)$$

Falls wir die Geschwindigkeit $\mathbf{v} = \frac{d\mathbf{x}}{dt}$ einführen, folgt aus (9.5.5) einfach

$$ds^2 = (c^2 - \mathbf{v}^2) dt^2, \quad (9.5.7)$$

und daher

$$d\tau = \sqrt{1 - \frac{\mathbf{v}^2}{c^2}} dt. \quad (9.5.8)$$



In dem Bezugssystem, in dem das Teilchen (momentan) ruht, gilt daher insbesondere $d\tau = dt$. Der Parameter τ ist also die Zeit in dem Ruhesystem des Teilchens; dies erklärt den Begriff Eigenzeit.

In der obigen Analyse haben wir vorausgesetzt, dass $v < c$, so dass wir mit Hilfe einer Lorentztransformation in das Ruhesystem des Teilchens transformieren können. Geometrisch bedeutet dies, dass die Weltlinie des Teilchens innerhalb des Lichtkegels durch jeden ihrer Punkte verläuft. Falls $v < c$, dann ist wegen (9.5.7) $ds^2 > 0$. Vektoren, deren Skalarprodukt bezüglich der Minkowski-Metrik positiv ist, nennt man *zeitartig*; falls das Skalarprodukt negativ ist, nennt man sie *raumartig*. [Der Vektor ds ist daher also zeitartig!]

Die Bedingung $v < c$ ist mit der Bewegungsgleichung verträglich (siehe später): ein Teilchen mit Anfangsgeschwindigkeit $v < c$ kann nie auf eine Geschwindigkeit, die grösser (oder gleich) die Lichtgeschwindigkeit ist, beschleunigt werden. Wir wählen in (9.5.8) stets das positive Vorzeichen, so dass $d\tau$ das gleiche Vorzeichen wie dt hat; streng genommen ist $d\tau$ dann kein Skalar, sondern lediglich ein Pseudoskalar, d.h. es transformiert sich unter Lorentztransformationen als

$$d\tau' = \text{sgn}(\Lambda_0^0) d\tau. \quad (9.5.9)$$

Dann definieren wir die (Pseudo)vektoren

$$u = \frac{dx}{d\tau}, \quad p = mu, \quad (9.5.10)$$

wobei $m > 0$ die Lorentz-invariante Masse des Teilchens ist. u wird die 4-er Geschwindigkeit und p der 4-er Impuls genannt. Beide Grössen transformieren sich unter den eigentlichen orthochronen Lorentztransformationen als

$$\hat{u}^\mu = \Lambda^\mu{}_\nu u^\nu, \quad \hat{p}^\mu = \Lambda^\mu{}_\nu p^\nu. \quad (9.5.11)$$

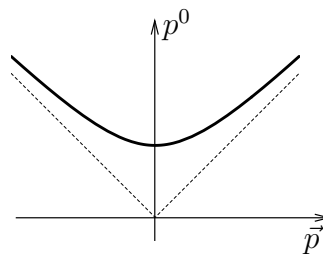
[Unter allgemeinen Lorentztransformationen tritt zusätzlicherweise der Faktor $\text{sgn}(\Lambda_0^0)$ auf — daher sind u und p lediglich Pseudovektoren.] Gleichung (9.5.4) impliziert dann,

dass

$$(u, u) = c^2, \quad (p, p) = m^2 c^2. \quad (9.5.12)$$

In Komponenten ist

$$u^\mu = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} (c, \mathbf{v}), \quad p^\mu = \frac{m}{\sqrt{1 - \frac{v^2}{c^2}}} (c, \mathbf{v}). \quad (9.5.13)$$



Insbesondere ist daher $p^0 > 0$; der 4-er Impuls liegt daher auf dem positiven Massenhypersboloid im \mathbb{R}^4 ,

$$(p^0)^2 - \mathbf{p}^2 = m^2 c^2. \quad (9.5.14)$$

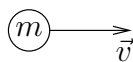
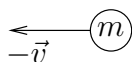
Analog zu (1.3.1) postuliert man, dass der Gesamtimpuls P^μ eines isolierten Systems erhalten ist. (Da P^μ ein 4-er Vektor ist, genügt dafür, dass der räumliche Anteil \mathbf{P} in jedem Inertialsystem erhalten ist). Im Unterschied zum nicht relativistischen Fall ist dann jedoch die Gesamtmasse nicht erhalten! Dies kann an dem Beispiel des symmetrischen Zerfalls illustriert werden:

vorher:



$$P^\mu = (Mc, \mathbf{0})$$

nacher:



$$P^\mu = \frac{1}{\sqrt{1 - v^2/c^2}} (2mc, \mathbf{0})$$

Falls P^μ erhalten ist, gilt

$$2m = M \sqrt{1 - v^2/c^2} < M, \quad (9.5.15)$$

und die Gesamtmasse ist nicht erhalten! Der Massendefekt, mal c^2 , ist

$$(M - 2m)c^2 = 2m \left(\frac{1}{\sqrt{1 - v^2/c^2}} - 1 \right) c^2 = 2 \cdot \frac{1}{2} m v^2 + \dots \quad (9.5.16)$$

Für $v \ll c$ ist das also gleich der nichtrelativistischen Energie der Zerfallsprodukte.

9.5.1 Zeitdilatation und Längenkontraktion

Eine überraschende Eigenschaft der SRT ist das Phänomen der *Zeitdilatation*. Um dies zu verstehen, machen wir das folgende Gedankenexperiment. Wir betrachten zwei Inertialsysteme S und \hat{S} , die sich mit Geschwindigkeit v gegeneinander bewegen. In dem System S liegt eine Uhr am Ursprung $\mathbf{x} = 0$. Dabei vergeht im System S genau die Zeit Δt zwischen zwei Schlägen der Uhr. In S finden diese beiden Ereignisse am selben Raumpunkt statt; sie sind daher durch

$$\Delta x = (\Delta t, \mathbf{0}) \quad (9.5.17)$$

voneinander getrennt. Nun transformieren wir diese Raum-Zeit Ereignisse in das Inertialsystem \hat{S} , das sich mit der Geschwindigkeit v in der x^1 Richtung relativ zu S bewegt.

Dann folgt aus der Lorentztransformation (siehe Kapitel 5.3)

$$\begin{aligned}\hat{t} &= \frac{t}{\sqrt{1 - \frac{v^2}{c^2}}} - \frac{x^1 v}{c^2} \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}, & \hat{x}^2 &= x^2 \\ \hat{x}^1 &= \frac{x^1}{\sqrt{1 - \frac{v^2}{c^2}}} - vt \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}, & \hat{x}^3 &= x^3\end{aligned}$$

dass sich die beiden Raum-Zeit Ereignisse im System \hat{S} gerade um

$$\Delta \hat{x} = \left(\frac{\Delta t}{\sqrt{1 - \frac{v^2}{c^2}}}, \frac{v \Delta t}{\sqrt{1 - \frac{v^2}{c^2}}} \right) \quad (9.5.18)$$

unterscheiden. Insbesondere finden die beiden Ereignisse natürlich nicht mehr am selben Raumpunkt in \hat{S} statt; da die Minkowski Distanz $c^2 t^2 - \mathbf{x}^2$ invariant ist, bedeutet das daher auch notwendigerweise, dass $\Delta \hat{t} \neq \Delta t$. Tatsächlich finden wir, dass

$$\Delta \hat{t} = \frac{\Delta t}{\sqrt{1 - \frac{v^2}{c^2}}}. \quad (9.5.19)$$

Für einen Beobachter, der die Uhr in Bewegung sieht, läuft sie mit einem um den Faktor

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (9.5.20)$$

gedehnten Zeitintervall gegenüber dem Intervall im Ruhesystem der Uhr — d.h. sie geht langsamer! Dies ist das Phänomen der Zeitdilatation, das zunächst sehr verwirrend erscheint (siehe zum Beispiel das ‘Zwillingsparadox’).

Dieser Effekt kann sehr schön am Zerfall von Myonen illustriert werden. Das Myon ist eine Art schwereres und instabiles Elektron. Seine Masse ist rund 200 mal grösser als die des Elektrons, und es kann spontan in ein Elektron und zwei Neutrinos zerfallen. Bringt man eine grosse Anzahl Myonen im Labor zur Ruhe und misst deren mittlere Lebensdauer, dann findet man das Resultat

$$\tau^{(0)}(\mu) = (2.19703 \pm 0.00004) \cdot 10^{-6} \text{ s}. \quad (9.5.21)$$

Macht man jedoch dieselbe Messung an einem Strahl von Myonen, die im Laborsystem mit der konstanten Geschwindigkeit v fliegen (zum Beispiel hageln viele atmosphärische Myonen ununterbrochen auf uns nieder!), so findet man statt $\tau^{(0)}(\mu)$ nun die mittlere Lebensdauer $\tau^{(v)}(\mu)$, wobei

$$\tau^{(v)}(\mu) = \gamma \tau^{(0)}(\mu). \quad (9.5.22)$$

(Dies kann mit einer Genauigkeit von ca. 1 Promille gemessen werden!) Schnelle Myonen leben daher (in unserem Bezugssystem) länger!

Ein weiterer, sehr verwandter Effekt, ist die sogenannte *Längenkontraktion*. Dazu stellen wir uns vor, dass wir zwei Markierungen an den Punkten

$$\mathbf{x}^{(A)} = (0, 0, 0), \quad \mathbf{x}^{(B)} = (L, 0, 0) \quad (9.5.23)$$

im Inertialsystem S vorgeben. (Diese Markierungen sind im System S statisch, d.h. sie ändern sich nicht in der Zeit.) Wir betrachten nun ein Inertialsystem \hat{S} , dessen Ursprung zur Zeit $t = \hat{t} = 0$ mit dem Ursprung in S übereinstimmt und sich mit Geschwindigkeit v in der x^1 -Richtung relativ zu S bewegt. In S erreicht der Ursprung von \hat{S} gerade zur Zeit $T = L/v$ den Punkt B .

Ein Beobachter, der am Ursprung in \hat{S} sitzt, erreicht den Punkt B jedoch zur Zeit

$$\hat{T} = \gamma \left(\frac{L}{v} - \frac{Lv}{c^2} \right) = \frac{L}{v} \gamma \left(1 - \frac{v^2}{c^2} \right) = \frac{L}{v} \sqrt{1 - \frac{v^2}{c^2}} = T/\gamma. \quad (9.5.24)$$

Da er sich mit der Geschwindigkeit v relativ zu S bewegt, deduziert er, dass der Abstand zwischen A und B gerade

$$\hat{L} = L \sqrt{1 - \frac{v^2}{c^2}} \quad (9.5.25)$$

ist. Für den sich bewegenden Beobachter erscheint daher der Abstand um den Faktor $1/\gamma$ verkürzt zu sein — das ist die sogenannte Lorentz-Kontraktion. Es ist klar, dass dieses Phänomen nicht die Abstände in der 2- oder 3-Richtung betrifft: die Lorentz Kontraktion bedeutet daher präziser, dass bewegte, räumlich ausgedehnte Objekte in der Richtung der Geschwindigkeit \mathbf{v} kontrahiert erscheinen; die Richtungen senkrecht zur Geschwindigkeit \mathbf{v} bleiben unverändert.

9.6 Lagrange Formulierung

Die Weltlinien freier Teilchen sind Geraden im \mathbb{R}^4 , und damit Geodäten im Sinne der Lorentz-invarianten Metrik $ds^2 = g_{\mu\nu} dx^\mu dx^\nu$, also charakterisiert durch das Variationsprinzip

$$\delta \int_{(1)}^{(2)} ds = 0 \quad (9.6.1)$$

bei festen Endpunkten im \mathbb{R}^4 . In jedem Inertialsystem können wir dies nach (9.5.4) schreiben als Hamilton-Prinzip

$$0 = \delta \int_{(1)}^{(2)} -mc^2 d\tau = \delta \int_{(1)}^{(2)} L_0(\mathbf{v}) dt, \quad (9.6.2)$$

wobei L_0 die Lagrange-Funktion

$$L_0(\mathbf{v}) = -mc^2 \sqrt{1 - \frac{\mathbf{v}^2}{c^2}} \quad (9.6.3)$$

ist. Der an sich willkürliche Faktor $-mc^2$ wird deshalb gewählt, weil dann für $v \ll c$

$$L_0(\mathbf{v}) = -mc^2 + \frac{1}{2}m\mathbf{v}^2 + \dots \quad (9.6.4)$$

äquivalent wird zur nicht-relativistischen Lagrange-Funktion des freien Teilchens: $\frac{1}{2}m\mathbf{v}^2$. In diesem Grenzfall kennen wir auch den Zusatzterm (5.4.19)

$$-e\left(\varphi - \frac{\mathbf{v}}{c} \cdot \mathbf{A}\right), \quad (9.6.5)$$

der die Kopplung des Teilchens an ein äusseres elektromagnetisches Feld beschreibt. Diesen übernehmen wir unverändert in die relativistische Lagrange-Funktion:

$$L(\mathbf{x}, \mathbf{v}, t) = -mc^2 \sqrt{1 - \mathbf{v}^2/c^2} - e\left(\varphi - \frac{\mathbf{v}}{c} \cdot \mathbf{A}\right). \quad (9.6.6)$$

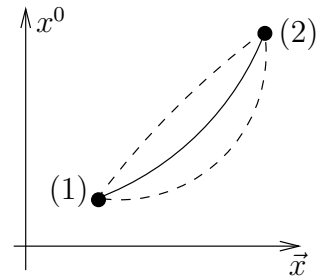
Fasst man $A = (\varphi, \mathbf{A})$ als 4-er Vektor auf (s. Vorlesung 'Elektrodynamik'), so ist

$$Ldt = -\left(mc^2 + \frac{e}{c}(u, A)\right)d\tau \quad (9.6.7)$$

lorentzinvariant, und das Hamiltonsche Variationsprinzip für die Weltlinie des Teilchens hat die lorentzinvariante Form

$$\delta \int_{(1)}^{(2)} \left(mc^2 + \frac{e}{c}(u, A)\right) d\tau = 0 \quad (9.6.8)$$

bei festen Endpunkten im \mathbb{R}^4 .



Wie schon zuvor ist L auch nicht eichinvariant: unter einer Eichtransformation $A_\nu \mapsto A_\nu - \partial_\nu \Lambda$ ändert sich L um ein totales Differential

$$L \mapsto L + \frac{e}{c} \left(\frac{\partial \Lambda}{\partial t} + \mathbf{v} \nabla \Lambda \right) = L + \frac{e}{c} \frac{d\Lambda}{dt}. \quad (9.6.9)$$

Insbesondere bleibt daher die Bewegungsgleichungen (d.h. die zugehörigen Euler-Lagrange Gleichungen) unverändert.

Die Bewegungsgleichungen eines geladenen Teilchens im elektromagnetischen Feld sind die Euler-Lagrange-Gleichungen zu (9.6.6):

$$\frac{d}{dt} \frac{m\mathbf{v}}{\sqrt{1 - \mathbf{v}^2/c^2}} = e \left(\mathbf{E} + \frac{\mathbf{v}}{c} \wedge \mathbf{B} \right), \quad (9.6.10)$$

denn an der nicht-relativistischen Rechnung von Kapitel 5.4 ist lediglich zu ersetzen

$$\frac{\partial}{\partial \mathbf{v}} \frac{m\mathbf{v}^2}{2} = m\mathbf{v} \quad \longrightarrow \quad -mc^2 \frac{\partial}{\partial \mathbf{v}} \sqrt{1 - \frac{\mathbf{v}^2}{c^2}} = \frac{m\mathbf{v}}{\sqrt{1 - \frac{\mathbf{v}^2}{c^2}}} \quad (9.6.11)$$

d.h. $m\mathbf{v} \rightarrow \mathbf{p}$. Durch (9.6.10) ist auch dp^0/dt bestimmt: Ableitung von (9.5.13) nach t liefert

$$c \frac{dp^0}{dt} = c \frac{\mathbf{p}}{p^0} \cdot \frac{d\mathbf{p}}{dt} = \mathbf{v} \cdot \frac{d\mathbf{p}}{dt} = e\mathbf{v} \cdot \mathbf{E}. \quad (9.6.12)$$

Dies ist der Energiesatz, denn rechts steht die Leistung der Lorentzkraft. Also ist

$$cp^0 = \frac{mc^2}{\sqrt{1 - \mathbf{v}^2/c^2}} \quad (9.6.13)$$

als relativistische *kinetische Energie* aufzufassen. Die Leistung (9.6.12) ist über ein endliches Zeitintervall beschränkt, falls \mathbf{E} es ist; dann ist es auch cp^0 , womit das Teilchen $v = c$ nicht erreichen kann. Für $v \ll c$ ist

$$\frac{mc^2}{\sqrt{1 - \mathbf{v}^2/c^2}} = mc^2 + \frac{1}{2}m\mathbf{v}^2 + \dots \quad (9.6.14)$$

Der zweite Term, $\frac{1}{2}m\mathbf{v}^2$ ist die nichtrelativistische kinetische Energie; den ersten Term, mc^2 , nennt man die *Ruheenergie* des Teilchens. Sie spielt zum Beispiel bei Zerfallsprozessen eine wichtige Rolle — siehe das Beispiel des symmetrischen Zerfalls.

Im Gegensatz zur nicht-relativistischen Mechanik macht die Formel für die kinetische Energie auch für Teilchen von verschwindender Masse Sinn: falls $m = 0$, dann ist die kinetische Energie einfach $E_{\text{kin}} = c|\mathbf{p}|$, und der 4-er Impuls ist einfach $p = (|\mathbf{p}|, \mathbf{p})$. Ein Teilchen ohne Masse trägt daher sowohl Energie, als auch Impuls (siehe Compton-Effekt!). Seine Geschwindigkeit ist immer gerade die Lichtgeschwindigkeit, unabhängig von seinem Impuls \mathbf{p} . Allerdings besitzt es kein Ruhesystem, da die Lorentz-Transformationen für $|\mathbf{v}| \rightarrow c$ divergieren. Ein Beispiel eines masselosen Elementarteilchens ist das Photon, das die Anregungen des elektromagnetischen Strahlungsfeldes beschreibt. Da Photonen masselos sind, ist es vielleicht nicht überraschend, dass man die Theorie des elektromagnetischen Strahlungsfeldes nicht auf der nicht-relativistischen Mechanik aufbauen kann, sondern dass sie in einem Rahmen formuliert werden muss, in dem die Lichtgeschwindigkeit als natürliche Grenzgeschwindigkeit auftritt. Natürlich reflektiert das wiederum lediglich, dass die Maxwell Gleichungen (die die Elektrodynamik beschreiben) relativistisch sind!

Der kanonische Impuls zu (9.6.6)

$$\mathbf{p} = \frac{\partial L}{\partial \mathbf{v}} = \frac{m\mathbf{v}}{\sqrt{1 - \mathbf{v}^2/c^2}} + \frac{e}{c}\mathbf{A} \quad (9.6.15)$$

unterscheidet sich vom räumlichen Anteil des kinetischen Impulses, wie auch schon im nicht-relativistischen Fall. Es folgt

$$\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)^2 + m^2c^2 = \frac{m^2\mathbf{v}^2}{1 - \mathbf{v}^2/c^2} + m^2c^2 = \frac{m^2c^2}{1 - \mathbf{v}^2/c^2} \quad (9.6.16)$$

und daraus die relativistische Hamiltonfunktion eines geladenen Teilchens in einem elektromagnetischen Feld:

$$\begin{aligned} H(\mathbf{x}, \mathbf{p}, t) &= \mathbf{p} \cdot \mathbf{v} - L = \frac{m\mathbf{v}^2}{\sqrt{1 - \mathbf{v}^2/c^2}} + mc^2 \sqrt{1 - \frac{\mathbf{v}^2}{c^2}} + e\varphi \\ &= c \left[\left(\mathbf{p} - \frac{e}{c} \mathbf{A}(\mathbf{x}, t) \right)^2 + m^2 c^2 \right]^{1/2} + e\varphi(\mathbf{x}, t). \end{aligned}$$

A Anhang: Die Legendre-Transformation

Definition. Die Teilmenge $A \subset \mathbb{R}^n$ ist **konvex**, falls

$$x, y \in A \implies \lambda x + (1 - \lambda)y \in A, \quad (0 \leq \lambda \leq 1).$$

Eine Funktion $f : D \rightarrow \mathbb{R}$ ($D \subset \mathbb{R}^n$, konvex) wird oft als konvex bezeichnet, falls

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \quad (1.0.17)$$

($x, y \in D, 0 \leq \lambda \leq 1$).

Fast äquivalent dazu, setzen wir stattdessen

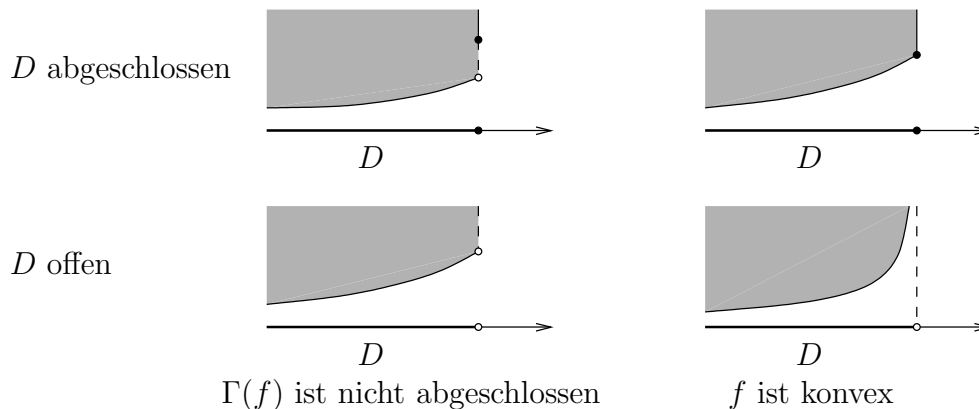
Definition. $f : D \rightarrow \mathbb{R}$ ist konvex, falls “die Menge oberhalb des Graphen”

$$\Gamma(f) = \{(x, t) \mid x \in D, t \geq f(x)\}$$

eine abgeschlossene konvexe Menge im \mathbb{R}^{n+1} ist.

Bemerkung. Diese Definition schliesst nebst (1.0.17) die Konvexität von D mit ein und regelt im Übrigen das Verhalten am Rand. Wir benutzen diese Definition, weil es nach (1.0.17) konvexe Funktionen gibt, für die die im Folgenden definierte, Legendre-Transformation nicht involutiv ist: Punkt ii) in folgendem Satz gilt dann nicht. (Man könnte verlangen, dass f differenzierbar sein muss, aber dies erweist sich als zu starke Forderungen, weil wir es in der Thermodynamik oft mit Funktionen zu tun haben, die nicht differenzierbar sind.)

Illustration für $n = 1$:



Definition. Sei $f : D \rightarrow \mathbb{R}$, ($D \subset \mathbb{R}^n$). Die *Legendretransformierte* f^* von f ist definiert für jene $p \in \mathbb{R}^n$, für welche

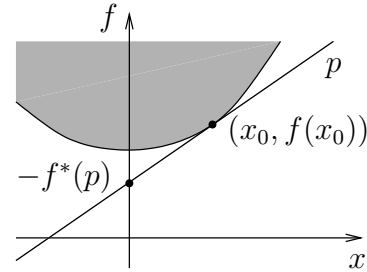
$$f^*(p) = \sup_{x \in D} ((p, x) - f(x))$$

endlich ist: $p \in D^*$, wobei $(p, x) = \sum_{i=1}^n p_i x_i$ das Skalarprodukt auf \mathbb{R}^n bezeichnet.

Bemerkung.

$$-f^*(p) = \inf_{x \in D} (f(x) - (p, x)) .$$

Geometrische Interpretation ($n = 1$): $-f^*(p)$ ist der untere Rand des Schattens von $\Gamma(f)$ auf der Ordinatenachse, geworfen durch Strahlen der Steigung p .



Satz. Sei $f : D \rightarrow \mathbb{R}$ ($D \subset \mathbb{R}^n$) eine reelle Funktion. Dann gilt

- (i) (für beliebiges f) $f^* : D^* \rightarrow \mathbb{R}$ ist konvex.
- (ii) Ist f konvex, so ist $f^{**} = f$ (mit $D^{**} = D$).
- (iii) Ist $f = f(x, y)$ konvex in (x, y) , so ist

$$f^{*(x \rightarrow p)}(p, y) = \sup_x [(p, x) - f(x, y)]$$

konkav in y .

Beweis. i) $s \geq f^*(p)$ bedeutet $s \geq (p, x) - f(x)$, ($x \in D$), sodass

$$\Gamma(f^*) = \bigcap_{x \in D} \{(p, s) \mid s \geq (p, x) - f(x)\} ,$$

was konvex ist, da es der Durchschnitt konvexer Mengen ist. Die Abgeschlossenheit von $\Gamma(f^*)$ folgt daraus, dass der Durchschnitt abgeschlossener Mengen wieder abgeschlossen ist.

ii) Nach Definition ist

$$f^*(p) \geq (p, x) - f(x) ,$$

und somit

$$f(x) \geq (p, x) - f^*(p) ,$$

für alle $x \in D$, $p \in D^*$. Also $D \subset D^{**}$ und $f(x) \geq f^{**}(x)$. Sei $(x_0, t) \notin \Gamma(f)$, d.h. (x_0, t) mit

$$\begin{cases} t < f(x_0) , & \text{falls } x_0 \in D , \\ t \text{ beliebig} , & \text{falls } x_0 \notin D . \end{cases}$$

Dann gibt es eine Ebene durch (x_0, t) , die unterhalb von $\Gamma(f)$ liegt, d.h. es existiert $p_0 \in \mathbb{R}^n$ mit

$$f(x) \geq t + (p_0, x - x_0), \quad (x \in D),$$

also mit

$$(p_0, x) - f(x) \leq (p_0, x_0) - t.$$

Es folgt $p_0 \in D^*$ und $f^*(p_0) \leq (p_0, x_0) - t$, d.h.

$$(p_0, x_0) - f^*(p_0) \geq t.$$

Nehme nun an, dass $\sup_{p \in D^*} ((p, x_0) - f^*(p)) < f(x_0)$ für $(x_0 \in D)$. Dann können wir immer ein t mit $\sup_{p \in D^*} ((p, x_0) - f^*(p)) < t < f(x_0)$ finden. Dies ist im Widerspruch zu $(p_0, x_0) - f^*(p_0) \geq t$. Also gilt $\sup_{p \in D^*} ((p, x_0) - f^*(p)) \geq f(x_0)$ für $(x_0 \in D)$. Analog lässt sich sehen, dass $\sup_{p \in D^*} ((p, x_0) - f^*(p)) = \infty$ für $(x_0 \notin D)$ gilt:

$$\sup_{p \in D^*} ((p, x_0) - f^*(p)) \begin{cases} \geq f(x_0), & (x_0 \in D) \\ = \infty, & (x_0 \notin D) \end{cases}$$

und somit $D^{**} = D$, $f^{**} = f$.

iii)

$$\sum_{i=1}^2 \alpha_i [(p, x_i) - f(x_i, y_i)] \leq (p, \sum_i \alpha_i x_i) - f(\sum_i \alpha_i x_i, \sum_i \alpha_i y_i) \leq f^*(p, \sum_i \alpha_i y_i);$$

Übergang zum Supremum über x_1, x_2 liefert $\sum_i \alpha_i f^*(p, y_i) \leq f^*(p, \sum_i \alpha_i y_i)$. □

Wir betrachten den Fall $n = 1$ im Speziellen; D ist nun ein Intervall; f sei konvex.

i) Ist p_0 die Steigung einer Stützgeraden bei x_0 , d.h.

$$f(x) \geq f(x_0) + p_0(x - x_0),$$

so ist

$$f^*(p_0) = p_0 x_0 - f(x_0). \quad (1.0.18)$$

ii) Ist f differenzierbar, so hat f in x_0 genau eine Stützgerade: die Tangente; also

$$f'(x_0) = p_0. \quad (1.0.19)$$

Ist f zudem strikt konvex (also f' strikt monoton wachsend) so ist (1.0.19) umkehrbar:

$$f^*(p) = px - f(x) \quad \text{mit } x = (f')^{-1}(p).$$

iii) Ist f konvex, so gilt (wie zuvor gezeigt) $f^{**} = f$. Falls f auch differenzierbar ist, dann kann gezeigt werden, dass auch f^* differenzierbar ist und es gilt mit (1.0.19)

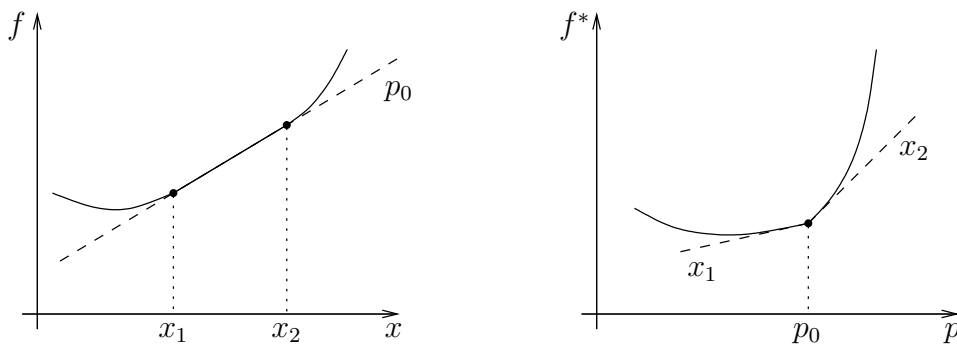
$$(f^*)'(p_0) = x_0 . \quad (1.0.20)$$

Mit (1.0.19) folgt damit, dass die Inverse von f' gleich $(f^*)'$ ist

$$(f^*)' = (f')^{-1} . \quad (1.0.21)$$

iv) Unter Legendretransformation:

Geradenstück \longleftrightarrow Ecke



Falls f linear auf $[x_1, x_2]$ ist, so gilt nach (1.0.18)

$$f^*(p_0) = p_0 x - f(x) , \quad (x \in [x_1, x_2]) ,$$

also für alle $p \in D^*$

$$f^*(p) \geq p x - f(x) = f^*(p_0) + x(p - p_0) ,$$

d.h. alle Geraden durch $(p_0, f^*(p_0))$ mit Steigungen $x \in [x_1, x_2]$ stützen f^* dort.

Hat umgekehrt f eine Ecke bei x_0 mit Steigungen $[p_1, p_2]$, so gilt wiederum nach (1.0.18)

$$f^*(p) = p x_0 - f(x_0)$$

für alle $p \in [p_1, p_2]$, d.h. $f^*(p)$ ist dort linear.

Literatur

- [G] H. Goldstein, *Klassische Mechanik*, Aula Verlag Wiesbaden, 9. Auflage (1987).
- [LL] L.D. Landau, E.M. Lifschitz, *Lehrbuch der Theoretischen Physik, Band I Mechanik*, Akademie Verlag Berlin, 11. Auflage (1976).
- [Sch] F. Scheck, *Theoretische Physik 1, Mechanik*, Springer Verlag Berlin, 7. Auflage (2003).
- [A] V.I. Arnold, *Mathematische Methoden der klassischen Mechanik*, Birkhäuser Verlag Basel (1988). [eher mathematisch]
- [St] N. Straumann, *Klassische Mechanik*, Springer Verlag Berlin (1987). [auch eher mathematisch]