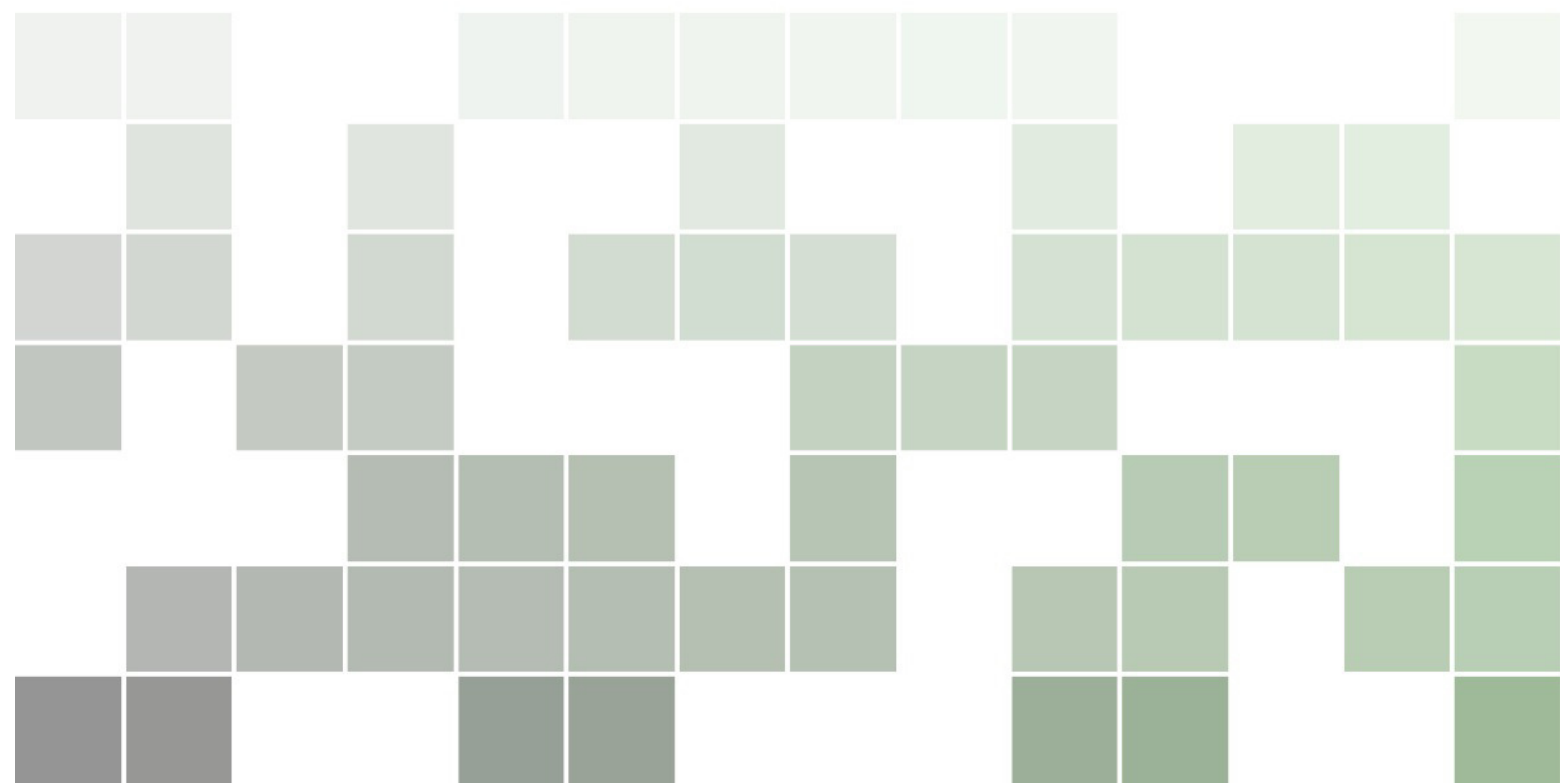


Theoretische Physik I

Nicolas Borghini

Version vom 31. Januar 2023



Nicolas Borghini

Universität Bielefeld, Fakultät für Physik

Homepage: <http://www.physik.uni-bielefeld.de/~borghini/>

Email: borghini at physik.uni-bielefeld.de

Inhaltsverzeichnis

Einleitung	1
Vorbemerkungen	2
Klassische Mechanik	7
I Newtonsche Mechanik	7
I.1 Grundbegriffe der newtonschen Mechanik 8	
I.1.1 Raumzeit der newtonschen Mechanik 8	
I.1.2 Beschreibung von mechanischen Systemen und ihrer Bewegung 9	
I.1.3 Mechanische Kräfte 10	
I.2 Newtonsche Gesetze 14	
I.2.1 Erstes newtonsches Gesetz 14	
I.2.2 Zweites newtonsches Gesetz 15	
I.2.3 Drittes newtonsches Gesetz 17	
I.2.4 „Viertes newtonsches Gesetz“ 18	
I.2.5 Energieerhaltung 18	
I.3 Inertialsysteme. Galilei-Transformationen 19	
I.3.1 Translationen 19	
I.3.2 Eigentliche Galilei-Transformationen 20	
I.3.3 Drehungen 21	
I.3.4 Allgemeine Galilei-Transformationen 22	
I.3.5 Kräfte und Galilei-Transformation 23	
I.4 Beschleunigte Bezugssysteme. Scheinkräfte 24	
I.4.1 Linear beschleunigte Bezugssysteme 24	
I.4.2 Rotierende Bezugssysteme 26	
Appendix zum Kapitel I32
I.A Kinematik eines Massenpunkts 32	
I.A.1 Grundbegriffe 32	
I.A.2 Koordinatensysteme 32	
I.A.3 Einfache Bewegungen 36	
II Newtonsche Mechanik: Anwendungen38
II.1 Mehrteilchensysteme 38	
II.1.1 Grundlagen 39	
II.1.2 Bewegung des Schwerpunkts 41	
II.1.3 Drehimpuls 41	
II.1.4 Energie 42	
II.1.5 Schwerpunktsystem 46	
II.2 Zwei-Körper-Systeme 47	
II.2.1 Separation der Bewegungsgleichungen 47	
II.2.2 Gekoppelte Punktmassen 49	
II.2.3 Kepler-Problem 49	
II.2.4 Streuung 56	

Appendix zum Kapitel II61
II.A Alternative Herleitung der Keplerschen Bahnkurven 61	
II.A.1 Vorbereitung: Binetsche Gleichung 61	
II.A.2 Kepler-Problem 62	
III Lagrange-Formalismus: Grundlagen64
III.1 Ein Resultat aus der Variationsrechnung 65	
III.1.1 Funktional 65	
III.1.2 Extremierung eines Funktionals 66	
III.2 Hamilton-Prinzip 68	
III.2.1 Definitionen 68	
III.2.2 Hamilton-Prinzip. Euler–Lagrange-Gleichungen 70	
III.2.3 Erste Beispiele 72	
III.2.4 Systeme mit Zwangsbedingungen 75	
III.3 Symmetrien und Erhaltungsgrößen 80	
III.3.1 Invarianz unter Raumzeit-Transformationen 80	
III.3.2 Noether-Theorem 84	
IV Lagrange-Formalismus: Anwendungen87
IV.1 Kleine Schwingungen 87	
IV.1.1 Eindimensionales Problem 87	
IV.1.2 Multidimensionales Problem 90	
IV.2 Starre Körper 94	
IV.2.1 Beschreibung des starren Körpers 94	
IV.2.2 Bewegungsgleichungen 97	
V Hamilton-Formalismus	106
V.1 Hamiltonsche Bewegungsgleichungen 106	
V.1.1 Kanonisch konjugierter Impuls 107	
V.1.2 Hamilton-Funktion 107	
V.1.3 Kanonische Bewegungsgleichungen 108	
V.1.4 Beispiele 110	
V.2 Phasenraum 111	
V.2.1 Phasenraumtrajektorien 111	
V.2.2 Satz von Liouville 115	
V.3 Poisson-Mechanik 118	
V.3.1 Phasenraum-Funktionen 118	
V.3.2 Poisson-Klammer 118	
V.3.3 Poisson-Klammer und Zeitentwicklung 120	
V.3.4 Kanonische Transformationen 122	
V.3.5 Poisson-Klammer und Symmetrien 125	
Klassische Elektrodynamik	131
VI Einleitung	131
VI.1 Grundbegriffe und Definitionen 132	
VI.2 Einheiten 134	
VII Elektrostatik	137
VII.1 Elektrostatistisches Potential 138	
VII.1.1 Skalarpotential 138	
VII.1.2 Poisson-Gleichung 138	
VII.1.3 Elektrisches Feld und Potential von Ladungen 139	

VII.1.4	Elektrostatische potentielle Energie	141
VII.2	Bestimmung des Skalarpotentials aus der Poisson-Gleichung	144
VII.2.1	Greensche Funktionen	144
VII.2.2	Lösung der Poisson-Gleichung auf \mathbb{R}^3	144
VII.2.3	Lösung der Poisson-Gleichung auf einem endlichen Gebiet von \mathbb{R}^3	146
VII.3	Multipolentwicklung	150
VII.3.1	Kartesische Multipolmomente	150
VII.3.2	Beispiele von elektrischen Multipolmomenten	152
VII.3.3	Wechselwirkung zwischen zwei Ladungsverteilungen	154
VII.3.4	Multipolentwicklung in Kugelkoordinaten	157
VIII	Magnetostatik	159
VIII.1	Grundbegriffe und -ergebnisse der Magnetostatik	159
VIII.1.1	Vektorpotential	160
VIII.1.2	Poisson-Gleichungen der Magnetostatik	161
VIII.1.3	Integrale Formulierung der Grundgleichungen der Magnetostatik	163
VIII.1.4	Magnetisches Feld induziert durch einfache Ladungsströme	165
VIII.1.5	Kraft zwischen zwei Stromkreisen	168
VIII.2	Multipolentwicklung	169
VIII.2.1	Multipolmomente einer Ladungsstromverteilung	169
VIII.2.2	Magnetisches Dipolmoment einer Leiterschleife	172
VIII.2.3	Magnetischer Dipol in einem äußeren magnetischen Feld	172
IX	Zeitabhängige elektromagnetische Felder	175
IX.1	Grundgesetze	176
IX.1.1	Maxwell-Gleichungen	176
IX.1.2	Bewegungsgleichungen für die elektrischen und magnetischen Felder	178
IX.2	Elektrodynamische Potentiale	179
IX.2.1	Definition	180
IX.2.2	Eichinvarianz	180
IX.2.3	Bewegungsgleichungen für die elektrodynamischen Potentiale	182
IX.3	Energie und Impuls des elektromagnetischen Feldes	183
IX.3.1	Energiedichte und -stromdichte des elektromagnetischen Feldes	183
IX.3.2	Impulsdichte und -stromdichte des elektromagnetischen Feldes	184
IX.4	Elektromagnetische Wellen im Vakuum	187
IX.4.1	Klassische Wellengleichung	187
IX.4.2	Elektromagnetische Wellen	190
IX.5	Klassische Theorie der Strahlung	193
IX.5.1	Greensche Funktion der klassischen Wellengleichung	193
IX.5.2	Retardierte Potentiale	195
IX.5.3	Multipolentwicklung	197
IX.5.4	Potentiale und Felder einer bewegten Punktladung	198
Grundlagen der Speziellen Relativitätstheorie		209
X	Mathematischer Apparat der Speziellen Relativitätstheorie	209
X.1	Einstein'sche Postulate	209
X.1.1	Motivation	209
X.1.2	Einstein'sche Postulate	209

X.2	Lorentz-Transformationen	211
X.2.1	Linienelement	211
X.2.2	Lorentz-Transformationen	212
X.2.3	Folgerungen	215
X.2.4	Minkowski-Raum	215
X.3	Vierervektoren und Vierertensoren	217
X.3.1	Lorentz-Skalare	217
X.3.2	Vierervektoren	217
X.3.3	Vierertensoren	221
X.3.4	Kovariante Formulierung eines physikalischen Gesetzes	223
XI	Relativistische Mechanik	225
XI.1	Bewegung eines freien relativistischen Teilchens	225
XI.1.1	Lagrange-Funktion und Wirkung eines freien Teilchens	225
XI.1.2	Impuls und Energie eines freien Teilchens	226
XI.2	Kovariante Formulierung des Grundgesetzes der Mechanik	228
XII	Relativistisch kovariante Formulierung der Elektrodynamik	229
XII.1	Lorentz-kovariante elektromagnetische Größen	229
XII.1.1	Elektromagnetischer Feldstärketensor	229
XII.1.2	Viererpotential	232
XII.1.3	Elektrischer Viererstrom	233
XII.1.4	Energieimpulstensor	233
XII.2	Relativistisch kovariante Formulierung der Grundgesetze	234
XII.2.1	Maxwell-Gleichungen	234
XII.2.2	Kontinuitätsgleichung	234
XII.2.3	Lorentz-Kraft und -Kraftdichte	235
XII.2.4	Energie- und Impulsbilanzgleichungen	235
XII.3	Weitere Resultate in relativistisch kovarianter Form	236
XII.3.1	Bewegungsgleichung für das Viererpotential	236
XII.3.2	Klassische Wellengleichung und ebene Wellen	236
XII.3.3	Retardiertes Viererpotential	237
XIII	Lagrange-Formulierung der Elektrodynamik	238
XIII.1	Ladungen und Ströme in einem elektromagnetischen Feld	238
XIII.1.1	Wiederholung: Lagrange-Funktion einer freien Punktladung	238
XIII.1.2	Punktladung in einem äußeren elektromagnetischen Feld	239
XIII.1.3	Ladungs- und Stromverteilungen in einem elektromagnetischen Feld	241
XIII.2	Elektromagnetisches Feld mit Quellen	242
XIII.2.1	Einführung in die klassische Feldtheorie	242
XIII.2.2	Elektromagnetisches Feld in Anwesenheit fester Quellen	243
XIII.2.3	Energieimpulstensor	244
Anhang		251
A	Tensoren auf einem Vektorraum	251
A.1	Vektoren, Linearformen und Tensoren	251
A.1.1	Vektoren	251
A.1.2	Linearformen	252
A.1.3	Tensoren	253
A.1.4	Metrischer Tensor	255
A.1.5	Warum Tensoren?	256
A.2	Basistransformation	256

B Drehungen	258
B.1 Isometrien im euklidischen Raum	258
B.2 Infinitesimale Drehungen	260
C Die δ-Distribution	262
C.1 Definition und erste Ergebnisse	262
C.1.1 Definition	262
C.1.2 Erste Eigenschaften	263
C.1.3 Darstellungen der δ -Distribution	263
C.1.4 Heaviside-Funktion	265
C.2 Rechenregeln	265
C.2.1 Skalierung	266
C.2.2 Ableitung	266
C.2.3 Substitution der Integrationsvariablen	266
C.3 Mehrdimensionale δ -Distributionen	267
D Legendre-Transformation	269
D.1 Legendre-Transformation einer Funktion einer Variablen	269
D.2 Legendre-Transformation einer Funktion mehrerer Variablen	270
E Vektoranalysis	271
F Kugelflächenfunktionen	272
Literaturverzeichnis	275

Einleitung

Allgemeine Einleitung.

Notationen, Konventionen, usw.

Allgemeine Literaturhinweise

(in alphabetischer Ordnung)

- Arnold, *Mathematical methods of classical mechanics* [1] (mehr mathematisch);
- Fließbach, *Lehrbuch zur theoretischen Physik I. Mechanik* [2] & *II. Elektrodynamik* [3];
- Goldstein, *Klassische Mechanik* [4] = *Classical Mechanics* [5]
- Greiner, *Klassische Mechanik I & II* [6, 7]; *Klassische Elektrodynamik* [8];
- Griffiths, *Elektrodynamik* [9] = *Introduction to Electrodynamics* [10];
- Jackson, *Klassische Elektrodynamik* [11] = *Classical Electrodynamics* [12]
- Landau & Lifschitz, *Lehrbuch der theoretischen Physik. Band I: Mechanik* [13] & *Band II: Klassische Feldtheorie* [14]
- Nolting, *Grundkurs Theoretische Physik. Band 1: Klassische Mechanik* [15], *Band 2: Analytische Mechanik* [16] & *Band 3: Elektrodynamik* [17].
- Scheck, *Theoretische Physik 1: Mechanik* [18] & *3: Klassische Feldtheorie* [19].

Vorbemerkungen

Physikalische Dimension

Physikalische Größen sind „dimensionsbehaftet“. Die *physikalische Dimension* bestimmt insbesondere die möglichen Einheiten, die bei einer Messung der Größe angewandt werden dürfen. Ist \mathcal{G} eine gegebene physikalische Größe, so wird die zugehörige Dimension üblicherweise mit $[\mathcal{G}]$ gekennzeichnet.⁽¹⁾

Diese physikalischen Dimensionen lassen sich alle als Produkte von Potenzen von nur 7 *Basisgrößen*, und zwar Länge (Symbol: L), Masse (M), Zeit (T), elektrische Stromstärke (I), (thermodynamische) Temperatur (Θ), Stoffmenge (N) und Lichtstärke (J). In der Mechanik und der Elektrodynamik sind nur die vier Basisgrößen L, M, T und I relevant, was ab jetzt angenommen wird. Für eine beliebige physikalische Größe \mathcal{G} kann man daher

$$[\mathcal{G}] = L^\alpha M^\beta T^\gamma I^\delta \quad (.1)$$

schreiben, wobei $\alpha, \beta, \gamma, \delta$ rationale Zahlen sind. Zum Beispiel gilt für Energie ($[E]$), Geschwindigkeit ($[v]$), Kraft ($[F]$), elektrische Ladung ($[Q]$)

$$[E] = ML^2T^{-2}, \quad [v] = LT^{-1}, \quad [F] = MLT^{-2}, \quad [Q] = TI.$$

Dazu wird „dimensionslosen“ Größen wie insbesondere reinen Zahlen die Dimension 1 zugeordnet.

Physikalische Größen $\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_n$ werden *dimensional unabhängig* genannt, falls das Produkt $[\mathcal{G}_1]^{\lambda_1} [\mathcal{G}_2]^{\lambda_2} \dots [\mathcal{G}_n]^{\lambda_n}$ nur dann dimensionslos ist, wenn jeder Exponent λ_i Null ist:

$$[\mathcal{G}_1]^{\lambda_1} [\mathcal{G}_2]^{\lambda_2} \dots [\mathcal{G}_n]^{\lambda_n} = 1 \quad \Leftrightarrow \quad \lambda_1 = \lambda_2 = \dots = \lambda_n = 0. \quad (.2)$$

Definitionsgemäß sind die Basisgrößen dimensional unabhängig voneinander.

Natürlich darf man statt (L,M,T,I) vier andere Basisgrößen verwenden, so lange sie dimensional unabhängig sind, wie z.B. (L,M,T,Q).⁽²⁾ Dabei gibt es eine klare Analogie zur linearen Algebra und zur möglichen Wahl von unterschiedlichen Basen auf einem gegebenen Vektorraum.

Wichtig ist, dass man nur Größen derselben Dimension addieren oder gleichsetzen kann. Somit kann eine Gleichung wie „Masse + Länge = Geschwindigkeit“ nie richtig sein.⁽³⁾ Dagegen dürfen Größen unterschiedlicher Dimensionen miteinander multipliziert werden, so dass z.B. die Gleichung „Beschleunigung mal Zeit = Geschwindigkeit“, entsprechend $LT^{-2} \cdot T = LT^{-1}$, sinnvoll ist.⁽³⁾

Entsprechend den unterschiedlichen Dimensionen von Ort, Geschwindigkeiten, Beschleunigung, Impuls, usw. sind die zugehörigen Vektoren ($\vec{r}, \vec{v}, \vec{a}, \vec{p}, \dots$) Elemente unterschiedlicher Vektorräume, auch wenn alle diese Vektoren reell und dreidimensional sind. Dieser Unterschied ist zwar nicht erkennbar, wenn man auf dem gleichen Bild den Ortsvektor (im Ortsraum) eines Körpers, seine Geschwindigkeit (im Geschwindigkeitsraum) und die darauf wirkenden Kräfte (im „Kraftraum“) darstellt, man darf aber die jeweiligen Pfeile nicht miteinander addieren.

⁽¹⁾Streng genommen sollte man noch zwischen der physikalischen Größe und deren mathematischen Darstellung \mathcal{G} unterscheiden, was hier der Kürze halber ignoriert wird.

⁽²⁾Somit wird in der Teilchen- und Hochenergiephysik statt (L,M,T) implizit die Basis (E,v,J) — wobei J hier den Drehimpuls bezeichnet — verwendet.

⁽³⁾Dementsprechend sollte die Leserin immer prüfen, dass ihre Gleichung „die richtige Dimension“ hat.

Notationen

Dreidimensionale Vektoren werden mit einem Pfeil geschrieben, wie z.B. \vec{p} oder \vec{x} .

Vierervektoren werden in einer Sans Serif-Schriftart geschrieben, wie z.B. \mathfrak{p} (Viererimpuls), \mathfrak{A} (Viererpotential), usw.

Die gleiche Schriftart wird auch für Lorentz-Tensoren höherer Stufe verwendet, z.B. für die Feldstärketensoren \mathfrak{F} , \mathfrak{G} , oder für den metrischen Tensor η . Dagegen werden die Komponenten dieser Vierervektoren und Lorentz-Tensoren in Kursivschrift geschrieben: p^μ , A^μ , $F^{\mu\nu}$, $\eta_{\mu\nu}$, usw.

Indizes

Lateinische Indizes i, j, k, l , usw. laufen über die drei möglichen räumlichen Koordinaten, d.h. über 1, 2, 3 oder x, y, z .

Ziemlich inkonsequent wird auf ihre Stelle (tief- oder hochgestellt) nicht aufgepasst, wenn sie sich auf rein dreidimensionale Größen beziehen, d.h. nicht auf die räumlichen Komponenten von Vierervektoren oder allgemeineren Lorentz-Tensoren. Somit gilt $v_i = v^i$ für die i -Komponente der Geschwindigkeit \vec{v} . Dagegen ist im Fall des Raumanteils eines Vierervektors oder eines Lorentz-Tensors die Stelle des Index wichtig.

Griechische Indizes μ, ν, ρ, σ , usw. laufen über die vier Raumzeit-Koordinaten 0, 1, 2, 3. Auf die Stelle dieser Lorentz-Indizes soll aufgepasst werden, wie in Abschn. ?? weiter betont wird.

Im ganzen Skript wird die Einsteinsche Summenkonvention verwendet, d.h. auf doppelt auftretende Indizes wird summiert. Somit gelten im Fall dreidimensionaler Indizes

$$a_i b_i = a^i b^i \equiv \sum_{i=1}^3 a_i b_i, \quad \mathcal{M}_{ii} \equiv \sum_{i=1}^3 \mathcal{M}_{ii},$$

entsprechend dem Skalarprodukt $\vec{a} \cdot \vec{b}$ bzw. der Spur $\text{Tr } \mathcal{M}$. Im Fall von Lorentz-Indizes sollte der eine kontravariant, der andere kovariant sein [s. Abschn. ??]:

$$a^\mu b_\mu = a_\mu b^\mu \equiv \sum_{\mu=0}^3 a^\mu b_\mu = \sum_{\mu=0}^3 a_\mu b^\mu, \quad T^\mu{}_\mu \equiv \sum_{\mu=0}^3 T^\mu{}_\mu,$$

entsprechend dem Viererprodukt $\mathfrak{a} \cdot \mathfrak{b}$ bzw. der Spur $\text{Tr } \mathfrak{T}$.

Erster Teil
Klassische Mechanik

KAPITEL I

Newton'sche Mechanik

-
- I.1 Grundbegriffe der newtonschen Mechanik 8
 - I.1.1 Raumzeit der newtonschen Mechanik 8
 - I.1.2 Beschreibung von mechanischen Systemen und ihrer Bewegung 9
 - I.1.3 Mechanische Kräfte 10
 - I.2 Newton'sche Gesetze 14
 - I.2.1 Erstes newtonsches Gesetz 14
 - I.2.2 Zweites newtonsches Gesetz 15
 - I.2.3 Drittes newtonsches Gesetz 17
 - I.2.4 „Viertes newtonsches Gesetz“ 18
 - I.2.5 Energieerhaltung 18
 - I.3 Inertialsysteme. Galilei-Transformationen 19
 - I.3.1 Translationen 19
 - I.3.2 Eigentliche Galilei-Transformationen 20
 - I.3.3 Drehungen 21
 - I.3.4 Allgemeine Galilei-Transformationen 22
 - I.3.5 Kräfte und Galilei-Transformation 23
 - I.4 Beschleunigte Bezugssysteme. Scheinkräfte 24
 - I.4.1 Linear beschleunigte Bezugssysteme 24
 - I.4.2 Rotierende Bezugssysteme 26
-

Die am meisten intuitive, und historisch die erste, Formulierung der Mechanik ist diejenige von Isaac Newton.^(a) Die nach ihm genannte Beschreibung beruht einerseits auf im Abschn. I.1 eingeführten Grundbegriffen, die als vom Anfang an gegeben betrachtet werden. Somit bilden Raum und Zeit den Rahmen, in welchen physikalische Systeme sich befinden und entwickeln, ohne dadurch beeinflusst zu werden. Dazu werden Änderungen des Bewegungszustands eines Systems, charakterisiert durch kinematische Größen, durch mechanische Kräfte verursacht.

Andererseits wird der Einfluss der letzteren auf physikalische Systeme durch Beziehungen zwischen diesen Kräften und kinematischen Größen bestimmt, und zwar hier durch die auf Newton zurückgehenden Gesetze, die in Abschn. I.2 dargelegt werden. Insbesondere führt laut dem zweiten newtonschen Gesetz eine Kraft zur instantanen Beschleunigung des Systems, auf welchem sie ausgeübt wird.

In der (modernen) Formulierung der newtonschen Gesetze spielen bestimmte Bezugssysteme zur Beschreibung der Bewegung eine besondere Rolle, und zwar die Inertialsysteme. Nach Angabe eines solchen Systems ist jedes weitere Bezugssystem, das sich relativ zum ersten in gleichförmiger geradliniger Bewegung befindet, ebenfalls ein Inertialsystem. Genauer nehmen die newtonschen Gesetze die gleiche Form in allen Koordinatensystemen an, die sich über eine Galilei-Transformation aus kartesischen Koordinaten in einem Inertialsystem erhalten lassen (Abschn. I.3).

In einem relativ zu Inertialsystemen beschleunigten Bezugssystem ist die Situation unterschiedlich. Das zweite newtonsche Gesetz lässt sich zwar noch in der gleichen Form ausdrücken, jedoch

^(a)I. NEWTON, 1643–1727

auf Kosten der Einführung von Scheinkräften auf den bewegten Körper, die in der Beschreibung der Bewegung in einem Inertialsystem nicht auftreten (Abschn. I.4).

I.1 Grundbegriffe der newtonschen Mechanik

Die Mechanik beschäftigt sich mit der Bewegung von physikalischen Systemen, d.h. mit der zeitlichen Änderung ihrer Position (und, bei ausgedehnten Systemen, ihrer Ausrichtung) im Raum. Zur genaueren Beschreibung der Bewegung müssen deshalb Modelle für die physikalischen Raum und Zeit spezifiziert werden (§ I.1.1). In § I.1.2 werden einige zusätzliche Basisdefinitionen und -Begriffe eingeführt, welche für die Formulierung der Gesetze der newtonschen Dynamik vorausgesetzt sind oder in der Beschreibung von Systemen bzw. von deren Bewegung benutzt werden. Schließlich werden mechanische Kräfte und ihre mathematische Modellierung in § I.1.3 eingeführt.

Die in diesem Abschnitt diskutierten Begriffe sind nicht auf den Rahmen der newtonschen Mechanik beschränkt, sondern bleiben ohne Änderung noch relevant in den anderen Formulierungen der klassischen Mechanik, die in Kap. III–V behandelt werden. Es sei hier schon erwähnt, dass manche Grundannahmen jenseits des nicht-relativistischen Rahmens nicht mehr erfüllt sind — wie z.B. die Absolutheit von Raum und Zeit in der relativistischen Mechanik —, während einige Begriffe an Bedeutung verlieren — beispielsweise wird die Bahnkurve in der Quantenmechanik nur als klassisches Analogon gesehen.

I.1.1 Raumzeit der newtonschen Mechanik

Nach Newton sind der Raum, in welchem physikalische Systeme sich befinden, und die Zeit, deren Vergehen die Entwicklung der Systeme erlaubt, *absolute* Größen. Das heißt, sie stellen einen allgemeinen Rahmen dar, in welchem physikalische Prozesse stattfinden, ohne durch diese Prozesse bzw. die daran beteiligten Systeme beeinflusst zu werden.

I.1.1 a Der Raum der newtonschen Mechanik

Der räumliche Rahmen der newtonschen Mechanik — und allgemeiner der nicht-relativistischen Physik — ist ein dreidimensionaler euklidischer Punktraum \mathcal{E}_3 , hiernach oft *Ortsraum* genannt, dessen Punkte alle äquivalent sind. Dieser Raum ist statisch, er ändert sich also nicht mit der Zeit.

Dem Ort eines (als punktförmig modellierten) Systems zu einer gegebenen Zeit wird ein Punkt $P \in \mathcal{E}_3$ zugeordnet. Zur Kennzeichnung dieses Orts werden einerseits *Bezugssysteme* eingeführt, entsprechend Beobachtern, die sich möglicherweise bewegen.

Andererseits können in jedem Bezugssystem *Koordinatensysteme* eingeführt werden, mit insbesondere einem Ursprungspunkt, der in diesem Skript oft mit O bezeichnet wird. Eine wichtige Eigenschaft von euklidischen Räumen ist die Existenz von überall im Raum geltenden *kartesischen Koordinatensystemen*, bestehend aus dem Nullpunkt O und drei zueinander orthogonalen Achsen mit festen Richtungen.

Nach der Angabe des Nullpunkts eines Bezugssystems kann man den (Orts)Vektor zwischen diesem Ursprung und jedem beliebigen Punkt $P \in \mathcal{E}_3$ betrachten. Die Menge aller solcher Vektoren — die im Folgenden mit Pfeilen gekennzeichnet werden, z.B. \vec{r} — bildet einen dreidimensionalen Vektorraum, auf dem eine euklidische Struktur definiert werden kann. Somit entspricht der Betrag $|\vec{r}| \equiv r$ des Vektors \vec{r} dem Abstand zwischen seinen Endpunkten im Punktraum. Hiernach wird oft der Einheitsvektor \vec{e}_r in Richtung von \vec{r} benutzt, d.h. der auf 1 normierte Vektor, für welchen $\vec{r} \equiv r \vec{e}_r$ gilt (falls $\vec{r} \neq \vec{0}$).

Wiederum ist der euklidische Vektorraum nach Festlegung eines kartesischen Koordinatensystems isomorph zum *Koordinatenraum* \mathbb{R}^3 aller möglichen 3-Tupel (x^1, x^2, x^3) von Koordinaten.

Diese eindeutige Beziehung wird günstig als⁽⁴⁾

$$\vec{r} = \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} \quad (\text{I.1a})$$

geschrieben, was eigentlich

$$\vec{r} = \sum_{i=1}^3 x^i \vec{e}_i \equiv x^i \vec{e}_i \quad (\text{I.1b})$$

mit den Basisvektoren $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ des Koordinatensystems bedeutet, wobei in der zweiten Gleichheit die einsteinsche Summenkonvention benutzt wird.

Für kartesische Koordinaten wird auch die Notation (x, y, z) bzw. $(\vec{e}_x, \vec{e}_y, \vec{e}_z)$ statt (x^1, x^2, x^3) bzw. $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ verwendet.

Hiernach wird der Vektorraum der Ortsvektoren entweder mit \mathcal{E}_3 — um seine euklidische Struktur zu betonen — oder mit \mathbb{R}^3 — wegen der Isomorphie mit dem Koordinatenraum — bezeichnet.

I.1.1 b Die Zeit in der newtonschen Mechanik

In der nicht-relativistischen Mechanik ist die Zeit der ein kontinuierlicher eindimensionaler Parameter — so dass Zeitpunkte als Elemente von \mathbb{R} modelliert werden —, der *universell* ist. Das heißt, die Zeit vergeht gleich schnell für alle Beobachter, unabhängig von ihrer Position und ihrer Bewegung. Insbesondere können zumindest prinzipiell alle Beobachter ihre jeweiligen Uhren zu einem bestimmten Zeitpunkt t_0 synchronisieren; dann bleiben diese Uhren auch in der Zukunft synchronisiert, d.h. wenn zwei Beobachter sich wieder treffen, zeigen ihre jeweiligen Uhren die gleiche Zeit $t_1 > t_0$, egal was ihre Bewegungen in der Zwischenzeit waren.

Bemerkung: Diese Synchronisierungseigenschaft gilt nicht mehr im relativistischen Kontext.

I.1.2 Beschreibung von mechanischen Systemen und ihrer Bewegung

I.1.2 a Idealisierte Systeme

Die Mechanik befasst sich nicht nur mit realistischen physikalischen Systemen, sondern auch mit Idealisierungen, deren Bewegung viel einfacher zu beschreiben und bestimmen ist.

Das einfachste solche Modell ist der *Massenpunkt* oder *Punktmasse*: dabei handelt es sich um einen punktförmigen Gegenstand, ohne innere Struktur, der nur eine Eigenschaft besitzt, und zwar seine Masse m .

Die Bewegung eines physikalischen Körpers lässt sich mit diesem Modell gut beschreiben, wenn die Ausdehnung des Körpers für das betrachtete Problem irrelevant ist.

Ist der Massenpunkt neben seiner Masse m noch mit einer elektrischen Ladung q versehen, so dass es in einem elektromagnetischen Feld einer entsprechenden Kraft unterliegt, dann spricht man von einer *Punktladung*.

Bemerkung: Wie in § I.2.2 weiter diskutiert wird, handelt es sich bei der hier mit m bezeichneten Größe um die *träge Masse* des Massenpunkts bzw. der Punktladung.

Ein weiteres idealisiertes Modell, und zwar jetzt für einen physikalischen Körper, dessen Ausdehnung eine Rolle spielt, ist das des *starrten Körpers*, der Thema des Abschn. IV.2 sein wird.

I.1.2 b Kinematische Größen

Die zeitliche Reihenfolge der sukzessiven Positionen eines (Massen)Punkts im Ortsraum bildet seine *Bahnkurve* oder *Trajektorie*. Unter Betrachtung der Vektorraum-Struktur des Ortsraums kann

⁽⁴⁾Hier steht das Gleichheitszeichen für „wird dargestellt durch“.

diese Bahnkurve als eine Funktion $\vec{x}(t) : \mathbb{R} \rightarrow \mathcal{E}_3$ angesehen werden — oder äquivalent, nach Angabe eines Koordinatensystems, als eine Abbildung von \mathbb{R} nach \mathbb{R}^3 .

Bemerkung: Mathematisch genauer sollte man zwischen der Trajektorie — was im engeren Sinne die geometrische Kurve in \mathcal{E}_3 oder \mathbb{R}^3 bezeichnet — und der *Zeit-Ort-Funktion* $\vec{x}(t)$ unterscheiden, wobei die letztere die Parametrisierung der Kurve — mathematisch ein *Weg* — durch die Zeit t ist. In der Tat kann die Position auf einer Bahnkurve nicht nur durch die Zeit, sondern durch andere Parameter parametrisiert werden, wie es im Beispiel 1 des § I.1.3 b der Fall sein wird.

Gegeben die durch die Zeit parametrisierte Bahnkurve $\vec{x}(t)$ eines Massenpunkts, ist dessen instantane *Geschwindigkeit* zur Zeit t die Rate der Änderung der Zeit-Ort-Funktion zu diesem Zeitpunkt, d.h.

$$\vec{v}(t) \equiv \frac{d\vec{x}(t)}{dt} \equiv \dot{\vec{x}}(t), \quad (\text{I.2})$$

wobei hier und im Folgenden der Überpunkt eine zeitliche Ableitung darstellt. Dabei ist der Vektor $\vec{v}(t)$ ein Tangentialvektor zur Bahnkurve im Punkt $\vec{x}(t)$. Die physikalische Dimension der Geschwindigkeit ist $[v] = \text{L T}^{-1}$, wobei L für Länge und T für Zeit (time) steht; dementsprechend ist die Einheit im SI-System der $\text{m} \cdot \text{s}^{-1}$.

Wiederum ist die *Beschleunigung* des Massenpunkts als die zweite Ableitung der Zeit-Ort-Funktion nach der Zeit definiert

$$\vec{a}(t) \equiv \frac{d^2\vec{x}(t)}{dt^2} \equiv \ddot{\vec{x}}(t) = \dot{\vec{v}}(t), \quad (\text{I.3})$$

mit physikalischer Dimension bzw. SI-Einheit $[a] = \text{L T}^{-2}$ bzw. $\text{m} \cdot \text{s}^{-2}$.

Schließlich wird der (*kinetische*) *Impuls* des Massenpunkts als das Produkt aus seiner Masse und Geschwindigkeit definiert

$$\vec{p}(t) \equiv m\vec{v}(t), \quad (\text{I.4})$$

während sein *Drehimpuls* bezüglich des Nullpunkts $\vec{r} = \vec{0}$ des Bezugssystems das Kreuzprodukt aus Ortsvektor und Impuls ist

$$\vec{L}(t) \equiv \vec{x}(t) \times \vec{p}(t). \quad (\text{I.5})$$

Die jeweiligen physikalischen Dimensionen und SI-Einheiten sind $[p] = \text{M L T}^{-1}$, in $\text{kg} \cdot \text{m} \cdot \text{s}^{-1}$, und $[L] = \text{M L}^2 \text{T}^{-1}$ mit Einheit $\text{kg} \cdot \text{m}^2 \cdot \text{s}^{-1}$, wobei M für Masse steht.

I.1.3 Mechanische Kräfte

Erfahrungsgemäß spiegeln Änderungen des Bewegungszustands eines physikalischen Körpers die Wirkung von äußeren Ursachen wider, welche (mechanische) *Kräfte* heißen. Diese werden hiernach allgemein mit dem Zeichen F bezeichnet. Die physikalische Dimension einer Kraft ist $[F] = \text{M L T}^{-2}$ und die zugehörige SI-Einheit ist das Newton, wobei $1 \text{ N} = 1 \text{ kg} \cdot \text{m} \cdot \text{s}^{-2}$.

Nach der Einführung der mathematischen Modellierung von mechanischen Kräften (§ I.1.3 a) wird die sog. Arbeit einer Kraft entlang einer Wegstrecke definiert (§ I.1.3 b). In vielen Situationen hängt diese Arbeit nur von den Endpunkten der zurückgelegten Strecke ab; in solchen Fällen entspricht die durch die Kraft verrichtete Arbeit (dem Negativen) der Änderung der potentiellen Energie des Systems (§ I.1.3 c).

Bemerkung: Neben den „wirksamen“ Kräften, die zu einer Beschleunigung führen — wie z.B. die für den freien Fall eines Körpers verantwortliche Schwerkraft —, gibt es auch Kräfte, deren „Rolle“ darin besteht, die möglichen Bewegungen einzuschränken. Ein Beispiel davon ist die durch einen

Tisch auf einen Körper ausgeübte Kraft, die zusammen mit der Schwerkraft dazu führt, dass die Bewegung des Körpers unter dem Einfluss anderer Kräfte in der Tischebene bleibt. Solche Kräfte werden oft als *Zwangskräfte* bezeichnet.

I.1.3a Mathematische Modellierung

Neben ihrem Betrag hat eine Kraft erfahrungsgemäß auch eine Richtung: somit ist die Schwerkraft auf einen Körper an der Erdoberfläche „nach unten“, Reibungskräfte sind „entgegengesetzt zur Bewegungsrichtung“⁽⁵⁾, usw. Dementsprechend werden Kräfte in der newtonschen Mechanik durch dreidimensionale Vektoren \vec{F} dargestellt.

Diese Modellierung durch Vektoren geht mit einer wichtigen Eigenschaft einher, und zwar mit einem *Superpositionsprinzip*. Laut dem letzteren sind Kräfte in der newtonschen Mechanik additiv, d.h. wenn zwei unterschiedliche Kräfte \vec{F}_1, \vec{F}_2 auf einen Körper wirken, dann ist ihre Resultierende durch die Summe $\vec{F}_1 + \vec{F}_2$ gegeben. Dank der Vektorraum-Struktur ist diese Summe wieder ein Vektor, d.h. kann eine Kraft darstellen.

Dieses Superpositionsprinzip gilt nicht mehr im Rahmen der Allgemeinen Relativitätstheorie.

Wie wir unten weiter sehen werden — vgl. der dritten Bemerkung nach Gl. (I.14b) —, kann die mathematische Form der Kraft auf einen bewegten Körper in den üblichen Fällen eine Funktion der Zeit t , der Position $\vec{x}(t)$ und der Geschwindigkeit $\vec{\dot{x}}(t)$ des Körpers sein, d.h. $\vec{F} = \vec{F}(t, \vec{x}(t), \vec{\dot{x}}(t))$. Falls \vec{F} nicht von der Geschwindigkeit abhängt, sondern nur vom Ort (und von der Zeit), wird $\vec{F}(t, \vec{r})$ *Kraftfeld* genannt.

Sei \vec{F} eine Kraft. Das zugehörige *Drehmoment* bezüglich eines Bezugspunkts O wird definiert als

$$\vec{\mathcal{M}} \equiv \vec{r} \times \vec{F}, \quad (\text{I.6})$$

wobei \vec{r} der Abstandsvektor von O zum Angriffspunkt der Kraft ist

I.1.3b Arbeit einer Kraft

Wird die Wirkung einer Kraft über ein Zeitintervall — oder genauer über die Strecke, welche das mechanische System in diesem Zeitintervall zurücklegt — passend integriert, so ergibt sich die Arbeit der Kraft.

Definition: Die durch eine Kraft \vec{F} geleistete *Arbeit* in der Verschiebung eines Körpers um das infinitesimale Wegelement $d\vec{\ell}$ entlang seiner Bahnkurve ist

$$dW = \vec{F} \cdot d\vec{\ell}. \quad (\text{I.7a})$$

Die hier verwendete Notation dW ist relativ standard, bedeutet aber nicht, dass dW das totale Differential einer Funktion W ist — was nur gilt, wenn \vec{F} konservativ ist, s. § I.1.3 c. Eine bessere Notation wäre δW , wie in der Thermodynamik oder der Statistischen Mechanik üblich ist.

Aus Gl. (I.7a) folgt die Arbeit einer Kraft entlang einer endlichen Wegstrecke \mathcal{C} , indem die elementaren Beiträge entlang infinitesimaler Wegelemente summiert werden. Daraus ergibt sich ein Kurvenintegral entlang des Wegs \mathcal{C} :

$$W = \int_{\mathcal{C}} \vec{F} \cdot d\vec{\ell}. \quad (\text{I.7b})$$

Die physikalische Dimension der Arbeit einer Kraft ist $[W] = \text{ML}^2\text{T}^{-2}$.

Zur Berechnung des Kurvenintegral in Gl. (I.7b) muss man in der Praxis eine Parametrisierung des Wegs \mathcal{C} einführen. Oft, aber nicht unbedingt, kann die Zeit t als Parameter benutzt werden.

⁽⁵⁾Genauer, zur lokalen Bewegungsrichtung im Punkt wo die Reibungskraft wirkt.

Dann lässt sich \mathcal{C} genau durch die Zeit-Ort-Funktion $\vec{x}(t)$ für $t \in [t_1, t_2]$ beschreiben, und das infinitesimale Wegelement im Kurvenintegral ist $d\vec{\ell} = \vec{v}(t) dt$. Somit lautet die Arbeit entlang dem Weg \mathcal{C}

$$W = \int_{\mathcal{C}} dW = \int_{\mathcal{C}} \vec{F} \cdot d\vec{\ell} = \int_{t_1}^{t_2} \vec{F} \cdot \vec{v}(t) dt \quad (\text{I.8})$$

und das Kurvenintegral wird zu einem gewöhnlichen Integral. Dabei ist das Produkt $\vec{F} \cdot \vec{v}(t)$ die (instantane) *Leistung* der Kraft.

Beispiel 1:

Betrachte man die 2-dimensionale Bewegung (unter irgendeinem nicht-spezifizierten Einfluss) eines Massenpunkts, welcher der Kraft

$$\vec{F}(x, y) = \begin{pmatrix} ay \\ b \end{pmatrix} \equiv \begin{pmatrix} F_x(x, y) \\ F_y(x, y) \end{pmatrix}$$

mit Konstanten a, b unterliegt. Der Massenpunkt bewegt sich von einem Ausgangspunkt O mit Koordinaten $(x=0, y=0)$ zu einem Endpunkt P mit $(x=L, y=L)$. Wir wollen die Arbeit von \vec{F} entlang zwei unterschiedlicher Wege $\mathcal{C}_1, \mathcal{C}_2$ von O nach P berechnen.

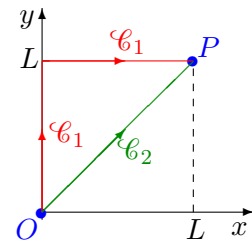


Abbildung I.1

Die Kurve \mathcal{C}_1 besteht aus zwei geraden Linienelementen mit $x=0, 0 \leq y \leq L$ bzw. $y=L, 0 \leq x \leq L$. Somit lässt sich das Kurvenintegral entlang \mathcal{C}_1 als Summe von zwei einfachen Integralen schreiben:

$$W_1 = \int_{\mathcal{C}_1} dW = \int_{\mathcal{C}_1} \vec{F}(x, y) \cdot d\vec{\ell} = \int_O^I \vec{F}(x=0, y) \cdot d\vec{\ell} + \int_I^P \vec{F}(x, y=L) \cdot d\vec{\ell},$$

mit dem Punkt I mit Koordinaten $(x=0, y=L)$. Entlang des Linienelements von O nach I bzw. von I nach P kann man für \mathcal{C}_1 einfach y bzw. x als Parameter benutzen, was zu

$$W_1 = \int_0^L F_y(x=0, y) dy + \int_0^L F_x(x, y=L) dx$$

führt. Ersetzt man die Komponenten F_x, F_y durch ihre Ausdrücke, so kommt

$$W_1 = \int_0^L b dy + \int_0^L aL dx = bL + aL^2.$$

Wiederum lässt sich die Kurve \mathcal{C}_2 als

$$\vec{x}(s) = \begin{pmatrix} x(s)=s \\ y(s)=s \end{pmatrix}$$

mit $s \in [0, L]$ parametrisieren; dann gilt $d\vec{\ell} = \frac{d\vec{x}}{ds} ds = \begin{pmatrix} 1 \\ 1 \end{pmatrix} ds$.

Dies gibt für die Arbeit von \vec{F} entlang \mathcal{C}_2

$$W_2 = \int_{\mathcal{C}_2} dW = \int_0^L \vec{F}(x(s), y(s)) \cdot \frac{d\vec{x}}{ds} ds.$$

Indem man das Skalarprodukt explizit schreibt, kommt

$$W_2 = \int_0^L (as + b) ds = \left[\frac{as^2}{2} + bs \right]_0^L = \frac{aL^2}{2} + bL.$$

Somit ist $W_2 \neq W_1$: die Arbeit der Kraft \vec{F} zwischen zwei Punkten hängt vom gewählten Weg ab.

Beispiel 2: Lorentz-Kraft

Eine bewegte Punktladung in einem Magnetfeld \vec{B} — in Abwesenheit von elektrischem Feld — unterliegt der Lorentz-Kraft^(b)

$$\vec{F} = q\vec{v} \times \vec{B}. \quad (\text{I.9})$$

Diese Kraft ist immer senkrecht zur Geschwindigkeit \vec{v} der Punktladung. Setzt man diese Kraft in Gl. (I.8) ein,

$$W = \int_{t_1}^{t_2} \vec{F} \cdot \vec{v}(t) dt = \int_{t_1}^{t_2} [q\vec{v}(t) \times \vec{B}] \cdot \vec{v}(t) dt,$$

so findet man sofort, dass die Lorentz-Kraft keine Arbeit verrichtet, denn das Integrand ist null.

I.1.3c Konservative Kräfte

Definition: Ein zeitunabhängiges Kraftfeld $\vec{F}(\vec{r})$ wird *konservativ* genannt, wenn es ein Skalarfeld⁽⁶⁾ $V(\vec{r})$ gibt, das

$$\vec{F}(\vec{r}) = -\vec{\nabla}V(\vec{r}) \quad (\text{I.10})$$

erfüllt. $V(\vec{r})$ wird *Potential* oder auch (und eigentlich genauer) *potentielle Energie* genannt.

In Gl. (I.10) bezeichnet der Nabla-Operator $\vec{\nabla}$, angewandt auf eine skalare Funktion auf \mathbb{R}^3 , den *Gradienten* dieser Funktion.

Für konservative Kraftfelder hängt die durch die Kraft zwischen zwei Punkten verrichtete Arbeit nicht vom Weg ab.

Beweis: Das Resultat folgt aus

$$W = \int_{\vec{r}_1}^{\vec{r}_2} \vec{F}(\vec{x}) \cdot d\vec{x} = - \int_{\vec{r}_1}^{\vec{r}_2} [\vec{\nabla}V(\vec{x})] \cdot d\vec{x} = -[V(\vec{r}_2) - V(\vec{r}_1)] \quad (\text{I.11})$$

unter Verwendung der Tatsache, dass $V(\vec{r})$ eine Stammfunktion von $\vec{\nabla}V(\vec{r})$ ist. \square

Bemerkung: Offensichtlich ist $V(\vec{r})$ nur bis auf eine additive Konstante eindeutig. Die Letztere wird oft so gewählt, dass das Potential im Unendlichen verschwindet.

Behauptung: Sei ein Kraftfeld $\vec{F}(\vec{r})$, definiert auf einem *einfach zusammenhängenden Gebiet* \mathcal{G} von \mathbb{R}^3 .⁽⁷⁾ Dann ist $\vec{F}(\vec{r})$ genau dann konservativ, wenn seine Rotation in jedem Punkt von \mathcal{G} verschwindet:

$$\vec{F}(\vec{r}) = -\vec{\nabla}V(\vec{r}) \quad \Leftrightarrow \quad \vec{\nabla} \times \vec{F}(\vec{r}) = \vec{0} \quad \forall \vec{r} \in \mathcal{G}. \quad (\text{I.12})$$

Dass die Rotation eines konservativen Kraftfeldes null ist, folgt direkt aus der Identität

$$\vec{\nabla} \times [\vec{\nabla}V(\vec{r})] = \vec{0},$$

die sich z.B. komponentenweise beweisen lässt: die i -te Komponente (in einem kartesischen Koordinatensystem) des Terms auf der linken Seite ist $\epsilon^{ijk} \partial_j \partial_k V(\vec{r})$, wobei ϵ^{ijk} das Levi-Civita-Symbol^(c) und ∂_ℓ die Ableitung nach der Komponenten x^ℓ des Ortsvektors ist. Dabei ist ϵ^{ijk} antisymmetrisch unter dem Austausch von j und k , während die zweite Ableitung $\partial_j \partial_k$ symmetrisch ist, so dass die Summe über alle Werte dieser Indizes Null ergibt.

Zum Beweis, dass sich ein rotationsfreies Kraftfeld $\vec{F}(\vec{r})$ als (das Negative des) Gradienten eines skalaren Potentials $V(\vec{r})$ schreiben lässt, soll man zunächst einen beliebigen Punkt \vec{r}_0 in \mathcal{G} wählen

⁽⁶⁾D.h. eine zahlenwertige Funktion des Ortsvektors \vec{r} .

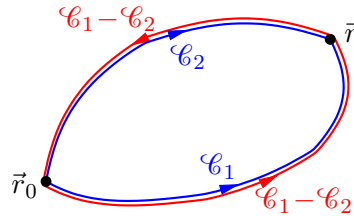
⁽⁷⁾Ein Gebiet wird als *einfach zusammenhängend* bezeichnet, wenn sich jeder geschlossene Weg im Gebiet stetig zu einem Punkt zusammenziehen lässt, ohne das Gebiet zu verlassen. Zum Beispiel ist eine Ebene (\mathbb{R}^2) einfach zusammenhängend, während \mathbb{R}^2 ohne einen Punkt nicht mehr einfach zusammenhängend ist.

^(b)H. A. LORENTZ, 1853–1926 ^(c)T. LEVI-CIVITA, 1873–1941

und V durch

$$V(\vec{r}) \equiv - \int_{\vec{r}_0}^{\vec{r}} \vec{F}(\vec{r}') \cdot d\vec{r}'$$

für jeden $\vec{r} \in \mathcal{G}$ definieren. Das Integral auf der rechten Seite der obigen Formel ist eindeutig definiert, wenn das Linienintegral $\int \vec{F} \cdot d\vec{x}$ von \vec{r}_0 nach \vec{r} unabhängig vom gewählten Weg ist. Betrachte man zwei unterschiedliche Wege $\mathcal{C}_1, \mathcal{C}_2$ in \mathcal{G} , die von \vec{r}_0 nach \vec{r} führen. Dann definiert



\mathcal{C}_2 trivial einen Weg $-\mathcal{C}_2$ (mit gleicher Kurve und umgekehrter Parametrisierung) von \vec{r} nach \vec{r}_0 , und somit einen geschlossenen Weg $\mathcal{C}_1 - \mathcal{C}_2$ mit Anfangs- und Endpunkt in \vec{r}_0 . Aus dem Satz von Stokes^(d) folgt für das Wegintegral der Kraft entlang $\mathcal{C}_1 - \mathcal{C}_2$

$$\oint_{\mathcal{C}_1 - \mathcal{C}_2} \vec{F}(\vec{r}') \cdot d\vec{r}' = \int_{\mathcal{S}} [\vec{\nabla} \times \vec{F}(\vec{r}')] \cdot d^2\vec{\mathcal{S}} = 0,$$

wobei \mathcal{S} die durch die Kurve $\mathcal{C}_1 - \mathcal{C}_2$ abgeschlossene Fläche bezeichnet, während die letzte Gleichheit aus der Annahme $\vec{\nabla} \times \vec{F} = \vec{0}$ folgt, die in jedem Punkt von \mathcal{S} gilt (dank der Annahme eines einfach zusammenhängenden Gebiets). Andererseits gilt

$$\oint_{\mathcal{C}_1 - \mathcal{C}_2} \vec{F}(\vec{r}') \cdot d\vec{r}' = \int_{\mathcal{C}_1} \vec{F}(\vec{r}') \cdot d\vec{r}' - \int_{\mathcal{C}_2} \vec{F}(\vec{r}') \cdot d\vec{r}',$$

woraus die gesuchte Unabhängigkeit des Wegintegrals vom Weg folgt. □

1.2 Newtonsche Gesetze

Basierend auf den in Abschn. I.1 eingeführten Begriffen, insbesondere auf der Existenz absoluter Raum und Zeit, können die dynamischen Gesetze, die die Wirkung von Kräften auf mechanische Systeme bestimmen, angegeben werden. Dabei besagen diese Gesetze aber nichts über die mathematische Form der Kräfte: diese hängen von der Situation ab und sollen durch zusätzliche Theorien oder Modelle präzisiert werden.

1.2.1 Erstes newtonsches Gesetz

Das erste newtonsche Gesetz — das auch *erstes Axiom*, *lex prima* oder *Trägheitsgesetz* genannt wird — besagt die Existenz von bevorzugten Bezugssystemen, in denen der Bewegungszustand eines Körpers sich nicht ändert, so lange keine Kraft auf ihn ausgeübt wird:

Erstes newtonsches Gesetz

Es gibt besondere Bezugssysteme, sog. Inertialsysteme, in denen ein Massenpunkt, der keiner Kraft unterliegt, in seinem Zustand der Ruhe oder gleichförmigen geradlinigen Bewegung beharrt. (I.13)

Anders gesagt bleibt in Inertialsystemen die Geschwindigkeit $\dot{\vec{x}}(t)$ eines *kräftefreien* Massenpunkts konstant, d.h. seine Beschleunigung $\ddot{\vec{x}}(t)$ verschwindet.

Bemerkungen:

* Eigentlich wurde dieses Prinzip schon durch Galilei^(e) formuliert, weshalb es auch seinen Namen trägt. Dazu ist der obige Ausdruck des Gesetzes in der Tat nicht die ursprüngliche Formulierung,

^(d)G. G. STOKES, 1819–1903 ^(e)G. GALILEI, 1564–1642

denn es gab bei Newton keinen Bezug auf Inertialsysteme.

* Das Gesetz setzt die Existenz eines kräftefreien Zustands implizit voraus. In der Praxis ist dies eine Idealisierung, denn kein Körper kann von der (anziehenden) Schwerkraft der anderen Körper im Universum isoliert werden.

Was aber praktisch realisierbar ist — zumindest in sehr guter Näherung —, ist dass die Resultierende der auf ein System wirkenden Kräfte verschwindet, was sich dann als äquivalent zur Abwesenheit von Kräfte herausstellt.

* Auf Inertialsysteme wird in Abschn. I.3 genauer eingegangen; als solches gilt für Experimente auf der Erde ein Bezugssystem, das sich relativ zu „Fixsternen“, oder besser weit entfernten Galaxien oder Quasars, nicht bewegt.

I.2.2 Zweites newtonsches Gesetz

Das zweite newtonsche Gesetz — auch *zweites Axiom*, *lex secunda* oder *Bewegungsgesetz* genannt — beschreibt die Änderung der Bewegung eines Körpers mit Impuls $\vec{p}(t)$ unter dem Einfluss einer Gesamtkraft \vec{F} :

Zweites newtonsches Gesetz

$$\text{In Inertialsystemen gilt } \frac{d\vec{p}(t)}{dt} \equiv \dot{\vec{p}}(t) = \vec{F}. \quad (\text{I.14a})$$

In dieser *Bewegungsgleichung* kann der Impuls $\vec{p}(t)$ durch die Masse m und die Geschwindigkeit $\vec{v}(t)$ des Körpers ausgedrückt werden. Wenn die Masse in der Bewegung konstant bleibt, so kann sie aus der Zeitableitung herausgezogen werden. Dann wird Gl. (I.14a) zu

$$m \frac{d^2\vec{x}(t)}{dt^2} = m \frac{d\vec{v}(t)}{dt} = m\vec{a}(t) = \vec{F}. \quad (\text{I.14b})$$

Das heißt, die Beschleunigung des Körpers ist proportional zur Kraft, welche die Bewegung beeinflusst, und antiproportional zur Masse m .

Bemerkungen:

- * Das erste Gesetz ist ein Spezialfall des zweiten mit $\vec{F} = \vec{0}$.
- * Die Masse m in der Bewegungsgleichung (I.14b) ist die sog. *träge Masse* des Massenpunkts, die ein Maß für die Trägheit des letzteren gegenüber Änderungen seines Bewegungszustands darstellt. Bei der trägen Masse handelt es sich um eine universale Eigenschaft des Körpers, welche die gleiche bleibt, egal welchen Kräften er unterliegt.
- * Im zweiten Gesetz wird implizit vorausgesetzt, dass die Kraft \vec{F} auf einen Körper möglicherweise von seiner Position \vec{x} und seiner Geschwindigkeit $\dot{\vec{x}}$ abhängen kann, sowie von der Zeit t , nicht aber von höheren Ableitungen.

Dank der letzteren Bemerkung stellt die vektorielle Bewegungsgleichung (I.14b) ein System gewöhnlicher Differentialgleichungen zweiter Ordnung dar. Wenn die mathematische Form der Kraft $\vec{F} = \vec{F}(t, \vec{x}(t), \dot{\vec{x}}(t))$ sowie Anfangsbedingungen $\vec{x}(t_0), \dot{\vec{x}}(t_0)$ zu irgend einer Referenzzeit t_0 gegeben sind, dann hat dieses System eine eindeutige Lösung $\vec{x}(t)$.

Physikalisch bedeutet diese Eindeutigkeit der Lösung, dass die Kenntnis des Bewegungszustands ($\vec{x}(t_0), \dot{\vec{x}}(t_0)$) zu einem beliebigen Zeitpunkt t_0 ausreicht, um den Bewegungszustand ($\vec{x}(t), \dot{\vec{x}}(t)$) zu jedem anderen Zeitpunkt t vollständig festzustellen. Dementsprechend ist die newtonsche Mechanik völlig *deterministisch*.

Beispiel: Massenpunkt im homogenen Schwerfeld mit Luftreibung

Betrachtet sei die Bewegung eines Massenpunkts mit Masse m unter dem Einfluss der Schwerkraft und einer Reibungskraft. Die Resultierende der auf ihn wirkenden Kraft lautet

$$\vec{F} = \vec{F}_S + \vec{F}_R, \quad (\text{I.15a})$$

wobei das Superpositionsprinzip für Kräfte (I.20) benutzt wurde. Dabei ist die Schwerkraft

$$\vec{F}_S \equiv m_S \vec{g}, \quad (\text{I.15b})$$

wobei die Schwerebeschleunigung \vec{g} prinzipiell nach unten gerichtet ist⁽⁸⁾ — obwohl das im Folgenden keine Rolle spielt —, während m_S die *schwere Masse* des Körpers bezeichnet. Experimentell ist die letztere proportional zur trägen Masse, $m_S \propto m$, wie sich im berühmten Ergebnis „im Vakuum fallen alle Körper gleich“ widerspiegelt. Mit der Wahl $m_S = m$ gilt $|\vec{g}| \simeq 9,8 \text{ m} \cdot \text{s}^{-2}$ auf der Erdoberfläche.

Hier werden zwei Annahmen gemacht, und zwar dass träge und schwere Masse erstens die gleiche Dimension M haben, und sich somit mit der gleichen Einheit quantifizieren lassen, und zweitens denselben numerischen Wert haben, was wiederum den Wert der Schwerkraftbeschleunigung g festlegt.

Für die Reibungskraft wird die Form der Stokes'schen Reibung

$$\vec{F}_R \equiv -\alpha_R \vec{v} \quad (\text{I.15c})$$

angenommen, mit einer positiven Konstanten α_R . Dabei ist \vec{v} die Geschwindigkeit des Massenpunkts relativ zum „Medium“ (z.B. Luft), das die Reibung verursacht.

Unter diesen Voraussetzungen lautet die Bewegungsgleichung (I.14b)

$$m \frac{d^2 \vec{x}(t)}{dt^2} = m \vec{g} - \alpha_R \vec{v}(t),$$

d.h., mit $\gamma_R \equiv \alpha_R/m$

$$\frac{d\vec{v}(t)}{dt} + \gamma_R \vec{v}(t) = \vec{g}. \quad (\text{I.16})$$

Somit ergibt sich eine lineare inhomogene gewöhnliche Differentialgleichung erster Ordnung. Ihre Lösung erfolgt wie üblich.

Zuerst wird die allgemeine Lösung \vec{v}_h der assoziierten homogenen Differentialgleichung

$$\frac{d\vec{v}_h(t)}{dt} + \gamma_R \vec{v}_h(t) = 0$$

gesucht, und zwar

$$\vec{v}_h(t) = \vec{C} e^{-\gamma_R t}$$

mit beliebigem $\vec{C} \in \mathbb{R}^3$.

Zweitens ist eine spezielle Lösung \vec{v}_s der homogenen Differentialgleichung gebraucht:

$$\frac{d\vec{v}_s(t)}{dt} + \gamma_R \vec{v}_s(t) = \vec{g}.$$

Zum Beispiel kann die stationäre (d.h. zeitunabhängige) Lösung

$$\vec{v}_s(t) = \frac{1}{\gamma_R} \vec{g}$$

gewählt werden.

Mit deren Hilfe kann als dritter Schritt die allgemeine Lösung der inhomogenen Differentialgleichung (I.16) als Summe aus $\vec{v}_h(t)$ und $\vec{v}_s(t)$ geschrieben werden:

$$\vec{v}(t) = \vec{v}_h(t) + \vec{v}_s(t) = \vec{C} e^{-\gamma_R t} + \frac{1}{\gamma_R} \vec{g}.$$

⁽⁸⁾Definitionsgemäß gibt die lokale Richtung der Schwerebeschleunigung die sog. Vertikale in einem Punkt.

Dabei soll noch die Integrationskonstante \vec{C} festgestellt werden, was unter Berücksichtigung der Anfangsbedingung (hier bei $t = t_0$)

$$\vec{v}_0 \equiv \vec{v}(t=t_0) = \vec{C} e^{-\gamma_{\text{R}} t_0} + \frac{1}{\gamma_{\text{R}}} \vec{g}$$

erfolgt. Aus der letzteren Gleichung folgt

$$\vec{C} = \vec{v}_0 e^{\gamma_{\text{R}} t_0} - \frac{e^{\gamma_{\text{R}} t_0}}{\gamma_{\text{R}}} \vec{g},$$

was schließlich zur Lösung

$$\vec{v}(t) = \vec{v}_0 e^{-\gamma_{\text{R}}(t-t_0)} + \frac{1 - e^{-\gamma_{\text{R}}(t-t_0)}}{\gamma_{\text{R}}} \vec{g} \quad (\text{I.17})$$

führt. Physikalisch findet man, dass der Einfluss der Anfangsgeschwindigkeit \vec{v}_0 wegen der Reibungskraft auf einer Zeitskala $\tau_{\text{R}} \equiv \gamma_{\text{R}}^{-1}$ exponentiell abnimmt. Dies führt zu einem stationären Endzustand, der unabhängig von der Anfangsbedingung ist.

Wenn nötig kann man auch die Bahnkurve des Massenpunkts über

$$\vec{x}(t) = \vec{x}_0 + \int_{t_0}^t \vec{v}(t') dt'$$

bestimmen, wobei $\vec{x}_0 \equiv \vec{x}(t=t_0)$ die Anfangsposition bei $t = t_0$ bezeichnet, und zwar

$$\vec{x}(t) = \vec{x}_0 + \frac{1 - e^{-\gamma_{\text{R}}(t-t_0)}}{\gamma_{\text{R}}} \vec{v}_0 + \frac{\gamma_{\text{R}}(t-t_0) + e^{-\gamma_{\text{R}}(t-t_0)} - 1}{\gamma_{\text{R}}^2} \vec{g}. \quad (\text{I.18})$$

Nach Abklingen der Exponentialterme, bleibt eine annähernd lineare Funktion der Zeit, in Kontrast zur quadratischen Abhängigkeit beim reibungsfreien Fall.

Die neugierige Leserin kann den Limes $\gamma_{\text{R}} \rightarrow 0$ in Gl. (I.17) und (I.18) nehmen — anhand einer Taylor-Entwicklung^(f) der Exponentialfunktion —, um die bekannten Ergebnisse des freien Falls wiederzuentdecken.

I.2.3 Drittes newtonsches Gesetz

Das dritte newtonsche Gesetz — auch bekannt als *drittes Axiom*, *lex tertia*, *Reaktionsprinzip* oder *Actio und Reactio* — verknüpft diejenigen Kräfte, welche zwei Körper aufeinander ausüben:

Drittes newtonsches Gesetz

Die von zwei Massenpunkten aufeinander ausgeübten Kräfte haben den gleichen Betrag aber entgegengesetzte Richtungen. (I.19a)

Wenn die zwei Körper jeweils durch 1 und 2 gekennzeichnet werden, und $\vec{F}_{2 \rightarrow 1}$ bzw. $\vec{F}_{1 \rightarrow 2}$ die auf den ersten durch den zweiten bzw. auf 2 durch 1 ausgeübte Kraft bezeichnet, dann gilt

$$\vec{F}_{2 \rightarrow 1} = -\vec{F}_{1 \rightarrow 2}. \quad (\text{I.19b})$$

Bemerkungen:

* Manchmal wird das Gesetz auch *Aktion gleich Reaktion* genannt: diese Bezeichnung hat aber den Nachteil, dass sie sich nur auf die Beträge der Kräfte bezieht, was irreführend sein kann.

* Wie in der Bemerkung in § II.1.1 a weiter diskutiert wird, schließt dieses Gesetz implizit die Existenz von Drei-Körper-Kräften aus.

^(f)B. TAYLOR, 1685–1731

* Laut dem dritten Gesetz ist die Reaktion von Körper 2 auf Körper 1 instantan, wenn Körper 1 eine Kraft auf Körper 2 ausübt. Dies wird in der Relativitätstheorie nicht mehr möglich sein, wenn sich die Körper nicht berühren.

1.2.4 „Viertes newtonsches Gesetz“

Als viertes newtonsches Gesetz, oder *lex quarta*, wird oft das Superpositionsprinzip für Kräfte hinzugefügt, das in § I.1.3 schon dargelegt wurde:

Viertes newtonsches Gesetz

Wirken auf einen Massenpunkt mehrere Kräfte $\vec{F}_1, \dots, \vec{F}_k$, so addieren sich diese vektoriell. (I.20a)

Das heißt, die resultierende Kraft auf den Massenpunkt ist

$$\vec{F} \equiv \sum_{i=1}^k \vec{F}_i. \quad (\text{I.20b})$$

Es ist dann diese resultierende Kraft, die in der mathematischen Formulierung des zweiten Gesetzes (I.14) auftritt.

1.2.5 Energieerhaltung

Eine erste Folgerung der newtonschen Gesetze ist der Zusammenhang zwischen der Änderung der kinetischen Energie eines Massenpunkts und der Arbeit der auf ihn ausgeübten Kräfte.

Definition: Die *kinetische Energie* eines Massenpunkts mit Masse m und Impuls \vec{p} bzw. Geschwindigkeit \vec{v} wird definiert als

$$T \equiv \frac{\vec{p}^2}{2m} = \frac{1}{2}m\vec{v}^2. \quad (\text{I.21})$$

Bemerkung: Die physikalische Dimension bzw. die SI-Einheit einer Energie ist — wie bei einer mechanischen Arbeit! — $[E] = \text{M L}^2 \text{T}^{-2}$ bzw. das Joule^(g) ($1 \text{ J} = 1 \text{ kg} \cdot \text{m}^2 \cdot \text{s}^{-2}$).

Für einen Massenpunkt, der nur konservativen Kräften unterliegt, gilt der

Theorem (Energieerhaltungssatz):

Wenn alle Kräfte konservativ sind, ist die Summe $T + V \equiv E$ aus kinetischer und potentieller Energie eines Massenpunkts, entsprechend seiner Gesamtenergie, erhalten. (I.22)

Die Erweiterung dieses Ergebnisses auf Systeme aus mehreren Massenpunkten wird in § II.1.4 a dargelegt.

Der Satz lässt sich beweisen, indem das Integral⁽⁹⁾

$$\int_{\vec{r}_1}^{\vec{r}_2} \vec{F}(\vec{x}) \cdot d\vec{x}$$

auf zwei unterschiedliche Weisen geschrieben wird. Da die Kraft konservativ ist, ist das Integral einerseits laut Gl. (I.11) gleich $V(\vec{r}_1) - V(\vec{r}_2)$.

⁽⁹⁾Die Eindeutigkeit der Definition des Integrals wurde im Beweis der Behauptung in § I.1.3c gezeigt.

^(g)J. P. JOULE, 1818–1889

Andererseits gilt, unter Verwendung der newtonschen Bewegungsgleichung (I.14b) und der Identität $d\vec{x} = \vec{v}(t) dt = \vec{p}(t) dt/m$

$$\int_{\vec{r}_1}^{\vec{r}_2} \vec{F}(\vec{x}) \cdot d\vec{x} = \int_{t_1}^{t_2} \frac{1}{m} \frac{d\vec{p}(t)}{dt} \cdot \vec{p}(t) dt = \int_{t_1}^{t_2} \frac{1}{m} \frac{d}{dt} \left[\frac{\vec{p}(t)^2}{2} \right] dt = T(t_2) - T(t_1). \quad (\text{I.23})$$

Somit gilt $V(\vec{r}_1) - V(\vec{r}_2) = T(t_2) - T(t_1)$, d.h.

$$T(t_1) + V(\vec{r}_1) = T(t_2) + V(\vec{r}_2),$$

wobei t_1 bzw. t_2 der Zeitpunkt ist, zu dem sich der Körper in \vec{r}_1 bzw. \vec{r}_2 befindet. \square

In diesem Beweis bedeutet Gl. (I.23), dass die Arbeit der auf das System wirkenden Kräfte gleich der Änderung der kinetischen Energie des Systems ist.

I.3 Inertialsysteme. Galilei-Transformationen

Das erste und das zweite newtonsche Gesetz beruhen auf der Existenz von besonderen Bezugssystemen, den Inertialsystemen, in denen ein kräftefreier Massenpunkt keine Beschleunigung erfährt. Angenommen, dass ein solches Inertialsystem \mathcal{B}_I existiert, dann gibt es noch eine ganze Menge anderer Inertialsysteme, bestehend aus den Bezugssystemen \mathcal{B}' , die mit \mathcal{B}_I über eine *Galilei-Transformation* verknüpft sind.

Sei O bzw. O' ein fester Bezugspunkt im Inertialsystem \mathcal{B}_I bzw. im Bezugssystem \mathcal{B}' . Der Ortsvektor eines geometrischen Punkts P relativ zu \mathcal{B}_I bzw. \mathcal{B}' wird mit $\vec{OP} \equiv \vec{r}$ bzw. $\vec{O'P} \equiv \vec{r}'$ bezeichnet. Auf ähnliche Weise können in jedem Bezugssystem unterschiedliche Zeiten t bzw. t' benutzt werden. Eine Transformation vom Inertialsystem \mathcal{B}_I zum Bezugssystem \mathcal{B}' ist dann eine Abbildung $(t, \vec{r}) \mapsto (t', \vec{r}')$, wobei t' und \vec{r}' Funktionen von t und \vec{r} sind.

Sei $\vec{x}(t)$ bzw. $\vec{x}'(t')$ die Trajektorie eines Massenpunkts bezüglich \mathcal{B}_I bzw. \mathcal{B}' . Für jede Klasse von „einfachen“ Galilei-Transformationen — Translationen (§ I.3.1), Galilei-Boosts (§ I.3.2) und Drehungen (§ I.3.3) — werden wir zeigen, dass die Beschleunigungen $d^2\vec{x}(t)/dt^2$ im Inertialsystem \mathcal{B}_I und $d^2\vec{x}'(t)/dt'^2$ im Bezugssystem \mathcal{B}' gleichzeitig verschwinden. Das heißt, wenn ein Massenpunkt kräftefrei ist, dann bleibt seine Geschwindigkeit in \mathcal{B}' konstant, so dass \mathcal{B}' entsprechend dem ersten newtonschen Gesetz ebenfalls inertial ist.

In § I.3.4 werden allgemeine Galilei-Transformationen eingeführt. Schließlich wird das Verhalten von Kräften unter Galilei-Transformationen in § I.3.5 diskutiert.

I.3.1 Translationen

Eine erste Klasse von einfachen Transformationen $(t, \vec{r}) \mapsto (t', \vec{r}')$ zwischen Inertialsystemen besteht aus den Translationen, entweder im Ortsraum oder in der Zeit.

I.3.1 a Räumliche Translationen

Betrachten wir zunächst den Fall eines Bezugssystems \mathcal{B}' , das sich aus dem Inertialsystem \mathcal{B}_I durch eine räumliche Translation um einen festen Vektor \vec{c} ableiten lässt. Das heißt, dass die Ortsvektoren \vec{r} bzw. \vec{r}' eines Punkts P bezüglich \mathcal{B}_I bzw. \mathcal{B}' über die Beziehung

$$\vec{r}' = \vec{r} - \vec{c} \quad (\text{I.24})$$

verknüpft sind (vgl. Abb. I.2). Es wird angenommen, dass die Beobachter in beiden Bezugssystemen die gleiche Zeit benutzen, d.h. $t' = t$.

Der gleiche Zusammenhang gilt für die Bahnkurven $\vec{x}(t)$ bzw. $\vec{x}'(t') = \vec{x}'(t)$ eines bewegten Massenpunkts. Somit gilt trivial nach Zeitableitung

$$\frac{d\vec{x}'(t')}{dt'} = \frac{d\vec{x}'(t)}{dt} = \frac{d\vec{x}(t)}{dt}, \quad \frac{d^2\vec{x}'(t')}{dt'^2} = \frac{d^2\vec{x}'(t)}{dt^2} = \frac{d^2\vec{x}(t)}{dt^2}.$$

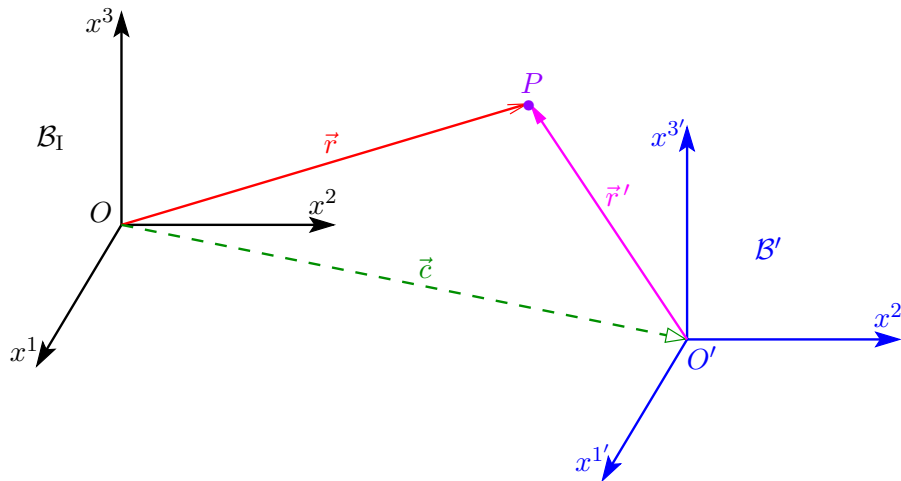


Abbildung I.2

Wenn keine Kräfte auf den Massenpunkt wirken, so dass $\ddot{\vec{x}}(t) = \vec{0}$, dann ist auch $\ddot{\vec{x}}(t') = \vec{0}$. Umgekehrt führt $\ddot{\vec{x}}(t') = \vec{0}$ zu $\ddot{\vec{x}}(t) = \vec{0}$, d.h. das erste newtonsche Gesetz (I.13) gilt genau dann in \mathcal{B}' , wenn es im Inertialsystem \mathcal{B}_I gilt.

Bemerkung: Führt man zwei Translationen (um Vektoren \vec{c}_1 und \vec{c}_2) hintereinander aus, so ist die resultierende Transformation ebenfalls eine Translation, und zwar um den Vektor $\vec{c}_1 + \vec{c}_2$.

I.3.1 b Translationen in der Zeit

Sei jetzt angenommen, dass das Bezugssystem \mathcal{B}' bezüglich des Inertialsystems \mathcal{B}_I ruht, und dass die gleichen Raumkoordinaten in \mathcal{B}' als in \mathcal{B}_I verwendet werden, d.h. $\vec{r} = \vec{r}'$ für Ortsvektoren. Dagegen unterscheidet sich der in \mathcal{B}' gewählte Nullpunkt der Zeit $t' = 0$ von dem Nullpunkt $t = 0$ in \mathcal{B}_I um eine Zeitverschiebung τ :

$$t' = t - \tau. \quad (\text{I.25})$$

Dies entspricht einer zeitlichen Translation, so wie Gl. (I.24) eine räumliche Translation darstellt.

Unter Verwendung der Kettenregel gilt dann für die Zeitableitungen in jedem Bezugssystem⁽¹⁰⁾

$$\frac{d}{dt'} = \frac{dt}{dt'} \frac{d}{dt} = \frac{d}{dt},$$

und eine ähnliche Gleichung für die zweite Ableitungen. Insbesondere ist $d^2\vec{x}(t)/dt^2 = \vec{0}$ genau äquivalent zu $d^2\vec{x}(t')/dt'^2 = \vec{0}$.

Bemerkung: Gleichung (I.25) bedeutet nicht nur, dass die Beobachter in \mathcal{B}_I und \mathcal{B}' unterschiedliche Nullpunkte der Zeit genommen haben, sondern auch, dass die Zeit gleich schnell in beiden Bezugssystemen vergeht — entsprechend der Universalität der Zeit in der newtonschen Mechanik (§ I.1.1 b). Das heißt, hier wird $t' = \alpha t - \tau$ mit $\alpha = 1$ betrachtet.

I.3.2 Eigentliche Galilei-Transformationen

Ein Bezugssystem \mathcal{B}' bewege sich gegenüber dem Inertialsystem \mathcal{B}_I mit konstanter Geschwindigkeit \vec{u} . Der Einfachheit halber wird angenommen, dass die an \mathcal{B}_I und \mathcal{B}' gebundenen Beobachter die gleichen Zeiten $t' = t$ benutzen; dazu verwenden sie Koordinatensysteme, deren Ursprungspunkte zur Zeit $t = 0$ übereinstimmen, und dessen Achsen parallel sind, wie in Abb. I.3 dargestellt wird.

⁽¹⁰⁾Die hier benutzte Notation bedeutet, dass für jede Funktion $f(t') = \varphi(t(t'))$ die Gleichung

$$\frac{df(t')}{dt'} = \frac{d\varphi(t(t'))}{dt}$$

gilt.

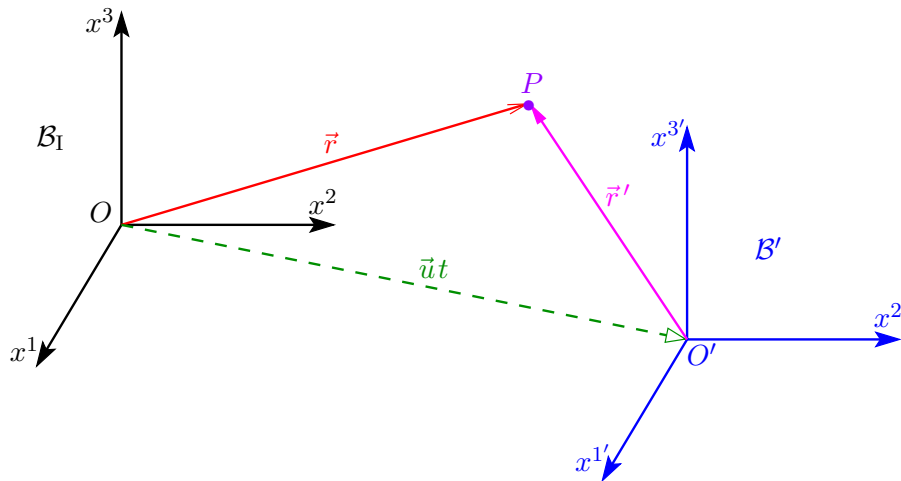


Abbildung I.3

Dann gilt für die Ortsvektoren eines bestimmten Punkts P in den beiden Bezugssystemen

$$\vec{r}' = \vec{r} - \vec{u}t. \quad (\text{I.26})$$

Eine solche Transformation nennt man *eigentliche Galilei-Transformation* oder auch *Galilei-Boost*.

Die Beziehung (I.26) führt für die Bahnkurven $\vec{x}(t)$ bzw. $\vec{x}'(t)$ eines bewegten Massenpunkts P sofort zu $\vec{x}'(t) = \vec{x}(t) - \vec{u}t$, und somit erstens zu

$$\frac{d\vec{x}'(t')}{dt'} = \frac{d}{dt} [\vec{x}(t) - \vec{u}t] = \frac{d\vec{x}(t)}{dt} - \vec{u}, \quad (\text{I.27})$$

wobei $t' = t$ verwendet wurde. Diese Beziehung bedeutet, dass die Beobachter in \mathcal{B}_I und \mathcal{B}' einem bewegten physikalischen System unterschiedliche Geschwindigkeiten zuordnen: Gleichung (I.27) stellt das (nicht-relativistische) *Additionsgesetz von Geschwindigkeiten* dar.

Eine zweite Zeitableitung liefert dann

$$\frac{d^2\vec{x}'(t')}{dt'^2} = \frac{d^2\vec{x}(t)}{dt^2},$$

so dass die durch die beiden Beobachter gemessenen Beschleunigungen gleich sind.

Falls keine Kräfte auf P wirken, so dass $\ddot{\vec{x}}(t) = \vec{0}$ in \mathcal{B}_I , dann gilt ebenfalls $\ddot{\vec{x}}'(t') = \vec{0}$ in \mathcal{B}' , d.h. was in einem der Bezugssysteme als kräftefrei aussieht, wird auch im anderen Bezugssystem als kräftefrei gesehen.

Bemerkung: Wie bei Translationen (im Raum oder in der Zeit) ergeben zwei hinter einander ausgeführte Galilei-Boosts mit jeweiligen „Boost-Geschwindigkeiten“ \vec{u}_1, \vec{u}_2 einen neuen Galilei-Boost mit Geschwindigkeit $\vec{u}_1 + \vec{u}_2$.

I.3.3 Drehungen

Die letzte Klasse „einfacher“ Galilei-Transformationen besteht aus den Drehungen um jede beliebige Achse im Raum. Sei somit angenommen, dass die Beobachter in \mathcal{B}_I und \mathcal{B}' Koordinatensysteme benutzen, deren Achsen gegenüber einander um einen konstanten Winkel θ um eine Richtung $\vec{e}_{\mathcal{R}}$ gedreht sind. Dagegen stimmen die Nullpunkte der Koordinatensysteme miteinander überein, und beide Beobachter benutzen die gleiche Zeit $t' = t$.

In diesem Fall lassen sich Ortsvektoren bezüglich \mathcal{B}' bzw. ihre Koordinaten aus denen bezüglich \mathcal{B}_I durch eine konstante 3×3 -Drehmatrix $\mathcal{R}_{\mathcal{B}'/\mathcal{B}_I}$ erhalten:⁽¹¹⁾

$$\vec{r}' = \mathcal{R}_{\mathcal{B}'/\mathcal{B}_I} \vec{r}. \quad (\text{I.28a})$$

⁽¹¹⁾Vgl. Anhang B über Drehmatrizen.

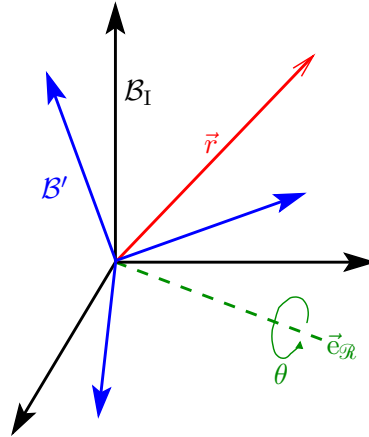


Abbildung I.4

Äquivalent lässt sich diese Beziehung komponentenweise schreiben:

$$x^{i'} = \mathcal{R}^{i'}_j x^j \quad \text{für } i' \in \{1, 2, 3\}, \quad (\text{I.28b})$$

wobei die $\mathcal{R}^{i'}_j$ für $i', j \in \{1, 2, 3\}$ die Matrixelemente von $\mathcal{R}_{\mathcal{B}'/\mathcal{B}_I}$ sind, während die $\{x^{i'}\}$ mit gestrichenen Indizes die Komponenten von \vec{r}' bezeichnen.

Bemerkung: Bei der hier betrachteten Drehung — sowie bei den Translationen in § I.3.1 oder den Galilei-Boosts in § I.3.2 — handelt es sich um eine sog. *passive* Transformation: der betrachtete Massenpunkt wird gegenüber anderen Objekten nicht gedreht (was eine *aktive* Drehung wäre). Stattdessen wird der Beobachter oder äquivalent das Bezugs- bzw. Koordinatensystem, in dem der Massenpunkt beschrieben wird, gedreht.

Beziehung (I.28a) bzw. (I.28b) gilt auch für die Bahnkurve eines Massenpunkts bezüglich jedes der Bezugssysteme bzw. für die Komponenten der Bahnkurve:

$$\vec{x}'(t') = \mathcal{R}_{\mathcal{B}'/\mathcal{B}_I} \vec{x}(t) \quad \text{bzw.} \quad x^{i'}(t') = \mathcal{R}^{i'}_j x^j(t) \quad \text{für } i' \in \{1, 2, 3\}.$$

Da diese Beziehungen linear und zeitunabhängig sind, geben sie nach zweifachen Zeitableitung unter Berücksichtigung von $t' = t$

$$\frac{d^2 \vec{x}'(t')}{dt'^2} = \mathcal{R}_{\mathcal{B}'/\mathcal{B}_I} \frac{d^2 \vec{x}(t)}{dt^2}$$

bzw.

$$\frac{d^2 x^{i'}(t')}{dt'^2} = \mathcal{R}^{i'}_j \frac{d^2 x^j(t)}{dt^2} \quad \text{für } i' \in \{1, 2, 3\}.$$

Somit führt $\vec{\ddot{x}}(t) = \vec{0}$ zu $\vec{\ddot{x}}'(t') = \vec{0}$. Da jede Drehmatrix invertierbar ist, gilt die Implikation auch in der umgekehrten Richtung.

Bemerkung: Wie bei Translationen oder Galilei-Boosts ist das Produkt zweier Drehmatrizen, entsprechend der Verkettung der assoziierten Drehungen, wieder eine Drehmatrix. Dabei ist aber das Produkt zweier Drehmatrizen im Allgemeinen nicht-kommutativ, d.h. die Reihenfolge, in der die Drehungen aufeinander folgen, ist wichtig.

I.3.4 Allgemeine Galilei-Transformationen

Generell sind die Galilei-Transformationen \mathcal{G} die linearen Abbildungen $(t, \vec{r}) \rightarrow (t', \vec{r}') = \mathcal{G}(t, \vec{r})$ zwischen den Zeiten und Positionen in zwei unterschiedlichen Inertialsystemen.

In § I.3.1–I.3.3 wurde gezeigt, dass räumliche und zeitliche Translationen, Galilei-Boosts und Drehungen Beispiele solcher Galilei-Transformationen sind. Allgemeiner definiert jedes „Produkt“ —

d.h. Hintereinanderausführung — von Translationen im Raum oder in der Zeit, Galilei-Boosts und Drehungen eine allgemeine Galilei-Transformation. Dabei hängt die resultierende Transformation von der Reihenfolge der Transformationen ab, genau wie es schon bei Drehungen der Fall ist.

Umgekehrt kann man zeigen, dass jede Galilei-Transformation zwischen zwei Inertialsystemen sich als Verkettung von Translationen, Drehungen und Galilei-Boosts zerlegen lässt.

Schließlich ist das Produkt zweier Galilei-Transformationen wieder eine Galilei-Transformation. Somit bilden diese Transformationen mit diesem Produkt mathematisch eine sog. *Gruppe*.⁽¹²⁾

Dabei benutzt man auch, dass es ein „neutrales Element“ für das Produkt gibt, und zwar die Identitätstransformation, dass jeder Galilei-Transformation eine inverse Transformation zugeordnet werden kann, und schließlich dass das Produkt assoziativ ist.

Da das Ergebnis eines Produkts von der Reihenfolge der „Multiplikatoren“ abhängt, wird die Gruppe als *nicht-kommutativ* oder äquivalent *nicht-abelsch*^(h) bezeichnet.

Bemerkung: Die Translationen — im Raum oder in der Zeit —, die Drehungen und die Galilei-Boosts bilden jeweils Untergruppen der Galilei-Gruppe.

I.3.5 Kräfte und Galilei-Transformation

In den vorigen Paragraphen wurde nur das Verhalten des Ortsvektors und seiner Ableitungen nach der Zeit unter Galilei-Transformationen von einem Inertialsystem zu einem anderen betrachtet. Daraus folgt, dass wenn die Beschleunigung $\ddot{\vec{x}}(t)$ eines Massenpunkts in einem Inertialsystem \mathcal{B}_I verschwindet, dann gilt $\ddot{\vec{x}}'(t) = \vec{0}$ für einen Beobachter, dessen Koordinatensystem mit demjenigen in \mathcal{B}_I über eine Galilei-Transformation zusammenhängt.

Um zu zeigen, dass das zweite newtonsche Gesetz (I.14) ebenfalls in beiden Bezugssystemen \mathcal{B}_I und \mathcal{B}' gleichzeitig gilt, soll noch das Transformationsgesetz für Kräfte präzisiert werden. Dafür soll man die Form der Kraft genauer berücksichtigen.

- Kräfte der Form $\vec{F} = F \vec{e}_F$ relativ zu \mathcal{B}_I , mit konstantem Betrag F und einer festen Richtung \vec{e}_F im Raum, transformieren sich unter Drehungen genau wie Ortsvektoren — indem \vec{e}_F für einen Inertialbeobachter in \mathcal{B}' gedreht aussieht —, während sie unter Translationen oder Galilei-Boosts invariant bleiben. Somit gilt die Gleichung $m\ddot{\vec{x}}(t) = \vec{F}$ sowohl in \mathcal{B}_I als (mit „gestrichenen“, transformierten Vektoren) in \mathcal{B}' .

Ein Beispiel für eine solche Kraft ist die Schwerkraft $m\vec{g}$ in einem homogenen Schwerfeld.

- Realistische ortsabhängige Kräfte hängen nicht vom Ortsvektor des Körpers ab, auf den sie ausgeübt werden, sondern nur von seinem Abstandsvektor von einem anderen Körper. Dies ist z.B. der Fall der newtonschen Gravitationskraft zwischen zwei Massenpunkten. Dieser Abstandsvektor, und somit die Kraft, bleibt unverändert unter Translationen oder Galilei-Boosts, während es sich unter Drehungen genau wie ein Ortsvektor bzw. eine Beschleunigung verhält. Wieder gilt $m\ddot{\vec{x}}(t) = \vec{F}$ sowohl in \mathcal{B}_I als in \mathcal{B}' .
- Bei geschwindigkeitsabhängigen Kräften ist eine zusätzliche Fallunterscheidung nötig.
 - Bei Reibungskräften, wie z.B. bei der Stokesschen Reibung (I.15c), ist die relevante Geschwindigkeit diejenige relativ zum Ruhesystem eines Mediums — z.B. der Luft oder Flüssigkeit, durch welche der Körper sich bewegt. Diese Relativgeschwindigkeit ändert sich nicht, bis auf Änderungen der beobachteten Richtung, unter den Transformationen der § I.3.1–I.3.3, so dass die Gleichung $m\ddot{\vec{x}}(t) = \vec{F}$ die gleiche Form in \mathcal{B}_I und in \mathcal{B}' annimmt.

⁽¹²⁾Diese Gruppe wird manchmal mit $\text{Gal}(3)$ bezeichnet.

^(h)N. H. ABEL, 1802–1829

- Der Fall der Lorentz-Kraft ist mehr problematisch. In Abwesenheit eines elektrischen Feldes \vec{E} lautet sie $\vec{F} = q\vec{v} \times \vec{B}$, mit \vec{B} einem statischen magnetischen Feld. Dabei ist \vec{v} die Geschwindigkeit der bewegten Punktladung relativ zum Bezugssystem, in dem $\vec{E} = \vec{0}$ und \vec{B} stationär ist. In Bezugssystemen, die sich relativ zu diesem Bezugssystem mit einer Geschwindigkeit \vec{u} bewegen, nehmen \vec{E} und \vec{B} unterschiedliche Formen an, wie man im Rahmen der Elektrodynamik nachprüfen kann. Die Felder ändern sich dabei so, dass die Lorentz-Kraft, einschließlich des Terms mit \vec{E} , (in Betrag) invariant bleibt. Dies gilt zumindest annähernd, solange $|\vec{u}|$ „klein“ ist, und zwar gegenüber der Vakuumlichtgeschwindigkeit. Für entsprechende Galilei-Boosts erhält die Gleichung $m\ddot{\vec{x}}(t) = \vec{F}$ die gleiche Form im gleichförmig bewegten Bezugssystem als in einem Inertialsystem.

Eigentlich transformieren sich die Gleichungen der Elektrodynamik unter Galilei-Boosts nicht, wie sie sollten, damit Galilei-Transformationen die richtigen Transformationen zwischen Inertialsystemen darstellen.

Abgesehen vom oben erwähnten Problem mit der Lorentz-Kraft nimmt die mathematische Formulierung (I.14) des zweiten newtonschen Gesetzes die gleiche Form in allen Bezugssystemen an, die sich aus einem Inertialsystem über eine Galilei-Transformation ableiten lassen. Deshalb sind diese Bezugssysteme auch Inertialsysteme.

1.4 Beschleunigte Bezugssysteme. Scheinkräfte

Im letzten Abschnitt haben wir gesehen, dass ein Bezugssystem \mathcal{B}' , das sich relativ zu einem Inertialsystem \mathcal{B}_I gleichförmig und geradlinig bewegt, oder dessen Koordinaten konstant verschoben oder gedreht gegenüber solchen von \mathcal{B}_I sind, ebenfalls ein Inertialsystem ist.

In diesem Abschnitt wird wieder ein Inertialsystem \mathcal{B}_I betrachtet, in denen die newtonschen Gesetze des Abschn. I.2 gelten. Jetzt bezeichnet \mathcal{B}' ein zweites Bezugssystem in *beschleunigter* Bewegung gegenüber \mathcal{B}_I . Wie wir sehen werden ist ein solches Bezugssystem kein Inertialsystem.

Ohne Beschränkung der Allgemeinheit wird angenommen, dass die Beobachter in \mathcal{B}_I und \mathcal{B}' die gleiche Zeit $t' = t$ verwenden. Ortsvektoren bezüglich eines in \mathcal{B}_I bzw. \mathcal{B}' festen Ursprungspunkts werden mit \vec{r} bzw. \vec{r}' bezeichnet. Wiederum steht $\vec{x}(t)$ bzw. $\vec{x}'(t) = \vec{x}'(t)$ für die Bahnkurve eines Massenpunkts relativ zu \mathcal{B}_I bzw. \mathcal{B}' .

Sei Σ ein physikalisches System, modelliert durch einen Massenpunkt mit (träger) Masse m . Laut dem zweiten newtonschen Gesetz (I.14) gilt im Inertialsystem \mathcal{B}_I

$$m \frac{d^2 \vec{x}(t)}{dt^2} = \vec{F}(t), \quad (\text{I.29})$$

wobei $\vec{F}(t)$ die (möglicherweise zeitabhängige) Resultierende der physikalischen Kräfte auf Σ bezeichnet. Im Folgenden wird die in \mathcal{B}' beobachtete Beschleunigung $d^2 \vec{x}'(t)/dt'^2 = d^2 \vec{x}'(t)/dt^2$ für unterschiedliche Bewegungen von \mathcal{B}' (relativ zu \mathcal{B}_I) berechnet, und zwar erstens für den Fall eines linear beschleunigten Bezugssystems (§ I.4.1), dann für ein rotierendes Bezugssystem (§ I.4.2). In beiden Fällen weicht $d^2 \vec{x}'(t)/dt^2$ von der im Inertialsystem beobachteten Beschleunigung $d^2 \vec{x}(t)/dt^2$ ab: die zusätzlichen Terme, multipliziert mit m , lassen sich aus der Sicht eines Beobachters in \mathcal{B}' als auf Σ wirkende *Scheinkräfte* interpretieren.

1.4.1 Linear beschleunigte Bezugssysteme

Betrachten wir zunächst den Fall eines Bezugssystems \mathcal{B}' , das sich relativ zum Inertialsystem \mathcal{B}_I mit einer zeitabhängigen Geschwindigkeit $\vec{u}(t)$ bewegt, wobei die Richtung von $\vec{u}(t)$ konstant bleibt, so dass die Bewegung geradlinig ist. Der Einfachheit halber wird angenommen, dass die an \mathcal{B}_I und

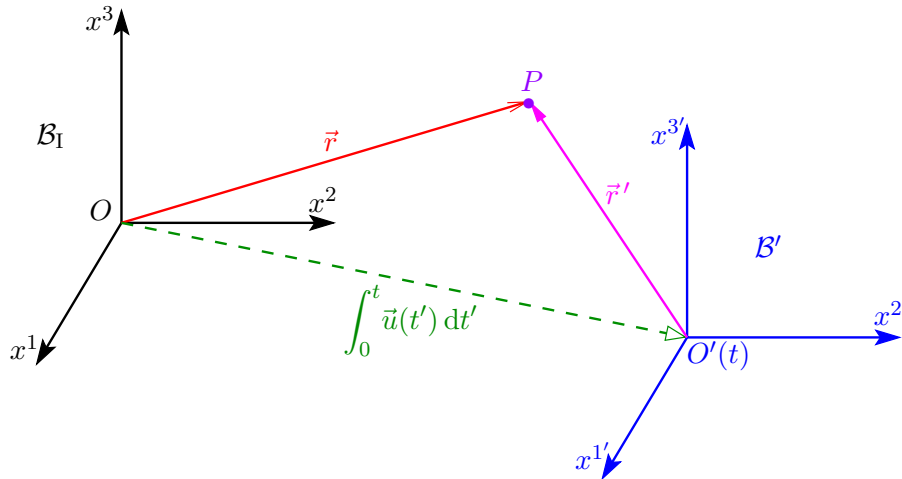


Abbildung I.5

\mathcal{B}' gebundenen Beobachter Koordinatensysteme verwenden, deren Nullpunkte O , O' zur Zeit $t = 0$ übereinstimmen (Abb. I.5). Dann gilt

$$\overrightarrow{OO'(t)} = \int_0^t \vec{u}(t') dt'.$$

Dementsprechend hängen die Ortsvektoren eines bestimmten geometrischen Punktes P bezüglich der beiden Bezugssysteme über

$$\vec{r}'(t) = \vec{r} - \int_0^t \vec{u}(t') dt' \quad (\text{I.30})$$

zusammen.

Zum Beispiel gilt im Fall eines konstant beschleunigten \mathcal{B}' : $\vec{u}(t) = \vec{a}t + \vec{u}(0)$, und somit

$$\vec{r}' = \vec{r} - \left(\frac{1}{2} \vec{a}t^2 + \vec{u}(0)t \right).$$

Die Beziehung (I.30) gilt auch zwischen den in \mathcal{B}_I und \mathcal{B}' gemessenen Zeit-Ort-Funktionen $\vec{x}(t)$ und $\vec{x}'(t)$ eines Massenpunkts. Eine erste Ableitung nach der Zeit gibt dann

$$\frac{d\vec{x}'(t)}{dt} = \frac{d\vec{x}(t)}{dt} - \vec{u}(t), \quad (\text{I.31})$$

entsprechend einem Additionstheorem für die Geschwindigkeiten. Dann ergibt eine zweite Ableitung

$$\frac{d^2\vec{x}'(t)}{dt^2} = \frac{d^2\vec{x}(t)}{dt^2} - \frac{d\vec{u}(t)}{dt}, \quad (\text{I.32})$$

d.h. die in beiden Bezugssystemen gemessenen Beschleunigungen weichen von einander ab.

In Abwesenheit von physikalischen Kräften, d.h. wenn $\vec{F} = \vec{0}$, gilt im Inertialsystem \mathcal{B}_I laut den ersten zwei newtonschen Gesetzen $d^2\vec{x}(t)/dt^2 = \vec{0}$. Dagegen wird $d^2\vec{x}'(t)/dt^2$ im Allgemeinen gemäß Gl. (I.32) ungleich Null sein: Aus der Sicht eines Beobachters im Bezugssystem \mathcal{B}' hat ein bezüglich \mathcal{B}_I ruhender, d.h. eigentlich kräftefreier, Massenpunkt die Beschleunigung $\ddot{\vec{x}}'(t) = -\ddot{\vec{u}}(t)$. Dementsprechend ist \mathcal{B}' kein Inertialsystem.

Um die mathematische Formulierung (I.14b) des zweiten newtonschen Gesetzes auch auf den Fall linear beschleunigter Bezugssystem zu verallgemeinern, soll der Beobachter in \mathcal{B}' eine *Scheinkraft*

$$\vec{F}_{\text{Schein}} \equiv -m\ddot{\vec{u}}(t) \quad (\text{I.33})$$

einführen. Dann gilt in \mathcal{B}' (in Abwesenheit von physikalischen Kräften)

$$m \frac{d^2\vec{x}'(t)}{dt^2} = \vec{F}_{\text{Schein}}.$$

Wenn es dazu physikalische Kräfte \vec{F} gibt, dann addieren sie sich vektoriell zu der Scheinkraft:

$$m \frac{d^2 \vec{x}'(t)}{dt^2} = \vec{F} + \vec{F}_{\text{Schein}}, \quad (\text{I.34})$$

was unter Berücksichtigung des in \mathcal{B}_I ausgedrückten zweiten newtonschen Gesetzes genau äquivalent zur Gl. (I.32) ist.

I.4.2 Rotierende Bezugssysteme

Jetzt wird angenommen, dass das Bezugssystem \mathcal{B}' gegenüber dem Inertialsystem \mathcal{B}_I rotiert. Dabei können sowohl die (momentane) Achse der Rotation als auch die Rotationsgeschwindigkeit von \mathcal{B}' bezüglich \mathcal{B}_I von der Zeit abhängen. Der Einfachheit halber wird angenommen, dass die Koordinatensysteme in \mathcal{B}' und \mathcal{B}_I den gleichen Ursprungspunkt haben.

Somit lassen sich Vektoren in \mathcal{B}' aus denen bezüglich \mathcal{B}_I durch eine zeitabhängige Drehmatrix $\mathcal{R}_{\mathcal{B}'/\mathcal{B}_I}(t) \in \text{SO}(3)$ erhalten [vgl. Gl. (I.28a)]; z.B. gilt für den Ortsvektor eines Punkts

$$\vec{r}' = \mathcal{R}_{\mathcal{B}'/\mathcal{B}_I}(t) \vec{r}. \quad (\text{I.35a})$$

Im Folgenden wird es günstiger sein, die Bahnkurve eines Massenpunkts bezüglich \mathcal{B}_I durch jene relativ zu \mathcal{B}' auszudrücken. Da die Drehmatrix $\mathcal{R}_{\mathcal{B}'/\mathcal{B}_I}(t)$ eine orthogonale Matrix ist, ist sie automatisch invertierbar, d.h. man darf problemlos

$$\vec{r} = [\mathcal{R}_{\mathcal{B}'/\mathcal{B}_I}(t)]^{-1} \vec{r}' \equiv \mathcal{R}(t) \vec{r}' \quad (\text{I.35b})$$

schreiben, wobei die kürzere Notation $\mathcal{R}(t)$ für die inverse Drehmatrix eingeführt wurde.⁽¹³⁾

Diese Beziehung gilt nicht nur für den Ortsvektor, sondern für jeden Vektor. Somit gilt für die in \mathcal{B}_I und \mathcal{B}' beobachteten Bahnkurven

$$\vec{x}(t) = \mathcal{R}(t) \vec{x}'(t) \quad (\text{I.36a})$$

und gleichfalls für die Kräfte

$$\vec{F}(t) = \mathcal{R}(t) \vec{F}'(t), \quad (\text{I.36b})$$

vgl. die Diskussion in § I.3.5.

⁽¹³⁾Es wird empfohlen, sich die etwa abstrakten Berechnungen dieses Paragraphen anhand eines Beispiels klarer zu machen. Z.B. kann man annehmen, dass \mathcal{B}' mit konstanter Rotationsgeschwindigkeit ω_0 um die x^3 -Achse des Koordinatensystems von \mathcal{B}_I rotiert. Die Beziehung zwischen den Basisvektoren der beiden Koordinatensysteme lautet dann

$$\begin{pmatrix} \vec{e}_{1'} \\ \vec{e}_{2'} \\ \vec{e}_{3'} \end{pmatrix} = \begin{pmatrix} \cos(\omega_0 t) & \sin(\omega_0 t) & 0 \\ -\sin(\omega_0 t) & \cos(\omega_0 t) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \vec{e}_1 \\ \vec{e}_2 \\ \vec{e}_3 \end{pmatrix}$$

oder äquivalent, nach Transposition

$$(\vec{e}_{1'} \ \vec{e}_{2'} \ \vec{e}_{3'}) = (\vec{e}_1 \ \vec{e}_2 \ \vec{e}_3) \begin{pmatrix} \cos(\omega_0 t) & -\sin(\omega_0 t) & 0 \\ \sin(\omega_0 t) & \cos(\omega_0 t) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Für die Koordinaten eines Vektors in \mathcal{B}' und \mathcal{B}_I gilt [vgl. Gl. (I.35a)]

$$\begin{pmatrix} x^{1'} \\ x^{2'} \\ x^{3'} \end{pmatrix} = \begin{pmatrix} \cos(\omega_0 t) & \sin(\omega_0 t) & 0 \\ -\sin(\omega_0 t) & \cos(\omega_0 t) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} \equiv \mathcal{R}_{\mathcal{B}'/\mathcal{B}_I}(t) \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix}.$$

Daraus folgt

$$\begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} \cos(\omega_0 t) & -\sin(\omega_0 t) & 0 \\ \sin(\omega_0 t) & \cos(\omega_0 t) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^{1'} \\ x^{2'} \\ x^{3'} \end{pmatrix} \equiv \mathcal{R}(t) \begin{pmatrix} x^{1'} \\ x^{2'} \\ x^{3'} \end{pmatrix}$$

wobei $\mathcal{R}(t)$ die in Gl. (I.35b) eingeführte Drehmatrix ist.

I.4.2a Bewegungsgleichung in einem rotierenden Bezugssystem

Leitet man zunächst die Beziehung (I.36a) nach der Zeit ab, so ergibt sich unter Verwendung der Produktregel

$$\frac{d\vec{x}(t)}{dt} = \frac{d\mathcal{R}(t)}{dt} \vec{x}'(t) + \mathcal{R}(t) \frac{d\vec{x}'(t)}{dt}. \quad (\text{I.37})$$

Dabei ist $d\mathcal{R}(t)/dt$ die Zeitableitung

$$\frac{d\mathcal{R}(t)}{dt} = \lim_{\delta t \rightarrow 0} \frac{\mathcal{R}(t+\delta t) - \mathcal{R}(t)}{\delta t} = \lim_{\delta t \rightarrow 0} \frac{1}{\delta t} [\mathcal{R}(t+\delta t)\mathcal{R}(t)^{-1} - \mathbb{1}_3] \mathcal{R}(t) \quad (\text{I.38})$$

mit der 3×3 -Einheitsmatrix $\mathbb{1}_3$, wobei das Produkt zweier Drehmatrizen $\mathcal{R}(t+\delta t)\mathcal{R}(t)^{-1} \equiv \mathcal{D}$ selbst eine Drehmatrix ist, die eigentlich von t und δt abhängt. Im Limes $\delta t = 0$ gilt $\mathcal{D} = \mathbb{1}_3$. Somit lässt sich \mathcal{D} bei „kleinem“ δt als

$$\mathcal{D} = \mathbb{1}_3 + \Omega(t)\delta t + \mathcal{O}((\delta t)^2) \quad (\text{I.39})$$

mit einer 3×3 -Matrix $\Omega(t)$ schreiben.

Genauer muss der Betrag des Produkts aus δt mit irgendeinem der Matrixelemente von $\Omega(t)$ klein gegenüber 1 sein. Bei Gl. (I.39) handelt es sich eigentlich um eine Taylor-Entwicklung, wobei $\Omega(t)$ die Ableitung von \mathcal{D} nach δt bei $\delta t = 0$ ist.

Diese Näherung kann in Gl. (I.38) eingesetzt werden, woraus sich

$$\frac{d\mathcal{R}(t)}{dt} = \Omega(t)\mathcal{R}(t) \quad (\text{I.40})$$

ergibt.

Da \mathcal{D} eine Drehmatrix ist, gilt $\mathcal{D}\mathcal{D}^\top = \mathcal{D}^\top\mathcal{D} = \mathbb{1}_3$. Die Transposition der Gl. (I.39) gibt

$$\mathcal{D}^\top = \mathbb{1}_3 + \Omega(t)^\top \delta t + \mathcal{O}((\delta t)^2),$$

woraus $\mathcal{D}\mathcal{D}^\top = \mathbb{1}_3 + [\Omega(t) + \Omega(t)^\top] \delta t + \mathcal{O}((\delta t)^2)$ folgt. Dies ist gleich $\mathbb{1}_3$ bis auf Terme der Ordnung $(\delta t)^2$ oder höher vorausgesetzt

$$\Omega(t)^\top = -\Omega(t),$$

d.h. wenn $\Omega(t)$ eine antisymmetrische (d.h. schiefsymmetrische) Matrix ist. Um diese vollständig zu charakterisieren, sind nur 3 reellen Zahlen nötig, z.B. die drei Einträge $[\Omega(t)]_{12}$, $[\Omega(t)]_{13}$, $[\Omega(t)]_{23}$. Definiert man 3 Zahlen $\omega^k(t)$, $k = \{1, 2, 3\}$ über die Beziehung⁽¹⁴⁾

$$\omega^k(t) \equiv -\frac{1}{2} \epsilon^{ijk} [\Omega(t)]_{ij}, \quad (\text{I.41})$$

so lässt sich die Matrix $\Omega(t)$ als

$$\Omega(t) = \begin{pmatrix} 0 & -\omega^3(t) & \omega^2(t) \\ \omega^3(t) & 0 & -\omega^1(t) \\ -\omega^2(t) & \omega^1(t) & 0 \end{pmatrix} \quad (\text{I.42})$$

schreiben.⁽¹⁵⁾ Mit dieser Parametrisierung gilt für jeden Vektor \vec{a} mit Komponenten (a^1, a^2, a^3)

$$\Omega(t)\vec{a} = \begin{pmatrix} 0 & -\omega^3(t) & \omega^2(t) \\ \omega^3(t) & 0 & -\omega^1(t) \\ -\omega^2(t) & \omega^1(t) & 0 \end{pmatrix} \begin{pmatrix} a^1 \\ a^2 \\ a^3 \end{pmatrix} = \begin{pmatrix} \omega^2(t)a^3 - \omega^3(t)a^2 \\ \omega^3(t)a^1 - \omega^1(t)a^3 \\ \omega^1(t)a^2 - \omega^2(t)a^1 \end{pmatrix},$$

⁽¹⁴⁾Die umgekehrte Beziehung zur Gl. (I.41) ist $[\Omega(t)]_{ij} = -\epsilon_{ijk}\omega^k(t)$.

⁽¹⁵⁾Für das in Fußnote (13) eingeführte Beispiel gilt

$$\mathcal{D} \equiv \mathcal{R}(t+\delta t)\mathcal{R}(t)^{-1} = \begin{pmatrix} \cos(\omega_0\delta t) & -\sin(\omega_0\delta t) & 0 \\ \sin(\omega_0\delta t) & \cos(\omega_0\delta t) & 0 \\ 0 & 0 & 1 \end{pmatrix} \underset{\delta t \rightarrow 0}{\sim} \begin{pmatrix} 1 & -\omega_0\delta t & 0 \\ \omega_0\delta t & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Somit ist in diesem Fall [vgl. Gl. (I.39)]

$$\Omega(t) = \begin{pmatrix} 0 & -\omega_0 & 0 \\ \omega_0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

und der über die Parametrisierung (I.41) assoziierte Vektor ist $\vec{\omega}(t) = \omega_0 \vec{e}_3$: der Betrag $|\vec{\omega}(t)|$ ist gleich demjenigen der Rotationsgeschwindigkeit, die Richtung \vec{e}_3 ist diejenige der Rotationsachse.

d.h.

$$\Omega(t)\vec{a} = \vec{\omega}(t) \times \vec{a}, \quad (\text{I.43})$$

wobei $\vec{\omega}(t)$ den Vektor mit Komponenten $(\omega^1(t), \omega^2(t), \omega^3(t))$ bezeichnet. Die physikalische Bedeutung dieses Vektors wird im § I.4.2 b diskutiert. Wie wir dort sehen wird, hat $\vec{\omega}(t)$ eine durchschaubare Interpretation, wenn er in \mathcal{B}_I „lebt“, d.h. wenn die $\omega^k(t)$ seine Komponenten relativ zu einem Koordinatensystem im Inertialsystem sind.

Ersetzt man die Ableitung $d\mathcal{R}(t)/dt$ in Gl. (I.37) mit Hilfe von Gl. (I.40), so kommt

$$\begin{aligned} \frac{d\vec{x}(t)}{dt} &= \Omega(t)\mathcal{R}(t)\vec{x}'(t) + \mathcal{R}(t)\frac{d\vec{x}'(t)}{dt} \\ &= \vec{\omega}(t) \times [\mathcal{R}(t)\vec{x}'(t)] + \mathcal{R}(t)\frac{d\vec{x}'(t)}{dt}, \end{aligned} \quad (\text{I.44})$$

wobei die zweite Gleichung aus Beziehung (I.43) folgt.

Sei jetzt $\vec{\omega}'(t) \equiv \mathcal{R}(t)^{-1}\vec{\omega}(t)$ der Vektor $\vec{\omega}(t)$ aus der Sicht eines Beobachters im rotierenden Bezugssystem \mathcal{B}' . Unter Verwendung der Identität

$$(\mathcal{R}\vec{a}) \times (\mathcal{R}\vec{b}) = \mathcal{R}(\vec{a} \times \vec{b}), \quad (\text{I.45})$$

die für jede Drehmatrix \mathcal{R} und jedes Paar von Vektoren \vec{a}, \vec{b} gilt, kann man jetzt

$$\frac{d\vec{x}(t)}{dt} = \mathcal{R}(t) \left[\vec{\omega}'(t) \times \vec{x}'(t) + \frac{d\vec{x}'(t)}{dt} \right] \quad (\text{I.46})$$

schreiben.

Beweis der Identität (I.45): es seien $\{a^i\}$ bzw. $\{b^i\}$ die kartesischen Koordinaten von \vec{a} bzw. \vec{b} und $\mathcal{R}^{i'}_j$ die Matrixelemente von \mathcal{R} . Dann gilt für $i' = 1, 2, 3$

$$[(\mathcal{R}\vec{a}) \times (\mathcal{R}\vec{b})]^{i'} = \epsilon^{i'}_{j'k'} \mathcal{R}^{j'}_j a^j \mathcal{R}^{k'}_k b^k.$$

Andererseits ist die Determinante der Drehmatrix \mathcal{R} durch

$$\det \mathcal{R} = \epsilon_{\ell'j'k'} \mathcal{R}^{\ell'}_1 \mathcal{R}^{j'}_2 \mathcal{R}^{k'}_3$$

gegeben, wobei eigentlich $\det \mathcal{R} = 1$. Äquivalent gilt nach einer Permutation $1 \rightarrow \ell, 2 \rightarrow j, 3 \rightarrow k$ der Matrixspalten

$$\epsilon_{\ell j k} \det \mathcal{R} = \epsilon_{\ell'j'k'} \mathcal{R}^{\ell'}_{\ell} \mathcal{R}^{j'}_j \mathcal{R}^{k'}_k,$$

wobei $\epsilon_{\ell j k}$ das Signum der Permutation ist. Multipliziert man diese Gleichung mit $\mathcal{R}^{i'}_{\ell}$ und summiert man über ℓ , so kommt $\epsilon_{\ell j k} \mathcal{R}^{i'}_{\ell} = \epsilon_{\ell'j'k'} \mathcal{R}^{i'}_{\ell'} \mathcal{R}^{\ell'}_{\ell} \mathcal{R}^{j'}_j \mathcal{R}^{k'}_k$. Dabei ist $\mathcal{R}^{\ell'}_{\ell}$ das ℓ' -Element der Matrix \mathcal{R} , d.h. auch das $\ell\ell'$ -Element der transponierten Matrix \mathcal{R}^T , so dass $\mathcal{R}^{i'}_{\ell'} \mathcal{R}^{\ell'}_{\ell}$, mit Summe über ℓ , das $i'\ell'$ -Element des Produkts $\mathcal{R}\mathcal{R}^T = \mathbf{1}_3$ ist. Somit gilt

$$\epsilon_{\ell j k} \mathcal{R}^{i'}_{\ell} = \epsilon_{\ell'j'k'} \delta^{i'\ell'} \mathcal{R}^{j'}_j \mathcal{R}^{k'}_k = \epsilon_{\ell'j'k'} \mathcal{R}^{j'}_j \mathcal{R}^{k'}_k.$$

Schließlich findet man

$$[(\mathcal{R}\vec{a}) \times (\mathcal{R}\vec{b})]^{i'} = \epsilon_{i'j'k'} \mathcal{R}^{j'}_j a^j \mathcal{R}^{k'}_k b^k = \epsilon_{\ell j k} \mathcal{R}^{i'}_{\ell} a^j b^k = [\mathcal{R}(\vec{a} \times \vec{b})]^{i'},$$

entsprechend dem gesuchten Ergebnis. \square

Das wiederholte Ableiten von Gl. (I.46) nach der Zeit unter Berücksichtigung der Produktregel liefert dann

$$\frac{d^2\vec{x}(t)}{dt^2} = \frac{d\mathcal{R}(t)}{dt} \left[\vec{\omega}'(t) \times \vec{x}'(t) + \frac{d\vec{x}'(t)}{dt} \right] + \mathcal{R}(t) \left[\frac{d\vec{\omega}'(t)}{dt} \times \vec{x}'(t) + \vec{\omega}'(t) \times \frac{d\vec{x}'(t)}{dt} + \frac{d^2\vec{x}'(t)}{dt^2} \right].$$

Der erste Term auf der rechten Seite lässt sich mithilfe der Gl. (I.40), (I.43) und (I.45) transformieren:

$$\begin{aligned} \frac{d\mathcal{R}(t)}{dt} \left[\vec{\omega}'(t) \times \vec{x}'(t) + \frac{d\vec{x}'(t)}{dt} \right] &= \Omega(t)\mathcal{R}(t) \left[\vec{\omega}'(t) \times \vec{x}'(t) + \frac{d\vec{x}'(t)}{dt} \right] \\ &= \vec{\omega}(t) \times \left\{ \mathcal{R}(t) \left[\vec{\omega}'(t) \times \vec{x}'(t) + \frac{d\vec{x}'(t)}{dt} \right] \right\} \\ &= [\mathcal{R}(t)\vec{\omega}'(t)] \times \left\{ \mathcal{R}(t) \left[\vec{\omega}'(t) \times \vec{x}'(t) + \frac{d\vec{x}'(t)}{dt} \right] \right\} \\ &= \mathcal{R}(t) \left\{ \vec{\omega}'(t) \times \left[\vec{\omega}'(t) \times \vec{x}'(t) + \frac{d\vec{x}'(t)}{dt} \right] \right\}. \end{aligned}$$

Somit ergibt sich insgesamt

$$\frac{d^2\vec{x}(t)}{dt^2} = \mathcal{R}(t) \left\{ \frac{d^2\vec{x}'(t)}{dt^2} + \vec{\omega}'(t) \times [\vec{\omega}'(t) \times \vec{x}'(t)] + 2\vec{\omega}'(t) \times \frac{d\vec{x}'(t)}{dt} + \frac{d\vec{\omega}'(t)}{dt} \times \vec{x}'(t) \right\}. \quad (\text{I.47})$$

Unter Berücksichtigung des zweiten newtonschen Gesetzes (I.29) (ausgedrückt im Inertialsystem \mathcal{B}_I) lässt sich der Term auf der linken Seite als \vec{F}/m umschreiben. Dabei kann man \vec{F} mithilfe der Gl. (I.36b) ersetzen. Nach Anwendung von \mathcal{R}^{-1} auf die daraus folgende Gleichung erhält man schließlich

$$m \frac{d^2\vec{x}'(t)}{dt^2} = \vec{F}' - m\vec{\omega}'(t) \times [\vec{\omega}'(t) \times \vec{x}'(t)] - 2m\vec{\omega}'(t) \times \frac{d\vec{x}'(t)}{dt} - m \frac{d\vec{\omega}'(t)}{dt} \times \vec{x}'(t). \quad (\text{I.48})$$

Die drei letzten Terme auf der rechten Seite sind wieder *Scheinkräfte*, die zur im Bezugssystem \mathcal{B}' gemessenen Beschleunigung beitragen.

Bemerkung: Alle drei neuen Scheinkräfte in Gl. (I.48) sind proportional zur trägen Masse m , so wie die in § I.4.1 gefundene Scheinkraft in einem linear beschleunigten System. Aus diesem Grund werden Scheinkräfte auch manchmal *Trägheitskräfte* genannt.

Wegen der nach heutigem (Herbst 2022) Wissenstand gültigen Proportionalität zwischen schwerer und träger Masse ist auch die Schwerkraft proportional zu m . Daher kann die Schwerkraft zumindest mathematisch als eine Scheinkraft betrachtet werden, die sich über eine passende Wahl von Bezugssystem annullieren lässt — was die grobe Grundidee von der Allgemeinen Relativitätstheorie darstellt.

I.4.2b Winkelgeschwindigkeit

Bevor wir die drei Scheinkräfte in Gl. (I.48) ausführlicher diskutieren, sollte die Bedeutung des darin auftretenden Vektors $\vec{\omega}'(t)$ bzw. von $\vec{\omega}(t)$ festgestellt werden. Zu diesem Zweck können wir Gl. (I.44) für den Fall eines bezüglich \mathcal{B}' ruhenden Massenpunkts betrachten. Für $d\vec{x}'(t)/dt = \vec{0}$ gilt nämlich

$$\frac{d\vec{x}(t)}{dt} = \vec{\omega}(t) \times [\mathcal{R}(t)\vec{x}'(t)] = \vec{\omega}(t) \times \vec{x}(t). \quad (\text{I.49})$$

Das heißt, die (Rate der) Änderung des in \mathcal{B}_I gemessenen Ortsvektors $\vec{x}(t)$ ist senkrecht auf $\vec{x}(t)$ und auf den Vektor $\vec{\omega}(t)$. Diese Gleichung beschreibt eine (instantane) Rotationsbewegung bezüglich \mathcal{B}_I um eine Achse, die durch den Ursprungspunkt des Koordinatensystems geht und entlang der Richtung von $\vec{\omega}(t)$ liegt. Dazu kann man anhand eines Beispiels⁽¹⁶⁾ prüfen, dass die zugehörige Rotationsgeschwindigkeit — d.h. die Rate der Änderung des Winkels von $\vec{x}(t)$ bezüglich einer festen Richtung — genau durch den Betrag $|\vec{\omega}(t)|$ gegeben wird. Entsprechend diesen Eigenschaften heißt der Vektor $\vec{\omega}(t)$ *Winkelgeschwindigkeit*.

Bemerkung: Genauer ist die Winkelgeschwindigkeit $\vec{\omega}(t)$ ein sog. *Pseudovektor* oder *axialer Vektor*, dessen Richtung von der Konvention für die Orientierung des Raums abhängt.

⁽¹⁶⁾... wie jenes aus der Fußnote (13)! Vgl. den entsprechenden Ausdruck von $\vec{\omega}(t)$ in Fußnote (15).

Sei $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ eine Orthonormalbasis des dreidimensionalen euklidischen Raums. In der üblichen Konvention wird die Rechte-Hand-Regel verwendet: wenn der Daumen bzw. der Zeigefinger der rechten Hand entlang \vec{e}_1 bzw. \vec{e}_2 zeigt, dann ist $\vec{e}_3 = \vec{e}_1 \times \vec{e}_2$ entlang des Mittelfingers, und $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ ist ein rechtshändiges System. Allgemeiner ist $(\vec{c}_1, \vec{c}_2, \vec{c}_1 \times \vec{c}_2)$ für jedes Paar von nicht-kollinearen Vektoren \vec{c}_1, \vec{c}_2 ebenfalls ein rechtshändiges System.

Mathematisch darf man auch problemlos die Linke-Hand-Regel benutzen, d.h. entscheiden, dass $\vec{c}_1 \times \vec{c}_2$ in die entgegengesetzte Richtung zeigen soll. Nun, in Gl. (I.49) hängen weder der Ortsvektor $\vec{x}(t)$ noch die Geschwindigkeit $d\vec{x}(t)/dt$ — die in die Richtung der Verschiebung des Ortsvektors zwischen den Zeitpunkten t und $t + \delta t$ zeigt, wenn δt klein genug ist — von der Wahl der Konvention für das Kreuzprodukt. Wenn auch $\vec{\omega}(t)$ nicht von dieser Konvention abhängt, wäre Gl. (I.49) problematisch: die linke Seite wäre unabhängig von der Konvention, die rechte Seite dagegen abhängig davon. Somit muss die Richtung von $\vec{\omega}(t)$ mit der Konvention für das Kreuzprodukt zusammenhängen.

Die „normalen“ Vektoren, wie z.B. der Ortsvektor, die nicht von der Konvention für die Orientierung des Raums abhängen, werden *polare Vektoren* genannt. Die Leserin kann sich überzeugen, dass polare Vektoren unter der Raumspiegelung $\vec{r} \rightarrow -\vec{r}$ (die ein rechtshändiges System in ein linkshändiges transformiert!) ihre Richtung umkehren, während axiale Vektoren ihre Richtung beibehalten.

I.4.2c Zentrifugalkraft

Der zweite Term auf der rechten Seite von Gl. (I.48) ist die *Zentrifugalkraft*

$$\vec{F}_{\text{zentrifugal}} \equiv -m\vec{\omega}'(t) \times [\vec{\omega}'(t) \times \vec{x}'(t)]. \quad (\text{I.50})$$

Zerlegt man den Ortsvektor $\vec{x}'(t)$ in die Summe aus einem Vektor $\vec{x}'_{\parallel}(t)$ parallel zur Winkelgeschwindigkeit $\vec{\omega}'(t)$ und einem Vektor $\vec{x}'_{\perp}(t)$ senkrecht darauf, so trägt nur $\vec{x}'_{\perp}(t)$ zum geklammerten Kreuzprodukt bei, und man findet

$$\vec{F}_{\text{zentrifugal}} = m|\vec{\omega}'(t)|^2 \vec{x}'_{\perp}(t).$$

Somit ist die Zentrifugalkraft orthogonal zur Richtung der instantanen Achse der Rotationsbewegung, wie in Abb. I.6 dargestellt wird.

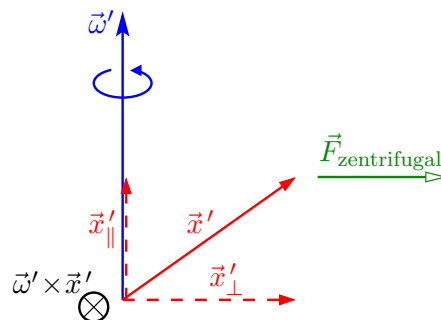


Abbildung I.6 – Zentrifugalkraft

Diese Scheinkraft wird auch durch Körper gespürt, die bezüglich des beschleunigten Bezugssystems \mathcal{B}' ruhen, falls sie nicht auf der Achse der Rotationsbewegung sitzen — d.h. für $\vec{x}'_{\perp}(t) \neq \vec{0}$.

I.4.2d Coriolis-Kraft

Die zweite in Gl. (I.48) auftretende Scheinkraft ist die *Coriolis⁽ⁱ⁾-Kraft*

$$\vec{F}_{\text{Coriolis}} \equiv -2m\vec{\omega}'(t) \times \vec{v}'(t), \quad (\text{I.51})$$

wobei $\vec{v}'(t) \equiv d\vec{x}'(t)/dt$ die Geschwindigkeit im nicht-Inertialsystem \mathcal{B}' ist. Diese Scheinkraft ist

⁽ⁱ⁾G. CORIOLIS, 1792–1843

offensichtlich senkrecht sowohl auf $\vec{v}'(t)$ als auf $\vec{\omega}'(t)$, und wirkt nicht auf Körper, die sich parallel zur instantanen Achse der Drehbewegung bewegen.

Ein an die Erde gebundenes Bezugssystem \mathcal{B}' ist wegen der Erdrotation in Drehbewegung relativ zu entfernten Galaxien. Somit erfahren Körper, die sich auf der Erde bewegen, eine Coriolis-Kraft. Diese dafür sorgt, dass horizontale Bewegungen, wie von z.B. Strömen in Ozeanen oder von Wolken, auf der Nord- bzw. Südhalbkugel nach rechts bzw. links abgelenkt werden.

I.4.2e Euler-Kraft

Der letzte Term in Gl. (I.48) ist die *Euler^(j)-Kraft*

$$\vec{F}_{\text{Euler}} \equiv -m \frac{d\vec{\omega}'(t)}{dt} \times \vec{x}'(t). \quad (\text{I.52})$$

Diese Scheinkraft tritt nur auf, wenn die Winkelgeschwindigkeit sich zeitlich ändert, entsprechend einer Änderung entweder der Rotationsgeschwindigkeit — d.h. des Betrags $|\vec{\omega}'(t)|$ — oder der Drehachse, oder selbstverständlich von beiden. Dabei spielt der Bewegungszustand — bewegt oder ruhend — des Körpers relativ zum beschleunigten Bezugssystem keine Rolle.

Falls die Rotationsachse konstant bleibt, so dass $\dot{\vec{\omega}}'(t)$ parallel zu $\vec{\omega}'(t)$ ist, findet man sofort, dass die Euler-Kraft tangential zur Bahnkurve des Körpers ist.

Literatur zum Kapitel I

- Fließbach, *Mechanik* [2] Teil I, Kap. 1–6.
- Greiner, *Klassische Mechanik I* [6] Kap. II und *Klassische Mechanik II* [7] Kap. I & II.
- Nolting, *Klassische Mechanik* [15] Kap. 2 & 3.
- Scheck, *Mechanik* [18] Kap. 1.

^(j)L. EULER, 1707–1783

Appendix zum Kapitel I

In diesem Anhang werden der Einfachheit halber ein paar Begriffe und Ergebnisse der Kinematik eines Massenpunkts gesammelt.

I.A Kinematik eines Massenpunkts

I.A.1 Grundbegriffe

Zur Beschreibung der Änderung des Ortsvektors \vec{r} eines Massenpunkts mit der Zeit t wird die *Zeit-Ort-Funktion* $\vec{x}(t)$ eingeführt. Die geometrische Kurve, entlang der sich der Massenpunkt bewegt, ist die *Bahnkurve* oder *Trajektorie*.

Die *Geschwindigkeit* — genauer: der Geschwindigkeitsvektor — ist die erste Ableitung der Zeit-Ort-Funktion nach der Zeit

$$\vec{v}(t) \equiv \frac{d\vec{x}(t)}{dt} \equiv \dot{\vec{x}}(t). \quad (\text{I.53})$$

Auf der rechten Seite der zweiten definierenden Gleichheit wird die Ableitung nach der Zeit mit einem Überpunkt bezeichnet („Newtonsche Notation“).

Wiederum ist die *Beschleunigung* die erste Ableitung des Geschwindigkeitsvektors nach der Zeit

$$\vec{a}(t) \equiv \frac{d\vec{v}(t)}{dt} \equiv \dot{\vec{v}}(t), \quad (\text{I.54})$$

d.h. die zweite Ableitung der Orts-Zeit-Funktion:

$$\vec{a}(t) = \frac{d^2\vec{x}(t)}{dt^2} \equiv \ddot{\vec{x}}(t). \quad (\text{I.55})$$

Geht man umgekehrt von der Angabe der Beschleunigung $\vec{a}(t)$ aus, liefern sukzessive Integrationen nach der Zeit den Geschwindigkeitsvektor $\vec{v}(t)$ und die Zeit-Ort-Funktion $\vec{x}(t)$. Dabei sollen Anfangsbedingungen (z.B. zu einem Zeitpunkt t_0) berücksichtigt werden:

$$\vec{v}(t) = \vec{v}(t_0) + \int_{t_0}^t \vec{a}(t') dt', \quad (\text{I.56a})$$

$$\vec{x}(t) = \vec{x}(t_0) + \int_{t_0}^t \vec{v}(t') dt' = \vec{x}(t_0) + \vec{v}(t_0)(t - t_0) + \int_{t_0}^t \left[\int_{t_0}^{t'} \vec{a}(t'') dt'' \right] dt'. \quad (\text{I.56b})$$

I.A.2 Koordinatensysteme

I.A.2a Kartesische Koordinaten

Sei gegeben eine Standard orthonormierte Basis $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ (oder kurz $\{\vec{e}_j\}_{j=1,2,3}$), die raumfest und zeitunabhängig ist. Die Zeit-Ort-Funktion wird durch deren Komponenten $\{x^j(t)\}$ auf dieser Basis festgelegt:

$$\vec{x}(t) = \sum_{j=1}^3 x^j(t) \vec{e}_j \equiv x^j(t) \vec{e}_j, \quad (\text{I.57})$$

wobei in der zweiten Gleichung die Einsteinsche Summenkonvention über den doppelt auftretenden Index j benutzt wird.

Laut Gl. (I.53) ist der Geschwindigkeitsvektor die Ableitung von $\vec{x}(t)$ nach der Zeit. Da die Basisvektoren $\{\vec{e}_j\}$ zeitunabhängig sind, ergibt sich

$$\vec{v}(t) = \sum_{j=1}^3 \dot{x}^j(t) \vec{e}_j. \quad (\text{I.58})$$

Somit ist die j -te Komponente $v^j(t)$ von $\vec{v}(t)$ gleich der ersten Ableitung nach der Zeit der j -ten Komponente $x^j(t)$ von $\vec{x}(t)$: $v^j(t) = \dot{x}^j(t)$.

Eine zweite Ableitung nach t gibt für die Beschleunigung

$$\vec{a}(t) = \sum_{j=1}^3 \ddot{x}^j(t) \vec{e}_j. \quad (\text{I.59})$$

Das heißt, die j -te Komponente von $\vec{a}(t)$ ist gleich der zweiten Ableitung der entsprechenden Komponente der Zeit-Ort-Funktion: $a^j(t) = \ddot{x}^j(t)$.

I.A.2b Ebene Polarkoordinaten

Seien kartesische Koordinaten (x^1, x^2) relativ zu einem Nullpunkt O in einer Ebene. Die Position des Endpunkts eines Ortsvektors $\vec{r} \neq \vec{0}$ (mit dem Ursprungspunkt in O) kann durch den Betrag $r \equiv |\vec{r}|$ und einen Winkel θ beschrieben werden, wobei θ konventionell im Gegenuhrzeigersinn relativ zur x^1 -Achse gemessen wird. Dann sind die kartesischen Koordinaten (x^1, x^2) des Endpunkts von \vec{r} gegeben durch

$$\begin{cases} x^1 = r \cos \theta \\ x^2 = r \sin \theta. \end{cases} \quad (\text{I.60})$$

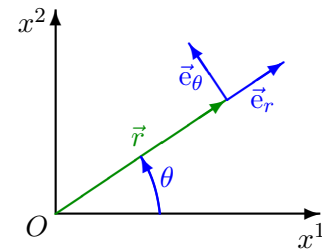


Abbildung I.7

Bei der Umkehrtransformation muss man vorsichtig sein. Einerseits gilt einfach

$$r = \sqrt{(x^1)^2 + (x^2)^2}, \quad (\text{I.61a})$$

d.h. $r \in \mathbb{R}_+$. Bei der Bestimmung des Polarwinkels θ hängt es davon ab, in welchem Quadranten der Ebene sich der Endpunkt von \vec{r} befindet. Möchte man einen Winkel im halboffenen Intervall $-\pi < \theta \leq \pi$ haben, so gilt

$$\theta = \arctan \frac{x^2}{x^1} + C \quad \text{für } x^1 \neq 0 \quad \text{mit } C = \begin{cases} 0 & \text{für } x^1 > 0 \\ -\pi & \text{für } x^1 < 0, x^2 < 0 \\ +\pi & \text{für } x^1 < 0, x^2 \geq 0 \end{cases} \quad (\text{I.61b})$$

Für $x^1 = 0$ ist $\theta = +\pi/2$ falls $x^2 > 0$, $\theta = -\pi/2$ falls $x^2 < 0$. Schließlich ist θ beliebig im Fall $r = 0$.

Gleichung (I.60) definiert eine Abbildung $(r, \theta) \mapsto (x^1, x^2)$, die jeden Punkt (r, θ) auf einen eindeutigen Punkt (x^1, x^2) abbildet. Offensichtlich werden aber mehrere Punkte der (r, θ) -Ebene auf denselben Punkt der (x^1, x^2) -Ebene abgebildet, wie z.B. (bei gegebenen r_0, θ_0) alle Punkte $(r_0, \theta_0 + 2n\pi)$ mit beliebiger $n \in \mathbb{Z}$.

Mit den Beschränkungen $r > 0$ und $\theta \in (-\pi, \pi]$ wird die Abbildung $(r, \theta) \mapsto (x^1, x^2)$ eineindeutig, d.h. jeder Punkt (r, θ) wird auf genau einen Punkt (x^1, x^2) abgebildet und umgekehrt. Dabei wird der Nullpunkt $(x^1 = 0, x^2 = 0)$ aber nicht erreicht, was unbefriedigend ist. Mit $r \in \mathbb{R}_+$ und $\theta \in (-\pi, \pi]$ ist $(r, \theta) \mapsto (x^1, x^2)$ nicht mehr eineindeutig, da alle Punkte $(r = 0, \theta)$ auf $(x^1 = 0, x^2 = 0)$ abgebildet werden.

Aus der Eineindeutigkeit — hier für $r \neq 0$ bzw. $(x^1, x^2) \neq (0, 0)$ — folgt die (lokale) Umkehrbarkeit einer Abbildung. Um zu wissen, ob eine Koordinatentransformation (hier bezeichnet d die Dimension des Raums: $d = 2$ für eine Ebene, $d = 3$ für den ganzen Ortsraum)

$$(x^{1'}, x^{2'}, \dots, x^{d'}) \mapsto (x^1, x^2, \dots, x^d) \quad (\text{I.62})$$

lokal umkehrbar ist, muss man die *Funktionaldeterminante*

$$\frac{\partial(x^1, x^2, \dots, x^d)}{\partial(x^{1'}, x^{2'}, \dots, x^{d'})} \equiv \det \begin{pmatrix} \frac{\partial x^1}{\partial x^{1'}} & \cdots & \frac{\partial x^1}{\partial x^{d'}} \\ \vdots & \ddots & \vdots \\ \frac{\partial x^d}{\partial x^{1'}} & \cdots & \frac{\partial x^d}{\partial x^{d'}} \end{pmatrix} \quad (\text{I.63})$$

betrachten: Die Abbildung (I.62) ist genau dann lokal umkehrbar, wenn die Determinante (I.63) ungleich null ist.

Im Fall von Polarkoordinaten ($x^{1'} = r, x^{2'} = \theta$) gibt Gl. (I.60) die Funktionaldeterminante

$$\frac{\partial(x^1, x^2)}{\partial(r, \theta)} = r,$$

d.h. man findet, dass die Abbildung $(r, \theta) \mapsto (x^1, x^2)$ fast überall lokal umkehrbar ist, außer bei $r = 0$.

Diesen Polarkoordinaten werden oft orthonormierte Basisvektoren $(\vec{e}_r, \vec{e}_\theta)$ wie in Abb. I.7 assoziiert: \vec{e}_r ist der Einheitsvektor entlang der Richtung von \vec{r} , d.h. $\vec{e}_r \equiv \vec{r}/r$ — wobei wieder $\vec{r} \neq \vec{0}$ angenommen wird —, und \vec{e}_θ ist orthogonal dazu. Durch die kartesischen Basisvektoren kann man sie als

$$\begin{cases} \vec{e}_r = \cos \theta \vec{e}_1 + \sin \theta \vec{e}_2 \\ \vec{e}_\theta = -\sin \theta \vec{e}_1 + \cos \theta \vec{e}_2 \end{cases} \quad (\text{I.64})$$

ausdrücken.

Eine wichtige Besonderheit dieser Basisvektoren ist, dass sie vom Ortsvektor \vec{r} selbst abhängen — oder genauer, vom Polarwinkel θ —, auch wenn diese Abhängigkeit üblicherweise nicht explizit geschrieben wird. Dies muss berücksichtigt werden, wenn man z.B. die Zeit-Ort-Funktion

$$\vec{x}(t) = r(t) \vec{e}_r \quad (\text{I.65})$$

nach der Zeit ableitet, denn \vec{e}_r wird eigentlich im Punkt $\vec{x}(t)$ bestimmt, d.h. hängt von der Zeit ab. Die Produktregel gibt dann

$$\vec{v}(t) \equiv \dot{\vec{x}}(t) = \dot{r}(t) \vec{e}_r + r(t) \dot{\vec{e}}_r,$$

was unter Nutzung von der Kettenregel

$$\dot{\vec{e}}_r = \frac{d\vec{e}_r}{d\theta} \frac{d\theta}{dt}$$

und von $d\vec{e}_r/d\theta = \vec{e}_\theta$ [vgl. Gl. (I.64)] zu

$$\vec{v}(t) = \dot{r}(t) \vec{e}_r + r(t) \dot{\theta}(t) \vec{e}_\theta \quad (\text{I.66})$$

führt. Das heißt, die Komponenten von $\vec{v}(t)$ in der mitbewegten Basis $(\vec{e}_r, \vec{e}_\theta)$ sind $v^r(t) = \dot{r}(t)$ und $v^\theta(t) = r(t) \dot{\theta}(t)$.

Wiederum gibt eine zweite Ableitung nach t die Beschleunigung

$$\vec{a}(t) = [\ddot{r}(t) - r(t) \dot{\theta}(t)^2] \vec{e}_r + [r(t) \ddot{\theta}(t) + 2\dot{r}(t) \dot{\theta}(t)] \vec{e}_\theta, \quad (\text{I.67})$$

wobei $\dot{\vec{e}}_\theta = \dot{\theta}(t) d\vec{e}_\theta/d\theta = -\dot{\theta}(t) \vec{e}_r$ benutzt wurde.

I.A.2c Zylinderkoordinaten

Geht man jetzt zurück zu einem dreidimensionalen Raum, kann man anstatt kartesischer Koordinaten (x^1, x^2, x^3) (relativ zu einem Nullpunkt O) *Zylinderkoordinaten* (ρ, θ, z) verwenden, die über die Beziehungen

$$\begin{cases} x^1 = \rho \cos \theta \\ x^2 = \rho \sin \theta \\ x^3 = z \end{cases} \quad (\text{I.68})$$

definiert sind, mit Polarkoordinaten (ρ, θ) in der (x^1, x^2) -Ebene. Selbstverständlich ist die Rücktransformation von (x^1, x^2) nach (ρ, θ) dieselbe wie bei Polarkoordinaten, während die Transformation von x^3 nach z trivial ist.

Die orthonormierten Basisvektoren $(\vec{e}_\rho, \vec{e}_\theta, \vec{e}_z)$ sind so gewählt, dass $(\vec{e}_\rho, \vec{e}_\theta)$ in der (x^1, x^2) -Ebene liegen:

$$\begin{cases} \vec{e}_\rho = \cos \theta \vec{e}_1 + \sin \theta \vec{e}_2 \\ \vec{e}_\theta = -\sin \theta \vec{e}_1 + \cos \theta \vec{e}_2 \\ \vec{e}_z = \vec{e}_3. \end{cases} \quad (\text{I.69})$$

Somit ist \vec{e}_z raumfest, \vec{e}_ρ und \vec{e}_θ aber nicht. Dann gilt für den Ortsvektor

$$\vec{r} = \rho \vec{e}_\rho + z \vec{e}_z. \quad (\text{I.70})$$

Für die Bahnkurve, Geschwindigkeit und Beschleunigung eines Massenpunkts findet man

$$\vec{x}(t) = \rho(t) \vec{e}_\rho + z(t) \vec{e}_z \quad (\text{I.71})$$

$$\vec{v}(t) = \dot{\rho}(t) \vec{e}_\rho + \rho(t) \dot{\theta}(t) \vec{e}_\theta + \dot{z}(t) \vec{e}_z \quad (\text{I.72})$$

$$\vec{a}(t) = [\ddot{\rho}(t) - \rho(t) \dot{\theta}(t)^2] \vec{e}_\rho + [\rho(t) \ddot{\theta}(t) + 2\dot{\rho}(t) \dot{\theta}(t)] \vec{e}_\theta + \ddot{z}(t) \vec{e}_z. \quad (\text{I.73})$$

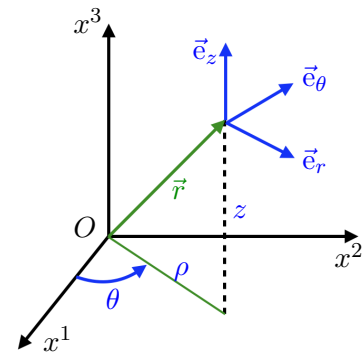


Abbildung I.8

I.A.2d Kugelkoordinaten

In Problemen mit annähernder Kugelsymmetrie um einen Punkt O kann man angepasste Koordinaten verwenden, und zwar *Kugelkoordinaten* (r, θ, φ) , deren Variablentransformation zu kartesischen Koordinaten (x^1, x^2, x^3) durch

$$\begin{cases} x^1 = r \sin \theta \cos \varphi \\ x^2 = r \sin \theta \sin \varphi \\ x^3 = r \cos \theta \end{cases} \quad (\text{I.74})$$

gegeben sind. Für die *Radialkoordinate* gilt $r \in \mathbb{R}_+$, für den *Polarwinkel* $\theta \in [0, \pi]$, und für den *Azimutwinkel* $\varphi \in [0, 2\pi)$.

Die entsprechende Funktionaldeterminante (I.63) lautet

$$\frac{\partial(x^1, x^2, x^3)}{\partial(r, \theta, \varphi)} = r^2 \sin \theta,$$

so dass die Transformation für $r \neq 0$ und $\theta \neq 0, \pi$ — d.h. außer der x^3 -Achse — umkehrbar ist.

Die zugehörigen orthogonalen Einheitsvektoren sind gegeben durch

$$\begin{cases} \vec{e}_r = \sin \theta \cos \varphi \vec{e}_1 + \sin \theta \sin \varphi \vec{e}_2 + \cos \theta \vec{e}_3 \\ \vec{e}_\theta = \cos \theta \cos \varphi \vec{e}_1 + \cos \theta \sin \varphi \vec{e}_2 - \sin \theta \vec{e}_3 \\ \vec{e}_\varphi = -\sin \varphi \vec{e}_1 + \cos \varphi \vec{e}_2. \end{cases} \quad (\text{I.75})$$

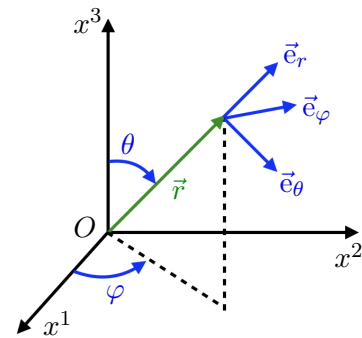


Abbildung I.9

Wieder sind diese Basisvektoren nicht raumfest, sondern hängen von (θ, φ) ab. Daraus prüft man die Beziehungen $d\vec{e}_r/d\theta = \vec{e}_\theta$ und $d\vec{e}_\theta/d\theta = -\vec{e}_r$ — wie bei Polarkoordinaten —, $d\vec{e}_r/d\varphi = -\sin\theta \vec{e}_\varphi$ und $d\vec{e}_\theta/d\varphi = -\cos\theta \vec{e}_\varphi$, und schließlich $d\vec{e}_\varphi/d\varphi = -\sin\theta \vec{e}_r - \cos\theta \vec{e}_\theta$ und $d\vec{e}_\theta/d\theta = 0$.

Für den Ortsvektor gilt

$$\vec{r} = r \vec{e}_r, \quad (\text{I.76})$$

d.h. \vec{e}_r ist wieder entlang \vec{r} . Daraus folgt sofort für die Zeit-Ort-Funktion eines Massenpunkts

$$\vec{x}(t) = r(t) \vec{e}_r. \quad (\text{I.77})$$

Sukzessiven Ableitung nach der Zeit liefern dann die Geschwindigkeit

$$\vec{v}(t) = \dot{r}(t) \vec{e}_r + r(t) \dot{\theta}(t) \vec{e}_\theta + r(t) \sin\theta(t) \dot{\varphi}(t) \vec{e}_\varphi \quad (\text{I.78})$$

und die Beschleunigung

$$\begin{aligned} \vec{a}(t) = & [\ddot{r}(t) - r(t) \dot{\theta}(t)^2 - r(t) \sin^2\theta(t) \dot{\varphi}(t)^2] \vec{e}_r \\ & + [r(t) \ddot{\theta}(t) + 2\dot{r}(t) \dot{\theta}(t) - r(t) \sin\theta(t) \cos\theta(t) \dot{\varphi}(t)^2] \vec{e}_\theta \\ & + [r(t) \sin\theta(t) \ddot{\varphi}(t) + 2\sin\theta(t) \dot{r}(t) \dot{\varphi}(t) + 2r(t) \cos\theta(t) \dot{\theta}(t) \dot{\varphi}(t)] \vec{e}_\varphi. \end{aligned} \quad (\text{I.79})$$

I.A.3 Einfache Bewegungen

I.A.3a Gleichförmige geradlinige Bewegungen

$\vec{a}(t) = \vec{0}$ für alle t mit den Anfangsbedingungen $\vec{x}(t_0) = \vec{x}_0$, $\vec{v}(t_0) = \vec{v}_0$ zu einer Zeit t_0 . Aus Gl. (I.56a) kommt sofort $\vec{v}(t) = \vec{v}_0$ für alle t . Dann gibt eine Integration über die Zeit

$$\vec{x}(t) = \vec{x}_0 + (t - t_0) \vec{v}_0.$$

I.A.3b Gleichmäßig beschleunigte Bewegung

Betrachte jetzt eine Bewegung mit der konstanten Beschleunigung $\vec{a}(t) = \vec{a}_0 \neq \vec{0}$ für alle t , und mit den Anfangsbedingungen $\vec{x}(t_0) = \vec{x}_0$, $\vec{v}(t_0) = \vec{v}_0$ zu einer Zeit t_0 . Eine erste Integration über t [Gl. (I.56a)] gibt

$$\vec{v}(t) = \vec{v}_0 + (t - t_0) \vec{a}_0 \text{ für alle } t.$$

Eine zweite Integration über die Zeit liefert dann

$$\vec{x}(t) = \vec{x}_0 + (t - t_0) \vec{v}_0 + \frac{(t - t_0)^2}{2} \vec{a}_0.$$

Wenn \vec{v}_0 und \vec{a}_0 nicht parallel zueinander sind, ist die Bahnkurve eine Parabel.

I.A.3c Kreisbewegung

Es wird jetzt angenommen, dass sich der Massenpunkt entlang eines Kreises \mathcal{C} mit Radius R bewegt. Der Einfachheit halber benutzt man zuerst ein System von Polarkoordinaten (r, θ) in der Ebene des Kreises mit dem Nullpunkt im Zentrum von \mathcal{C} . Sei $\theta(t_0) = \theta_0$ der Polarwinkel des Massenpunkts zur Zeit t_0 .

Offensichtlich bleibt die Radialkoordinate $r(t)$ des Massenpunkts konstant mit der Zeit: $r(t) = R$. Somit nehmen die Geschwindigkeit (I.66) und die Beschleunigung (I.67) die einfacheren Formen

$$\vec{v}(t) = R \dot{\theta}(t) \vec{e}_\theta, \quad \vec{a}(t) = -R \dot{\theta}(t)^2 \vec{e}_r + R \ddot{\theta}(t) \vec{e}_\theta \quad (\text{I.80})$$

an. Dabei muss nicht vergessen werden, dass \vec{e}_r und \vec{e}_θ vom Polarwinkel $\theta(t)$ abhängen.

Zu jeder Zeit ist der Geschwindigkeitsvektor parallel zum instantanen Basisvektor \vec{e}_θ , d.h. tangential zum Kreis. Die Änderungsrate des Polarwinkels $\dot{\theta}(t) \equiv \omega(t)$ wird *Winkelgeschwindigkeit* genannt. Mit deren Hilfe lautet der Betrag der Geschwindigkeit $|\vec{v}(t)| = R|\omega(t)|$.

Wiederum lässt sich die Beschleunigung in zwei Komponenten zerlegen:

- eine Radialkomponente entlang \vec{e}_r , die *Zentripetalbeschleunigung* $a^r(t) = -R\dot{\theta}(t)^2 = -R\omega(t)^2$, d.h. $a^r(t) = -|\vec{v}(t)|^2/R$;
- eine Komponente entlang \vec{e}_θ , die *Tangentialbeschleunigung* $a^\theta(t) = R\ddot{\theta}(t) = R\dot{\omega}(t)$.

Für eine gleichförmige Kreisbewegung gilt $\ddot{\theta}(t) = \dot{\omega}(t) = 0$ zu jeder Zeit t , d.h. $\omega(t)$ bleibt konstant. In diesem Fall die Beschleunigung rein radial (und zum Kreiszentrum ausgerichtet), und die Zeit-Ort-Funktion wird zu

$$r(t) = R \quad , \quad \theta(t) = \theta_0 + \omega_0(t - t_0),$$

mit dem konstanten Wert ω_0 der Winkelgeschwindigkeit.

Benutzt man jetzt kartesische Koordinaten (x, y) in der Ebene des Kreises, wieder mit dem Nullpunkt im Kreiszentrum, so lautet die Zeit-Ort-Funktion (genauer: deren Komponenten) für eine gleichförmige Kreisbewegung mit Winkelgeschwindigkeit ω_0

$$x(t) = R \cos[\theta_0 + \omega_0(t - t_0)] \quad , \quad y(t) = R \sin[\theta_0 + \omega_0(t - t_0)].$$

KAPITEL II

Newton'sche Mechanik: Anwendungen

II.1	Mehrteilchensysteme	38
II.1.1	Grundlagen	39
II.1.2	Bewegung des Schwerpunkts	41
II.1.3	Drehimpuls	41
II.1.4	Energie	42
II.1.5	Schwerpunktsystem	46
II.2	Zwei-Körper-Systeme	47
II.2.1	Separation der Bewegungsgleichungen	47
II.2.2	Gekoppelte Punktmassen	49
II.2.3	Kepler-Problem	49
II.2.4	Streuung	56

In diesem Kapitel werden einige wichtigen Anwendungen des im Kap. I eingeführten newtonschen Formalismus vorgestellt.

Zuerst wird die genauere Formulierung der Bewegungsgleichungen für ein mechanisches System aus mehreren Massenpunkten unter Berücksichtigung der newtonschen Gesetze in Abschn. II.1 dargelegt. Dabei lässt sich die Bewegung in zwei Anteile zerlegen, und zwar einerseits die globale Bewegung des Systems — oder genauer seines Schwerpunkts —, andererseits die Bewegung der Teile des Systems relativ zueinander.

Danach wird der besondere Fall von Zwei-Körper-Systemen im Kap. II.2 betrachtet, wobei der entwickelte Formalismus auf einige Beispiele angewandt wird.

II.1 Mehrteilchensysteme

Die in Abschn. I.2 eingeführten newtonschen Gesetze bzw. der in § I.2.5 gefundene Energiesatz beziehen sich auf der Bewegung eines einzigen Massenpunkts unter dem Einfluss äußerer Kräfte bzw. in einem äußeren Potential — obwohl das dritte Gesetz eigentlich schon Aufschluss über die Kräfte zwischen zwei Körpern gibt.

Im Gegensatz befasst sich dieser Abschnitt mit der Bewegung eines Systems Σ aus $N \geq 2$ Massenpunkten, die hiernach auch als „Teilchen“ bezeichnet werden. Ein solches System wird *Mehrteilchensystem* genannt. Im Folgenden werden die unterschiedlichen Massenpunkte mit Indizes a , $b = 1, 2, \dots, N$ gekennzeichnet: ihre Ortsvektoren relativ zu einem festen Bezugssystem werden mit $\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N \equiv \{\vec{r}_a\}_{1 \leq a \leq N}$ bezeichnet; ihre Bahnkurven mit $\{\vec{x}_a(t)\}_{1 \leq a \leq N}$; ihre (als konstant angenommenen) Massen mit $\{m_a\}_{1 \leq a \leq N}$; ihre Impulse mit $\{\vec{p}_a(t)\}_{1 \leq a \leq N}$; usw.

Als einfache Beispiele eines solchen Systems kann man entweder die Erde um die Sonne⁽¹⁷⁾ oder zwei durch eine Feder gekoppelte Massen betrachten, entsprechend $N = 2$. Der Fall $N = 3$ kann z.B. mit dem System {Sonne + Erde + Mond} illustriert werden.⁽¹⁷⁾ Unter Vernachlässigung des

⁽¹⁷⁾Dabei werden die Sonne, die Erde, der Mond, die Sterne als Punktmassen modelliert, was ziemlich sinnvoll ist, denn ihre jeweiligen Radien sind viel kleiner als ihr typischer Abstand.

interstellaren Mediums sind Galaxien Systeme aus $N \approx 10^9$ – 10^{12} Sternen.⁽¹⁷⁾ Schließlich stellen die Moleküle in einem Kubikmeter von Gas ein Beispiel mit $N \approx 10^{24}$ dar.

Die üblichen Größen zur Beschreibung eines Mehrteilchensystems, sowie die in der klassischen Mechanik üblichen Modellierung der für ein solches System relevanten Kräfte, werden in § II.1.1 dargelegt. In § II.1.2–II.1.4 werden Aussagen über die Zeitentwicklung einiger charakteristischen Größen — Impuls, Drehimpuls und Energie — hergeleitet. Insbesondere wird sich herausstellen, dass diese Größen bei Mehrteilchensystemen, die mit ihrer Umgebung nicht wechselwirken, erhalten sind, d.h. sie bleiben konstant in der Zeit. Schließlich befasst sich § II.1.5 mit dem „geeigneten“ Bezugssystem, in welchem die Teilchenbewegungen entkoppelt von der globalen Bewegung des Systems sind.

II.1.1 Grundlagen

II.1.1 a Modellierung des Systems. Bewegungsgleichungen

Um die Bewegungsgleichung eines der Massenpunkte des Systems Σ zu erhalten, sollte man zunächst die Kräfte auf ein solches Teilchen kennen. Der Einfachheit halber wird die Bewegung des Mehrteilchensystems in einem Inertialsystem \mathcal{B} untersucht wird, so dass keine Scheinkräfte vorhanden sind. Somit wird angenommen, dass die gesamte Kraft auf Teilchen a sich als

$$\vec{F}_a = \vec{F}_{a,\text{ext}} + \sum_{\substack{b=1 \\ b \neq a}}^N \vec{F}_{b \rightarrow a} \quad (\text{II.1})$$

schreiben lässt, wobei eine mögliche Zeit- oder Ortsabhängigkeit nicht geschrieben wurde. Dabei stellt $\vec{F}_{a,\text{ext}}$ die Resultierende aus den auf a wirkenden *externen* Kräften dar, die auch *äußere* Kräfte genannt werden. Dieser Term beschreibt den Einfluss von Ursachen, die nicht teil vom betrachteten Mehrteilchensystem sind, wie z.B. eines Schwerfeldes, in dem die Teilchen sich befinden.

Dagegen steht $\vec{F}_{b \rightarrow a}$ für eine *innere Kraft*, und zwar für diejenige, die das b -te Teilchen von Σ auf das Teilchen a ausübt. Damit die Massenpunkte des Mehrteilchensystems sich gegenseitig beeinflussen, sollten diese inneren Kräfte nicht identisch Null sein.

Um Gleichungen kürzer zu machen, kann man eine innere Kraft $\vec{F}_{a \rightarrow a} \equiv \vec{0}$ definieren. Mit diesem Trick läuft z.B. die Summe in Gl. (II.1) über alle Werte von b von 1 bis N , ohne a auszuschließen.

Bemerkung: Die Zerlegung (II.1) enthält eigentlich eine wichtige Annahme, und zwar dass es keine Drei-Körper-Kräfte im System gibt. Das heißt, Teilchen a unterliegt keiner Kraft, die durch Teilchen b und b' *gemeinsam* verursacht wird: der ganze Einfluss dieser beiden Teilchen auf a lässt sich als Summe von Zwei-Körper-Kräften $\vec{F}_{b \rightarrow a} + \vec{F}_{b' \rightarrow a}$ beschreiben. A fortiori gibt es keine Vier-, Fünf- usw. Körper-Kräfte. Diese Annahme liegt eigentlich der ganzen newtonschen Mechanik zu Grunde, denn sie ist für das dritte Gesetz (I.19) nötig.

Laut dem zweiten newtonschen Gesetz (I.14) lautet die dynamische *Bewegungsgleichung* für jedes Teilchen $a = 1, \dots, N$

$$\frac{d\vec{p}_a(t)}{dt} = m_a \frac{d^2\vec{x}_a(t)}{dt^2} = \vec{F}_a = \vec{F}_{a,\text{ext}} + \sum_{b=1}^N \vec{F}_{b \rightarrow a} \quad \text{für jedes } a \in \{1, 2, \dots, N\}. \quad (\text{II.2})$$

Die N vektoriellen Gleichungen (II.2) bzw. die $3N$ entsprechenden skalaren Gleichungen, die sich aus Projektion auf drei Koordinatenachsen ergeben, bilden ein System aus gekoppelten Differentialgleichungen, welche die Bewegung des Mehrteilchensystems beschreiben.

II.1.1 b Grundbegriffe und Definitionen

Die *Gesamtmasse* des Mehrteilchensystems Σ wird definiert als

$$M \equiv \sum_{a=1}^N m_a. \quad (\text{II.3})$$

Diese Definition sieht zwar trivial aus, sie gilt aber nur als newtonsche, nicht-relativistische Näherung. Eigentlich ist die Masse eines gebundenen Systems kleiner als die Summe seiner Bestandteile — z.B. ist die Masse eines Atomkerns kleiner als die Summe der Massen dessen Protonen und Neutronen —, wobei die Differenz, der „Massendefekt“, mit der Bindungsenergie des Systems zusammenhängt.

Zur Charakterisierung der Bewegung von Σ lohnt es sich, den *Schwerpunkt* des Mehrteilchensystems einzuführen. Dabei handelt es sich erstens um den geometrischen Punkt mit Ortsvektor

$$\vec{R} \equiv \frac{1}{M} \sum_{a=1}^N m_a \vec{r}_a. \quad (\text{II.4})$$

Entsprechend dieser Definition spricht man auch von dem *Massenmittelpunkt*. In ähnlicher Weise lautet die Bahnkurve des Schwerpunkts

$$\vec{X}(t) \equiv \frac{1}{M} \sum_{a=1}^N m_a \vec{x}_a(t). \quad (\text{II.5})$$

In einem zweiten Schritt kann man diesem geometrischen Schwerpunkt die Gesamtmasse M des Mehrteilchensystems zuordnen, so dass als *Schwerpunkt* jetzt einen fiktiven Massenpunkt mit Ortsvektor \vec{R} bzw. Trajektorie $\vec{X}(t)$ bezeichnet wird.

Addiert man die Impulse bzw. Drehimpulse aller Teilchen von Σ , so erhält man den *Gesamtimpuls*

$$\vec{P}(t) \equiv \sum_{a=1}^N \vec{p}_a(t) \quad (\text{II.6})$$

bzw. den *Gesamtdrehimpuls*

$$\vec{L}(t) \equiv \sum_{a=1}^N \vec{x}_a(t) \times \vec{p}_a(t) \quad (\text{II.7})$$

des Mehrteilchensystems. Dabei findet man einfach, dass der Gesamtimpuls $\vec{P}(t)$ mit der Gesamtmasse M und der Geschwindigkeit $d\vec{X}(t)/dt$ des Schwerpunkts zusammenhängt, und zwar über

$$\vec{P}(t) = M \frac{d\vec{X}(t)}{dt}. \quad (\text{II.8})$$

Diese Beziehung folgt sofort aus $\vec{P}(t) = \sum_{a=1}^N m_a \frac{d\vec{x}_a(t)}{dt} = M \frac{d}{dt} \left[\frac{1}{M} \sum_{a=1}^N m_a \vec{x}_a(t) \right]$. \square

Definition: Eine Funktion $f(t, \{\vec{r}_a\}, \{\vec{v}_a\})$ der Ortsvektoren und Geschwindigkeiten der N Teilchen heißt *Erhaltungsgröße* oder *Konstante der Bewegung*, wenn sie für alle Lösungen $\{\vec{r}_a = \vec{x}_a(t)\}$, $\{\vec{v}_a = \dot{\vec{x}}_a(t)\}$ der Bewegungsgleichungen (II.2) konstant in der Zeit bleibt:

$$\frac{d}{dt} f(t, \{\vec{x}_a(t)\}, \{\dot{\vec{x}}_a(t)\}) = 0. \quad (\text{II.9})$$

Bemerkungen:

- * Der Wert der Konstanten hängt im Allgemeinen von der Lösung ab!
- * Die Zeitableitung in Definition (II.9) ist eine totale Ableitung, nicht nur eine partielle Ableitung nach dem ersten Argument (t) der Funktion f .

II.1.2 Bewegung des Schwerpunkts

Summiert man die individuellen Bewegungsgleichungen (II.2) über alle Teilchen bzw. alle Werte von a , so ergibt sich

$$\frac{d\vec{P}(t)}{dt} = \sum_{a=1}^N \vec{F}_{a,\text{ext}} + \sum_{a,b=1}^N \vec{F}_{b \rightarrow a}. \quad (\text{II.10})$$

Unter Verwendung der Gl. (II.8) ist der linke Term dieser Gleichung einfach gleich dem Produkt aus der Gesamtmasse und der Beschleunigung des Schwerpunkts, $M d^2\vec{X}(t)/dt^2$.

Für den zweiten Term im rechten Glied von Gl. (II.10) kann man schreiben

$$\sum_{a,b=1}^N \vec{F}_{b \rightarrow a} = \frac{1}{2} \sum_{a,b=1}^N (\vec{F}_{b \rightarrow a} + \vec{F}_{a \rightarrow b}) = \frac{1}{2} \sum_{a,b=1}^N (\vec{F}_{b \rightarrow a} + \vec{F}_{a \rightarrow b}) = \vec{0}, \quad (\text{II.11})$$

wobei die erste Gleichung trivial ist, während die zweite aus einer Umbenennung der Indizes im zweiten Summanden in den Klammern, und die dritte aus dem dritten newtonschen Axiom (I.19) folgt.

Somit führt Gl. (II.10) insgesamt zur Bewegungsgleichung

$$\frac{d\vec{P}(t)}{dt} = M \frac{d^2\vec{X}(t)}{dt^2} = \sum_{a=1}^N \vec{F}_{a,\text{ext}}. \quad (\text{II.12a})$$

Diese Gleichung stellt die mathematische Formulierung des *Schwerpunktsatzes* dar:

Der Schwerpunkt eines N -Teilchensystems bewegt sich nur unter dem Einfluss der Summe der äußeren Kräfte auf die Teilchen, und zwar wie ein Massenpunkt mit der Masse M , an dem diese Summe angreift. (II.12b)

Falls keine äußeren Kräfte vorliegen — d.h. für ein sog. *abgeschlossenes System*, ohne Wechselwirkung mit seiner Umgebung —, oder wenn ihre Resultierende verschwindet, dann gilt einfach $d\vec{P}(t)/dt = \vec{0}$, d.h. der Gesamtimpuls $\vec{P}(t)$ ist erhalten:

Wenn die Resultierende aller äußeren Kräfte auf ein Mehrteilchensystem Σ verschwindet, dann ist der Gesamtimpuls von Σ eine Erhaltungsgröße. (II.12c)

II.1.3 Drehimpuls

Als nächstes können wir jetzt die Zeitentwicklung des durch Gl. (II.7) definierten Gesamtdrehimpulses untersuchen. Leitet man dessen Ausdruck nach der Zeit ab, so ergibt sich nach Anwendung der Produktregel

$$\frac{d\vec{L}(t)}{dt} = \sum_{a=1}^N \frac{d\vec{x}_a(t)}{dt} \times \vec{p}_a(t) + \sum_{a=1}^N \vec{x}_a(t) \times \frac{d\vec{p}_a(t)}{dt}.$$

Da die Geschwindigkeit $d\vec{x}_a(t)/dt$ parallel zum Impuls ist, verschwindet der erste Term auf der rechten Seite. Unter Verwendung der Bewegungsgleichung (II.2) für Teilchen a erhält man dann

$$\frac{d\vec{L}(t)}{dt} = \sum_{a=1}^N \vec{x}_a(t) \times \vec{F}_a = \sum_{a=1}^N \vec{x}_a(t) \times \vec{F}_{a,\text{ext}} + \sum_{a,b=1}^N \vec{x}_a(t) \times \vec{F}_{b \rightarrow a}, \quad (\text{II.13})$$

wobei die zweite Gleichung aus der Aufspaltung (II.1) der Kraft auf a folgt.

Jeder Beitrag $\vec{x}_a(t) \times \vec{F}_{a,\text{ext}}$ für $a \in \{1, \dots, N\}$ ist das *Drehmoment* der äußeren Kraft $\vec{F}_{a,\text{ext}}$, d.h. der erste Term auf der rechten Seite von Gl. (II.13) ist die Resultierende dieser Drehmomente.

Ähnlich wie in Gl.(II.11) kann man den zweiten Term im Ausdruck ganz rechts von Gl. (II.13) umschreiben, und zwar als

$$\sum_{a,b=1}^N \vec{x}_a(t) \times \vec{F}_{b \rightarrow a} = \frac{1}{2} \sum_{a,b=1}^N [\vec{x}_a(t) \times \vec{F}_{b \rightarrow a} + \vec{x}_b(t) \times \vec{F}_{a \rightarrow b}].$$

Laut dem dritten newtonschen Gesetz kann $\vec{F}_{a \rightarrow b}$ durch $-\vec{F}_{b \rightarrow a}$ ersetzt werden, woraus sich

$$\sum_{a,b=1}^N \vec{x}_a(t) \times \vec{F}_{b \rightarrow a} = \frac{1}{2} \sum_{a,b=1}^N [\vec{x}_a(t) - \vec{x}_b(t)] \times \vec{F}_{b \rightarrow a}$$

ergibt. Falls die Kraft $\vec{F}_{b \rightarrow a}$ entlang des Abstandvektors $\vec{x}_b(t) - \vec{x}_a(t)$ von Teilchen a nach Teilchen b liegt — was z.B. der Fall für die Coulomb-Kraft zwischen zwei Punktladungen oder für die newtonschen Schwerkraft zwischen zwei Punktmassen ist —, wird $\vec{F}_{b \rightarrow a}$ *Zentralkraft* genannt. In diesem Fall verschwindet das Kreuzprodukt auf der rechten Seite der obigen Gleichung, und die Gl. (II.13) vereinfacht sich zu

$$\frac{d\vec{L}(t)}{dt} = \sum_{a=1}^N \vec{x}_a(t) \times \vec{F}_{a,\text{ext}}. \quad (\text{II.14a})$$

Dieses Ergebnis stellt den *Drehimpulssatz* für ein Mehrteilchensystem dar, und zwar

Für ein System aus N Teilchen, zwischen denen nur Zentralkräfte wirken, ist die Rate der Änderung des Gesamtdrehimpulses gleich der Resultierenden von den äußeren Drehmomenten. (II.14b)

Falls keine äußeren Kräfte vorliegen, oder wenn die Resultierende ihrer Drehmomente verschwindet, gilt $d\vec{L}(t)/dt = \vec{0}$, d.h. der Gesamtdrehimpuls $\vec{L}(t)$ ist erhalten:

Wenn die Resultierende aller äußeren Drehmomente auf ein Mehrteilchensystem mit nur zentralen inneren Kräften verschwindet, dann ist der Gesamtdrehimpuls des Systems eine Erhaltungsgröße. (II.14c)

Bemerkungen:

* Zentralkräfte werden oft definiert als Kräfte, die immer entlang der Richtung bis zu einem festen Punkt — dem *Kraftzentrum* — liegen. Die hier adoptierte Definition, laut der die Zweikörperkraft $\vec{F}_{b \rightarrow a}$ immer in Richtung des Abstandsvektors zwischen Teilchen a und b zeigt, lässt sich ähnlich betrachten: wird die Position vom Massenpunkt b als fest angenommen, ist die durch Teilchen b auf jedes andere Teilchen a ausgeübte Kraft eine Zentralkraft bezüglich der Position von b .

* Hier wurde in der Herleitung der Ergebnisse (II.13)–(II.14c) angenommen, dass die inneren Kräfte zentral sind. Eigentlich ist die Erhaltung des Gesamtdrehimpulses in abgeschlossenen Systemen viel allgemeiner — wie wir in § III.3.1 c sehen werden, spiegelt diese Erhaltung die Isotropie des Raums wider — und gilt auch für Systeme mit nicht-zentralen Kräften.

II.1.4 Energie

II.1.4a Energiesatz

Die kinetische Energie des a -ten Teilchens eines Systems ist $T_a = \vec{p}_a^2/2m_a$, vgl. Definition (I.21). Logischerweise wird die *gesamte kinetische Energie* eines Mehrteilchensystems definiert als

$$T \equiv \sum_{a=1}^N \frac{\vec{p}_a^2}{2m_a}. \quad (\text{II.15})$$

Sei jetzt angenommen, dass sowohl die äußeren als die inneren Kräfte, die auf die Körper des Mehrteilchensystems wirken, konservativ sind. Das heißt einerseits, dass es Potentiale $V_{a,\text{ext}}(\vec{r}_a)$ existieren, die

$$\vec{F}_{a,\text{ext}} = -\vec{\nabla}_a V_{a,\text{ext}}(\vec{r}_a) \quad \text{für } a \in \{1, \dots, N\} \quad (\text{II.16})$$

erfüllen, wobei $\vec{\nabla}_a$ den Gradienten bezüglich dem Ortsvektor \vec{r}_a bezeichnet.⁽¹⁸⁾ Andererseits gibt es paarweise Wechselwirkungspotentiale V_{ab} , die nur vom Abstand $|\vec{r}_a - \vec{r}_b|$ abhängen, aus denen die inneren Kräfte abgeleitet werden können, und zwar über

$$\vec{F}_{b \rightarrow a} = -\vec{\nabla}_a V_{ab}(|\vec{r}_a - \vec{r}_b|) \quad \text{für } a, b \in \{1, \dots, N\}. \quad (\text{II.17})$$

In diesem Fall ist die Zweikörperkraft der Form

$$\vec{F}_{b \rightarrow a} = f_{ba}(|\vec{r}_a - \vec{r}_b|) \vec{e}_{ba}$$

mit einer Funktion f_{ba} einer reellen Variablen und dem Einheitsvektor $\vec{e}_{ba} \equiv (\vec{r}_a - \vec{r}_b)/|\vec{r}_a - \vec{r}_b|$ entlang des Abstandsvektors der beiden Teilchen. Somit ist die Kraft zentral.

Damit das dritte newtonsche Gesetz erfüllt wird, muss $V_{ab} = V_{ba}$ für jedes Paar a, b gelten.

Die Kettenregel gibt nämlich $-\vec{\nabla}_a V_{ab}(|\vec{r}_a - \vec{r}_b|) = \vec{\nabla}_b V_{ab}(|\vec{r}_a - \vec{r}_b|)$.

Die *gesamte potentielle Energie* V des Mehrteilchensystems ist dann

$$V \equiv \sum_{a=1}^N V_{a,\text{ext}}(\vec{r}_a) + \sum_{1 \leq a < b \leq N} V_{ab}(|\vec{r}_a - \vec{r}_b|). \quad (\text{II.18})$$

Dabei muss die Wechselwirkungsenergie jedes Teilchenpaars nur einmal mitgezählt werden, weshalb die Summe über Paare mit $a < b$ eingeschränkt wurde. Äquivalent kann man die Summe im zweiten Term auf der rechten Seite noch anders umschreiben, und zwar entweder als die Hälfte der Summe über alle Werte von a und b mit der Einschränkung $a \neq b$, oder als die halbe Summe über alle Werte von a und b ohne Einschränkung, jedoch mit der Konvention $V_{aa} \equiv 0$:

$$\sum_{1 \leq a < b \leq N} V_{ab} = \frac{1}{2} \sum_{a=1}^N \sum_{\substack{b=1 \\ b \neq a}}^N V_{ab} = \frac{1}{2} \sum_{a,b=1}^N V_{ab} \quad \text{mit } V_{aa} = 0.$$

Bemerkung: Die Gesamtkraft \vec{F}_a auf das a -te Teilchen kann noch durch das Gesamtpotential V [Gl. (II.18)] ausgedrückt werden:

$$\vec{F}_a = -\vec{\nabla}_a V, \quad (\text{II.19})$$

wobei V als Funktion der N Ortsvektoren $\vec{r}_1, \dots, \vec{r}_N$ zu betrachten ist.

Unter den obigen Annahmen über die Kräfte ist die Summe

$$T + V \equiv \sum_{a=1}^N T_a + \left[\sum_{a=1}^N V_{a,\text{ext}} + \sum_{1 \leq a < b \leq N} V_{ab}(|\vec{r}_a - \vec{r}_b|) \right] \quad (\text{II.20})$$

aus allen kinetischen und potentiellen Energien eine Konstante der Bewegung, die *Gesamtenergie* des Systems.

Dieses Ergebnis stellt den *Energieerhaltungssatz* für ein Mehrteilchensystem dar, und zwar

Die Gesamtenergie eines Systems aus N Teilchen mit nur konservativen inneren und äußeren Kräften ist eine Erhaltungsgröße. (II.21)

⁽¹⁸⁾Hier ist diese Notation noch überflüssig, denn $V_{a,\text{ext}}$ ist nur Funktion der Ortskoordinaten des a -ten Körpers. Dagegen hängt V_{ab} von \vec{r}_a und \vec{r}_b ab, oder das in Gl. (II.18) definierte Potential V ist Funktion der Positionen aller Bestandteile des Systems, so dass die Präzisierung der Variablen, nach denen abgeleitet wird, dann nötig wird.

Beweis: Einerseits gilt unter Verwendung der Bewegungsgleichung (II.2)

$$\frac{dT}{dt} = \sum_{a=1}^N \frac{dT_a}{dt} = \sum_{a=1}^N \frac{\vec{p}_a(t)}{m} \cdot \frac{d\vec{p}_a(t)}{dt} = \sum_{a=1}^N \vec{v}_a(t) \cdot \vec{F}_a, \quad (\text{II.22})$$

mit der Geschwindigkeit \vec{v}_a von Teilchen a . Andererseits gibt die Kettenregel

$$\frac{dV}{dt} = \frac{d}{dt} V(\vec{x}_1(t), \dots, \vec{x}_N(t)) = \sum_{a=1}^N \left[\vec{\nabla}_a V(\vec{x}_1(t), \dots, \vec{x}_N(t)) \right] \cdot \frac{d\vec{x}_a(t)}{dt}.$$

Der Term in eckigen Klammern ist genau das Negative der Kraft auf Teilchen a , vgl. Gl. (II.19), so dass

$$\frac{dV}{dt} = - \sum_{a=1}^N \vec{F}_a \cdot \vec{v}_a(t) \quad (\text{II.23})$$

Insgesamt liefern Gl. (II.22) und (II.23) das gesuchte Ergebnis $d(T+V)/dt = 0$. \square

Bemerkung: Wenn einige der Kräfte nicht-konservativ sind, gilt Gl. (II.22) immer noch. Dagegen gilt Gl. (II.23) nicht mehr, denn nicht jede Kraft \vec{F}_a lässt sich aus einem Potential ableiten. Spaltet man die Kräfte in einen konservativen und einen nicht-konservativen Anteil auf, und definiert man die gesamte potentielle Energie der ersteren, so findet man einfach, dass die Rate $d(T+V)/dt$ der Änderung der Gesamtenergie gleich der gesamten Leistung der nicht-konservativen Kräfte ist:

$$\frac{d}{dt}(T+V) = \sum_{a=1}^N \vec{F}_{a,\text{diss.}} \cdot \vec{v}_a(t), \quad (\text{II.24})$$

wobei das hier verwendete Kürzel „diss.“ für *dissipative Kraft* steht, weil diese Kräfte zur nicht-Erhaltung der Energie, zu ihrer *Dissipation*, führen.

In der Praxis wird die dissipierte Energie nicht vernichtet, sondern in eine andere „nicht-mechanische“ Form umgewandelt, und zwar Wärme. Unter Berücksichtigung der letzteren bleibt die Gesamtenergie erhalten, entsprechend dem ersten Hauptsatz der Thermodynamik.

II.1.4 b Virialsatz

Wie kurz gesehen wurde bleibt die Gesamtenergie $T+V$ eines Mehrkörpersystems mit nur konservativen Kräften erhalten. Dagegen sind die kinetische und potentielle Energien im allgemeinen nicht getrennt konstant, sondern wandeln sich ständig in einander um. Man kann aber noch eine Aussage über die zeitlichen Mittelwerte von T und V — oder genauer, im allgemeinen Fall, vom Gradienten von V — machen.

Sei f eine Funktion einer reellen Variablen, und zwar hiernach der Zeit. Der zeitliche Mittelwert von $f(t)$ wird als

$$\langle f \rangle_t \equiv \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau f(t) dt \quad (\text{II.25})$$

definiert, wobei es angenommen wird, dass der Limes des Integrals existiert.

Unter Verwendung dieser Definition gilt den *Virialsatz*

$$2 \langle T \rangle_t = \left\langle \sum_{a=1}^N (\vec{\nabla}_a V) \cdot \vec{x}_a(t) \right\rangle_t. \quad (\text{II.26})$$

Beweis: Die Bewegungsgleichungen (II.2) geben

$$\sum_{a=1}^N m_a \ddot{\vec{x}}_a(t) \cdot \vec{x}_a(t) = \sum_{a=1}^N \vec{F}_a \cdot \vec{x}_a(t) = - \sum_{a=1}^N (\vec{\nabla}_a V) \cdot \vec{x}_a(t)$$

wobei die linke Seite noch als

$$\sum_{a=1}^N m_a \ddot{\vec{x}}_a(t) \cdot \vec{x}_a(t) = \frac{d}{dt} \left[\sum_{a=1}^N m_a \dot{\vec{x}}_a(t) \cdot \vec{x}_a(t) \right] - \sum_{a=1}^N m_a \dot{\vec{x}}_a(t)^2$$

geschrieben werden kann. Somit sind die zeitlichen Mittelwerte der Terme auf den rechten Seiten der letzten zwei Gleichungen gleich. Dabei gilt unter Nutzung der Definition (II.25)

$$\begin{aligned} \left\langle \frac{d}{dt} \left[\sum_{a=1}^N m_a \dot{\vec{x}}_a(t) \cdot \vec{x}_a(t) \right] \right\rangle_t &= \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau \frac{d}{dt} \left[\sum_{a=1}^N m_a \dot{\vec{x}}_a(t) \cdot \vec{x}_a(t) \right] dt \\ &= \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \sum_{a=1}^N m_a \dot{\vec{x}}_a(t) \cdot \vec{x}_a(t) \Big|_0^\tau. \end{aligned}$$

Wenn die Produkte $\dot{\vec{x}}_a(t) \cdot \vec{x}_a(t)$ immer abgegrenzt bleiben, nimmt die Summe im letzten Term einen endlichen Wert an, so dass der Grenzwert Null ist. Somit ergibt sich

$$\left\langle -\sum_{a=1}^N (\vec{\nabla}_a V) \cdot \vec{x}_a(t) \right\rangle_t = \left\langle -\sum_{a=1}^N m_a \dot{\vec{x}}_a(t)^2 \right\rangle_t.$$

Nach trivialen Umschreibung ist dies genau das gesuchte Ergebnis (II.26). \square

Für abgeschlossene Systeme — d.h. $V_{a,\text{ext}} = 0$ für alle Teilchen a — mit Wechselwirkungspotentialen der Form

$$V_{ab}(|\vec{r}_b - \vec{r}_a|) = \alpha_{ab} |\vec{r}_a - \vec{r}_b|^n \quad (\text{II.27})$$

mit $n \in \mathbb{Z}$ und $\alpha_{ab} \in \mathbb{R}$ für alle $a, b \in \{1, \dots, N\}$ nimmt der Virialsatz eine einfachere Form an.

In der Tat gelten

$$\vec{\nabla}_a |\vec{r}_b - \vec{r}_c| = (\delta_{ab} - \delta_{ac}) \frac{\vec{r}_b - \vec{r}_c}{|\vec{r}_b - \vec{r}_c|} \quad \text{und} \quad \frac{dV_{bc}}{d|\vec{r}_b - \vec{r}_c|} = n \frac{V_{bc}}{|\vec{r}_b - \vec{r}_c|},$$

woraus unter Anwendung der Kettenregel

$$\vec{\nabla}_a V = \vec{\nabla}_a \left[\sum_{b,c=1}^N V_{bc} \right] = \sum_{b,c=1}^N \frac{dV_{bc}}{d|\vec{r}_b - \vec{r}_c|} \vec{\nabla}_a |\vec{r}_b - \vec{r}_c| = n \sum_{b,c=1}^N \frac{V_{bc}}{|\vec{r}_b - \vec{r}_c|^2} (\vec{r}_b - \vec{r}_c) (\delta_{ab} - \delta_{ac})$$

folgt. Beim Durchführen des (Skalar)Produkts mit \vec{r}_a und der Summe über a führen die Terme $(\delta_{ab} - \delta_{ac})\vec{r}_a$ zu $\vec{r}_b - \vec{r}_c$, und daher zu

$$\sum_{a=1}^N (\vec{\nabla}_a V) \cdot \vec{r}_a = n \sum_{b,c=1}^N V_{bc} = 2nV.$$

Insgesamt vereinfacht sich der Virialsatz (II.26) zu

$$2 \langle T \rangle_t = n \langle V \rangle_t. \quad (\text{II.28a})$$

Daraus folgt dann für die Gesamtenergie — die wegen ihrer Erhaltung natürlich gleich ihrem zeitlichen Mittelwert ist

$$E = \langle E \rangle_t = \langle T \rangle_t + \langle V \rangle_t = \left(\frac{n}{2} + 1 \right) \langle V \rangle_t = \left(\frac{2}{n} + 1 \right) \langle T \rangle_t. \quad (\text{II.28b})$$

Ein erstes wichtiges Anwendungsbeispiel ist das von Massenpunkten, die unter einander über das newtonsche Gravitationspotential $V_{ab} \propto 1/|\vec{r}_a - \vec{r}_b|$ wechselwirken, entsprechend Gl. (II.27) mit $n = -1$. Diese Massenpunkte können z.B. die Sterne eines Kugelsternhaufens oder die Mitglieder eines Galaxienhaufens darstellen.⁽¹⁹⁾ In diesem Fall lautet der Virialsatz

$$2 \langle T \rangle_t = - \langle V \rangle_t \quad (\text{II.29a})$$

und die Gesamtenergie

$$E = \frac{1}{2} \langle V \rangle_t = - \langle T \rangle_t. \quad (\text{II.29b})$$

Da $\langle T \rangle_t$ genau wie T positiv ist, gilt $E < 0$ — wie wir im Fall $N = 2$ in § II.2.3 wieder finden werden.

⁽¹⁹⁾... unter der Annahme, dass relativistische Effekte keine Rolle spielen.

Ein zweites Beispiel ist das von gekoppelten harmonischen Oszillatoren, d.h. mit quadratischen Potentialen $V_{ab} \propto (\vec{r}_a - \vec{r}_b)^2$, entsprechend dem Fall $n = 2$. Dann kommen

$$\langle T \rangle_t = \langle V \rangle_t \quad \text{und} \quad E = 2 \langle T \rangle_t. \quad (\text{II.30})$$

Da dieses Problem (zumindest prinzipiell) analytisch lösbar ist, indem die Oszillatoren entkoppelt werden können (vgl. § IV.1.2), kann dieses Ergebnis direkt wiedergefunden werden. Eigentlich gelten nämlich die zwei Gleichungen schon für einen einzelnen harmonischen Oszillator.

II.1.5 Schwerpunktsystem

Sei Σ ein Mehrteilchensystem. Als *Schwerpunktsystem* wird ein Bezugssystem \mathcal{B}_Σ^* bezeichnet, in welchem der Schwerpunkt von Σ ruht, und welches gegenüber einem Inertialsystem \mathcal{B}_I nicht rotiert. Um beide Bedingungen zu erfüllen, nimmt man oft ein Koordinatensystem in \mathcal{B}_Σ^* mit dem Schwerpunkt von Σ als Ursprungspunkt und den Achsen parallel zu denen eines Koordinatensystems von \mathcal{B}_I .

Hiernach werden die bezüglich \mathcal{B}_Σ^* gemessenen Größen mit Sternchen gekennzeichnet: Ortsvektor \vec{r}^* , Teilchenbahnkurven $\vec{x}_a^*(t)$, Impulse $\vec{p}_a^*(t)$, usw.

Bemerkung: Das Schwerpunktsystem ist nicht unbedingt ein Inertialsystem! Dies gilt nur bei abgeschlossenen Mehrteilchensystemen, die keiner resultierenden äußeren Kraft unterliegen. In solchen Fällen wird der Schwerpunkt laut der Bewegungsgleichung (II.12a) bzw. dem Schwerpunktsatz (II.12b) nicht beschleunigt.

Aus Gl. (II.5), abgeleitet nach der Zeit, und der Tatsache, dass der Schwerpunkt bezüglich \mathcal{B}_Σ^* ruht, folgt sofort für den Gesamtimpuls des Mehrteilchensystems

$$\vec{P}^*(t) \equiv \sum_{a=1}^N \vec{p}_a^*(t) = \vec{0}. \quad (\text{II.31})$$

Diese Beziehung kann auch als Definition des Schwerpunktsystems — das auf Englisch öfter *center-of-momentum frame* genannt wird, wobei „momentum“ = Impuls — angesehen werden.

Wird das Koordinatensystem von \mathcal{B}_Σ^* so gewählt, dass der Schwerpunkt von Σ im Nullpunkt sitzt, so gilt automatisch $\vec{X}^*(t) = 0$ zu jeder Zeit t .

Der Gesamtdrehimpuls \vec{L}^* eines Mehrteilchensystems Σ bezüglich des zugehörigen Schwerpunktsystems⁽²⁰⁾

$$\vec{L}^*(t) \equiv \sum_{a=1}^N \vec{x}_a^*(t) \times \vec{p}_a^*(t) \quad (\text{II.32})$$

wird *Eigendrehimpuls* des Systems Σ genannt. Wiederum wird die gesamte kinetische Energie bezüglich \mathcal{B}_Σ^*

$$T^* \equiv \sum_{a=1}^N \frac{\vec{p}_a^{*2}}{2m_a} \quad (\text{II.33})$$

als *innere kinetische Energie* des Systems bezeichnet.

Sei jetzt \mathcal{B} ein beliebiges Bezugssystem mit der einzigen Einschränkung, dass es gegenüber dem Schwerpunktsystem \mathcal{B}_Σ^* nicht rotiert. Es gelten dann zwei nach König^(k) genannte Ergebnisse, und zwar

⁽²⁰⁾... noch besser, relativ zum eigenen Schwerpunkt, indem der letztere als Ursprungspunkt genommen wird.

^(k)J. S. KÖNIG, 1712–1757

Theorem (1. Satz von König): Der Gesamtdrehimpuls \vec{L} des Systems Σ relativ zu einem festen Punkt in \mathcal{B} , z.B. dem Nullpunkt, ist die Summe aus dem Eigendrehimpuls \vec{L}^* des Systems und dem Drehimpuls bezüglich \mathcal{B} des Schwerpunkts von Σ , versehen mit der Gesamtmasse M :

$$\vec{L}(t) = \vec{L}^*(t) + \vec{X}(t) \times \vec{P}(t) \quad \text{mit} \quad \vec{P}(t) = M \frac{d\vec{X}(t)}{dt}, \quad (\text{II.34})$$

wobei $\vec{X}(t)$ die Bahnkurve des Schwerpunkts von Σ relativ zu \mathcal{B} bezeichnet.

Theorem (2. Satz von König): Die gesamte kinetische Energie T des Systems Σ relativ zu \mathcal{B} ist die Summe aus der inneren kinetischen Energie T^* des Systems und der kinetischen Energie bezüglich \mathcal{B} des Schwerpunkts von Σ , versehen mit der Gesamtmasse M :

$$T(t) = T^*(t) + \frac{\vec{P}(t)^2}{2M}. \quad (\text{II.35})$$

Beweis: **to do!**

Diese Sätze bedeuten, dass die Bewegung eines System aus N Körpern in zwei unabhängige Anteile zerlegt werden kann, und zwar in die „innere“ Bewegung der Teilchen relativ zum Schwerpunkt des Systems und die „globale“ (Translations-)Bewegung des Schwerpunkts durch den Raum. Dementsprechend finden theoretische Untersuchungen des Systems meistens in seinem Schwerpunktsystem statt.

II.2 Zwei-Körper-Systeme

Der Formalismus des vorigen Abschnitts kann natürlich auf den Fall eines Systems aus zwei Körpern angewandt werden. In diesem Fall treten aber einige wichtigen Vereinfachungen aus, beginnend bei den Bewegungsgleichungen, die sich oft entkoppeln (§ II.2.1). Somit können verschiedene Probleme ziemlich ausführlich behandelt werden: harmonisch gekoppelte Massen (§ II.2.2), Körper in newtonscher Gravitationswechselwirkung (§ II.2.3), oder die Streuung zweier Teilchen (§ II.2.4).

II.2.1 Separation der Bewegungsgleichungen

II.2.1 a Variablen

Genau wie in § II.1.1 b wird die Schwerpunktkoordinate definiert durch

$$\vec{X}(t) \equiv \frac{m_1 \vec{x}_1(t) + m_2 \vec{x}_2(t)}{m_1 + m_2}. \quad (\text{II.36a})$$

Da es nur zwei Körper gibt, kann man auch eine *Relativkoordinate* einführen:

$$\vec{x}(t) \equiv \vec{x}_1(t) - \vec{x}_2(t). \quad (\text{II.36b})$$

Bei dem Übergang von den Variablen $\vec{x}_1(t)$, $\vec{x}_2(t)$ zu den neuen Variablen $\vec{X}(t)$, $\vec{x}(t)$ handelt es sich um eine lineare Transformation. Die entsprechende Rücktransformation lautet

$$\vec{x}_1(t) = \vec{X}(t) + \frac{m_2}{m_1 + m_2} \vec{x}(t), \quad \vec{x}_2(t) = \vec{X}(t) - \frac{m_1}{m_1 + m_2} \vec{x}(t), \quad (\text{II.36c})$$

wie sich leicht nachprüfen lässt.

Für die Beschreibung der Bewegung wird es sich lohnen, die *reduzierte Masse*

$$\mu \equiv \frac{m_1 m_2}{m_1 + m_2} \quad (\text{II.37})$$

einzuführen.

II.2.1 b Bewegungsgleichungen

Aus den allgemeinen Bewegungsgleichungen [vgl. Gl. (II.2)]

$$\begin{aligned} m_1 \frac{d^2 \vec{x}_1(t)}{dt^2} &= \vec{F}_{1,\text{ext}} + \vec{F}_{2 \rightarrow 1}, \\ m_2 \frac{d^2 \vec{x}_2(t)}{dt^2} &= \vec{F}_{2,\text{ext}} + \vec{F}_{1 \rightarrow 2} \end{aligned}$$

lassen sich Bewegungsgleichungen für $\vec{X}(t)$ und $\vec{x}(t)$ herleiten.

Erstens liefert eine Summe die Bewegungsgleichung des Schwerpunkts, und zwar

$$(m_1 + m_2) \frac{d^2 \vec{X}(t)}{dt^2} = \vec{F}_{1,\text{ext}} + \vec{F}_{2,\text{ext}} \quad (\text{II.38a})$$

unabhängig von den inneren Kräften, in Übereinstimmung mit dem Schwerpunktsatz (II.12).

Andererseits ergibt sich nach Division jeder Gleichung durch die darin auftretende Masse und anschließender Subtraktion

$$\frac{d^2 \vec{x}(t)}{dt^2} = \frac{\vec{F}_{1,\text{ext}}}{m_1} - \frac{\vec{F}_{2,\text{ext}}}{m_2} + \frac{\vec{F}_{2 \rightarrow 1}}{m_1} - \frac{\vec{F}_{1 \rightarrow 2}}{m_2},$$

d.h. unter Verwendung des dritten newtonschen Gesetzes $\vec{F}_{1 \rightarrow 2} = -\vec{F}_{2 \rightarrow 1}$ und der reduzierten Masse (II.37)

$$\frac{d^2 \vec{x}(t)}{dt^2} = \frac{\vec{F}_{1,\text{ext}}}{m_1} - \frac{\vec{F}_{2,\text{ext}}}{m_2} + \frac{\vec{F}_{2 \rightarrow 1}}{\mu}. \quad (\text{II.38b})$$

Im Allgemeinen hängt diese Gleichung noch von den äußeren Kräften ab.

Sei jetzt angenommen, dass das Zwei-Körper-System abgeschlossen ist, $\vec{F}_{a,\text{ext}} = \vec{0}$ für $a = 1, 2$. Dann vereinfachen sich die Bewegungsgleichungen (II.38) zu

$$\frac{d^2 \vec{X}(t)}{dt^2} = \vec{0}, \quad (\text{II.39a})$$

entsprechend der Erhaltung des Gesamtimpulses (II.12c) für ein System ohne äußere Kräfte, und

$$\mu \frac{d^2 \vec{x}(t)}{dt^2} = \vec{F}_{2 \rightarrow 1}. \quad (\text{II.39b})$$

Dabei hängt die Kraft auf der rechten Seite a priori noch von den Positionen $\vec{x}_1(t)$, $\vec{x}_2(t)$ der beiden Körper — d.h. äquivalent von $\vec{X}(t)$ und $\vec{x}(t)$ — und von ihren Geschwindigkeiten ab. Falls die Zweikörperkraft $\vec{F}_{2 \rightarrow 1}$ nur Funktion von der Relativkoordinate $\vec{x}(t)$ und von ihrer Zeitableitung ist, dann ist Gl. (II.39b) die Bewegungsgleichung eines einzelnen fiktiven Massenpunkts mit der reduzierten Masse μ . Somit wird das ursprüngliche Zwei-Körper-Problem in zwei entkoppelte einfachere Ein-Körper-Probleme transformiert, und zwar einerseits für den Schwerpunkt, andererseits für den fiktiven Massenpunkt.

Ist die Kraft $\vec{F}_{2 \rightarrow 1}$ nun zentral, d.h. parallel zu $\vec{x}_1(t) - \vec{x}_2(t)$, dann bewegt sich dieser Massenpunkt in einem Zentralkraftfeld mit Kraftzentrum in $\vec{x}_2(t)$. In diesem Fall ist der Drehimpuls des fiktiven Teilchens

$$\vec{\ell} = \vec{x}(t) \times \mu \frac{d\vec{x}(t)}{dt}$$

eine Erhaltungsgröße, und man spricht von einem Zentralkraftproblem.

Die Erhaltung von $\vec{\ell}$ lässt sich leicht beweisen: aus der Produktregel folgt

$$\frac{d\vec{\ell}}{dt} = \mu \frac{d\vec{x}(t)}{dt} \times \frac{d\vec{x}(t)}{dt} + \vec{x}(t) \times \mu \frac{d^2 \vec{x}(t)}{dt^2}.$$

Der erste Term ist offensichtlich gleich $\vec{0}$, während der zweite dank Gl. (II.39b) das Kreuzprodukt $\vec{x}(t) \times \vec{F}_{2 \rightarrow 1}$ aus zwei parallelen Vektoren, also ebenfalls gleich $\vec{0}$, ist. \square

In den folgenden § II.2.2–II.2.4 werden wir Beispiele solcher Zentralkraftprobleme sehen, im Fall *konservativer* Kraftfelder, $\vec{F}_{2 \rightarrow 1}(\vec{r}_1 - \vec{r}_2) \equiv \vec{F}(\vec{r}) = -\vec{\nabla}V(r)$.

Bemerkung: Wenn eine der Massen viel größer als die andere ist, z.B. $m_2 \gg m_1$, dann gelten $\vec{X}(t) \simeq \vec{x}_2(t)$ — d.h. die Position des Schwerpunkts stimmt fast mit jener des schweren Partners überein —, $\vec{x}_1(t) \simeq \vec{X}(t) + \vec{x}(t)$ und $\mu \simeq m_1$ — d.h. die Eigenschaften des fiktiven Teilchens sind ähnlich denen des leichten Körpers.

II.2.2 Gekoppelte Punktmassen

Betrachten wir als erstes Beispiel zwei Massen m_1, m_2 , die über eine Feder gekoppelt sind. Es wird angenommen, dass die Bewegung der Massen eindimensional entlang der horizontalen x -Achse bleibt, und dass es keine äußere Kraft gibt.

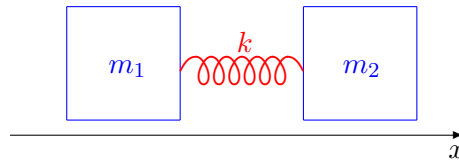


Abbildung II.1 – Gekoppelte Massen

Die Feder wird als ein *harmonischer Oszillator* modelliert, d.h. die dadurch verursachte Kraft ist proportional zur Auslenkung aus einer Ruhelänge. Somit lautet die Kraft auf Masse 1

$$\vec{F}_{2 \rightarrow 1} = k(x_2 - x_1 - \ell_0) \vec{e}_x \tag{II.40}$$

mit der Federkonstante k , der Ruhelänge ℓ_0 der Feder, und den Positionen x_1, x_2 der Massen — die als Punktmassen zu betrachten sind.

Die Bewegungsgleichungen für $x_1(t), x_2(t)$ lauten

$$m_1 \frac{d^2 x_1(t)}{dt^2} = k[x_2(t) - x_1(t) - \ell_0], \quad m_2 \frac{d^2 x_2(t)}{dt^2} = -k[x_2(t) - x_1(t) - \ell_0], \tag{II.41}$$

wobei das dritte newtonsche Gesetz benutzt wurde, um die Kraft $\vec{F}_{1 \rightarrow 2}$ auszudrücken. Das Addieren der beiden Gleichungen führt zur Zeitunabhängigkeit der Schwerpunktkoordinate $X(t)$ — wie zu erwarten war!

Aus den Gl. (II.41) folgt noch

$$\mu \frac{d^2}{dt^2} [x_1(t) - x_2(t)] = k[x_2(t) - x_1(t) - \ell_0],$$

d.h., für die Relativkoordinate $x(t) \equiv x_1(t) - x_2(t)$

$$\mu \frac{d^2 x(t)}{dt^2} = -k[x(t) + \ell_0]. \tag{II.42}$$

Das ist die Bewegungsgleichung für einen einzelnen Massenpunkt mit der reduzierten Masse μ , der an einem Kraftzentrum harmonisch gebunden ist.

Die Lösung ist dann der Form $x(t) = -\ell_0 + A \cos(\omega t + \varphi)$ mit $\omega \equiv \sqrt{k/\mu}$ und zwei Konstanten A, φ , die durch Anfangsbedingungen festgestellt sind. Daraus erhält man über die Rücktransformation (II.36c) die Zeit-Ort-Funktionen $x_1(t)$ und $x_2(t)$.

II.2.3 Kepler-Problem

Die newtonsche Gravitationskraft zwischen zwei Massenpunkten mit Massen m_1, m_2 ist eine konservative Zentralkraft, gegeben durch

$$\vec{F}(|\vec{x}|) = -\frac{G_N m_1 m_2}{|\vec{x}|^2} \frac{\vec{x}}{|\vec{x}|} = -\vec{\nabla}V(|\vec{x}|) \tag{II.43a}$$

mit dem Potential

$$V(|\vec{x}|) = -\frac{G_N m_1 m_2}{|\vec{x}|}. \quad (\text{II.43b})$$

Dabei bezeichnet \vec{x} der Abstandsvektor zwischen den Massen und G_N die newtonsche Gravitationskonstante, $G_N \simeq 6,674 \cdot 10^{-11} \text{ m}^3 \cdot \text{kg}^{-1} \cdot \text{s}^{-2}$.

Ziel dieses Abschnitt ist die Untersuchung der einfachsten möglichen Bewegungen, zu denen dieses Potential führen kann, und insbesondere die Herleitung der drei durch Kepler⁽¹⁾ empirisch ermittelten Gesetze. Zu diesem Zweck wird die Bewegung zweier Massenpunkte betrachtet, die über das Potential (II.43b) miteinander wechselwirken, während sie keiner äußeren Kraft unterliegen. Diese Fragestellung wird als *Kepler-Problem* bezeichnet.

Bemerkung: Das Gravitationspotential ist nur bis auf eine additive Konstante definiert, die bei der Gradientenbildung nicht beiträgt. In Gl. (II.43b) wurde diese Konstante so gewählt, dass das Potential für unendlich große Abstände, also wenn die Kraft Null ist, verschwindet.

II.2.3 a Vorbereitungen

In Übereinstimmung mit den allgemeinen Ergebnissen des § II.2.1 kann man die triviale Bewegung des Schwerpunkts von der mehr interessanten Relativbewegung separieren, wobei die letztere die Zeitänderung der Relativkoordinate \vec{x} für ein fiktives Teilchen mit der reduzierten Masse μ ist.

Da es sich bei dem newtonschen Potential (II.43b) um ein Zentralpotential handelt, ist der Drehimpuls $\vec{\ell}$ für die Relativbewegung erhalten. Insbesondere bleibt seine Richtung konstant. Da sowohl $\vec{x}(t)$ als $\dot{\vec{x}}(t)$ senkrecht auf $\vec{\ell}$ sind, bleibt die Bahnkurve in einer Ebene senkrecht auf $\vec{\ell}$ — der sog. Ekliptik. Dementsprechend lohnt es sich, ein Koordinatensystem zu wählen, in dem die Richtung von $\vec{\ell}$ entlang einer der Achsen ist, z.B. der z -Achse.

Die Wahl der Koordinaten in der Bewegungsebene bleibt noch offen. Man kann zuerst feste kartesische Basisvektoren \vec{e}_x , \vec{e}_y einführen, und dementsprechend die Bahnkurve durch $(x(t), y(t))$ parametrisieren. Da das Wechselwirkungspotential (II.43b) nur vom Abstand $r(t) \equiv |\vec{x}(t)|$ zum Kraftzentrum abhängt, ist es sinnvoll, eine Parametrisierung der Bahnkurve zu wählen, in dem dieser Abstand eine direkte Rolle spielt. Dazu dient noch der Winkel $\theta(t)$ der Relativkoordinate relativ zu einer festen Richtung in der Ebene, z.B. zur x -Achse, zur Charakterisierung von $\vec{x}(t)$.

Somit lautet die Relativkoordinate

$$\vec{x}(t) = r(t) \vec{e}_r(t), \quad (\text{II.44})$$

wobei die instantane Richtung des Einheitsvektors $\vec{e}_r(t)$ in Radialrichtung von der Position des (fiktiven) bewegten Körpers und damit indirekt von der Zeit abhängt: bezüglich der Basisvektoren eines festen kartesischen Koordinatensystems gilt

$$\vec{e}_r(t) = \cos \theta(t) \vec{e}_x + \sin \theta(t) \vec{e}_y. \quad (\text{II.45a})$$

Man bezeichnet mit $\vec{e}_\theta(t)$ einen orthogonal auf $\vec{e}_r(t)$ stehenden Einheitsvektor in der Bewegungsebene (s. Abb. II.2)

$$\vec{e}_\theta(t) = -\sin \theta(t) \vec{e}_x + \cos \theta(t) \vec{e}_y, \quad (\text{II.45b})$$

so dass $\vec{e}_r(t)$, $\vec{e}_\theta(t)$ und $\vec{e}_z(t) \equiv \vec{e}_r(t) \times \vec{e}_\theta(t)$ die Basisvektoren eines *mitbewegten Koordinatensystems* bilden, wobei $\vec{e}_z(t)$ hier eigentlich zeitunabhängig ist.

Die Ableitung der Gl. (II.45a) nach der Zeit gibt⁽²¹⁾

$$\frac{d\vec{e}_r(t)}{dt} = -\frac{d\theta(t)}{dt} \sin \theta(t) \vec{e}_x + \frac{d\theta(t)}{dt} \cos \theta(t) \vec{e}_y = \dot{\theta}(t) \vec{e}_\theta(t). \quad (\text{II.46a})$$

⁽²¹⁾Auf ähnliche Weise gilt

$$\frac{d\vec{e}_\theta(t)}{dt} = -\frac{d\theta(t)}{dt} \cos \theta(t) \vec{e}_x - \frac{d\theta(t)}{dt} \sin \theta(t) \vec{e}_y = -\dot{\theta}(t) \vec{e}_r(t). \quad (\text{II.46b})$$

⁽¹⁾J. KEPLER, 1571–1630

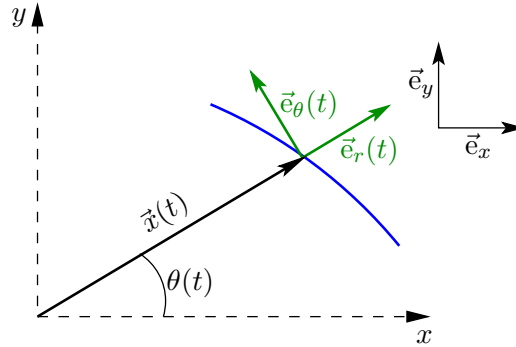


Abbildung II.2

Somit lautet die Relativgeschwindigkeit, unter Verwendung der Kettenregel beim Ableiten der Gl. (II.44)

$$\frac{d\vec{x}(t)}{dt} = \frac{dr(t)}{dt} \vec{e}_r(t) + r(t) \frac{d\vec{e}_r(t)}{dt} = \dot{r}(t) \vec{e}_r(t) + r(t) \dot{\theta}(t) \vec{e}_\theta(t). \quad (II.47)$$

Schließlich ist der Drehimpuls der Relativbewegung gegeben durch

$$\vec{\ell}(t) \equiv \vec{x}(t) \times \mu \dot{\vec{x}}(t) = \mu r(t)^2 \dot{\theta}(t) \vec{e}_z \quad (II.48)$$

mit konstanter Richtung und der konstanten Komponente

$$\ell \equiv \mu r(t)^2 \dot{\theta}(t) \quad (II.49)$$

entlang \vec{e}_z .

II.2.3b Zweites Keplersches Gesetz

Im Intervall zwischen zwei sukzessiven Zeitpunkten t und $t + dt$ ändert sich der Polarwinkel $\theta(t)$ der Bahnkurve um $d\theta \simeq \dot{\theta}(t) dt$. Die Ortsvektoren $\vec{x}(t)$ und $\vec{x}(t + dt)$ und die Bahnkurve grenzen ein Flächenelement ab, dessen Flächeninhalt annähernd durch

$$dS = \frac{1}{2} r(t) r(t) d\theta = \frac{r(t)^2}{2} \frac{d\theta(t)}{dt} dt$$

gegeben ist.

Ersetzt man die Zeitableitung $\dot{\theta}(t)$ mithilfe der Gl. (II.49), so kommt der „Flächensatz“

$$\frac{dS}{dt} = \frac{\ell}{2\mu} = \text{Konstante}, \quad (II.50)$$

d.h. (zweites Keplersches Gesetz)

Der Ortsvektor (von dem Stern zum Planeten) überstreicht in gleichen Zeiten gleich große Flächen. (II.51)

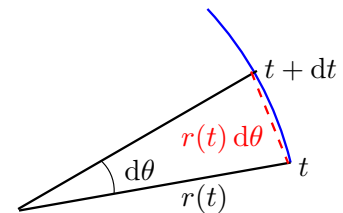


Abbildung II.3

Dabei beziehen sich die Worte zwischen Klammern auf den historischen durch Kepler betrachteten Fall eines Planeten um einen viel schwereren Stern, $m_2 \gg m_1$. Im Fall zweier (ungefähr) gleich schweren Körper, wie z.B. die zwei Sterne eines Doppelsternsystems, geht die Relativkoordinate von einem Stern zum anderen.

Bemerkung: In der Herleitung dieses Ergebnisses wurde die genau Form des Potentials $V(r)$ nicht benutzt, d.h. der Flächensatz gilt für jedes Zentralkraftproblem.

II.2.3c Bahnkurve. Erstes und drittes Keplersches Gesetz

Da die newtonsche Gravitationskraft konservativ ist, ist die Gesamtenergie

$$E = T(t) + V(t) = \frac{1}{2} \mu \left[\frac{d\vec{x}(t)}{dt} \right]^2 + V(r(t))$$

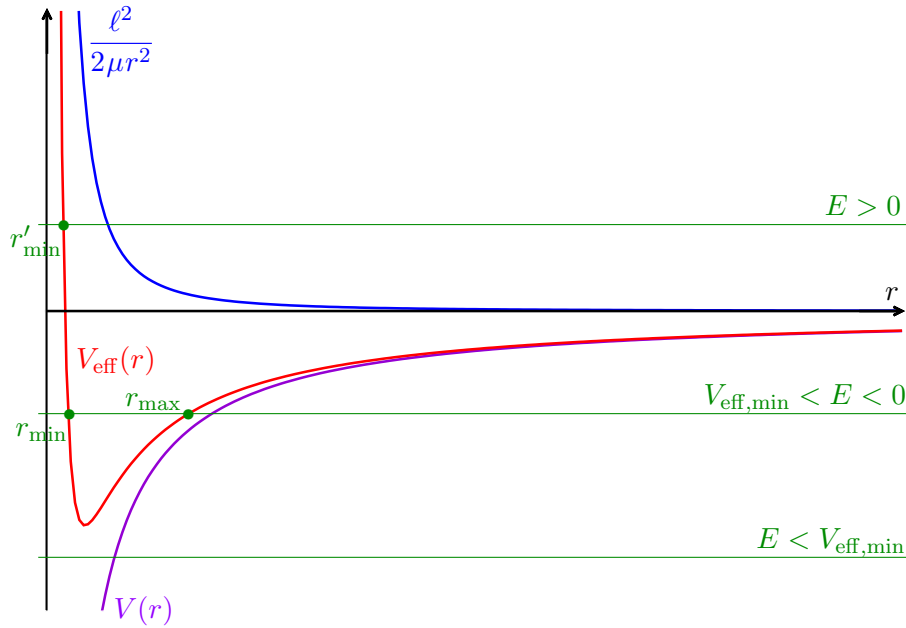


Abbildung II.4 – Effektives Potential $V_{\text{eff}}(r)$ für das Kepler-Problem.

der Relativbewegung eine Erhaltungsgröße. Unter Verwendung der expliziten Form (II.47) der Geschwindigkeit gilt noch

$$E = \frac{\mu}{2} \left[\dot{r}(t)^2 + r(t)^2 \dot{\theta}(t)^2 \right] + V(r(t)).$$

Mithilfe der Gl. (II.49) kann man die Zeitableitung $\dot{\theta}(t)$ durch $\ell/\mu r(t)^2$ ersetzen, so dass sich die totale Energie noch als

$$E = \frac{\mu}{2} \dot{r}(t)^2 + \frac{\ell^2}{2\mu r(t)^2} + V(r(t)) \quad (\text{II.52})$$

umschreiben lässt, unabhängig von der Winkelvariablen θ . Dabei ist $\dot{r}(t)$ die radiale Geschwindigkeit. Betrachtet man dementsprechend Gl. (II.52) als totale Energie für die eindimensionale Bewegung in Radialrichtung, so ist der erste Term auf der rechten Seite die zugehörige kinetische Energie, während die zwei letzten ein *effektives Potential*

$$V_{\text{eff}}(r) \equiv V(r) + \frac{\ell^2}{2\mu r^2} \quad (\text{II.53})$$

darstellen. Dann gilt

$$\dot{r}(t)^2 = \frac{2}{\mu} [E - V_{\text{eff}}(r)]. \quad (\text{II.54})$$

Beim newtonschen Gravitationspotential $V(r) = -\alpha/r$ mit $\alpha = G_{\text{N}} m_1 m_2 > 0$ ist das resultierende effektive Potential eine nicht-monotone Funktion von r , erstmal abnehmend dann wachsend. Je nach den relativen Werten von der Gesamtenergie E und dem minimalen Wert $V_{\text{eff},\text{min}} < 0$ des effektiven Potentials kommen drei verschiedene Möglichkeiten vor, die in Abb. II.4 graphisch gezeigt werden:

- Für $E < V_{\text{eff},\text{min}}$ ist die rechte Seite von Gl. (II.54) immer negativ, so dass keine Lösung der Gleichung mit reeller Radialgeschwindigkeit existieren kann: diesen mathematischen Werten von E und ℓ entspricht keine physikalisch erlaubte Lösung des Kepler-Problems.
- Für $V_{\text{eff},\text{min}} < E < 0$ kann die rechte Seite von Gl. (II.54) nicht-negativ sein, und zwar für Werte $r_{\text{min}} \leq r \leq r_{\text{max}}$, wobei r_{min} , r_{max} die zwei Lösungen der Gleichung $V_{\text{eff}}(r) = E$ sind. Da

r_{\max} endlich ist, bleibt das fiktive Teilchen immer in endlichem Abstand vom Kraftzentrum, und ist somit daran „gebunden“.

Bei den Abständen r_{\min} und r_{\max} verschwindet die Radialgeschwindigkeit \dot{r} — die eigentlich ihr Vorzeichen ändern wird: dabei handelt es sich um *Umkehrpunkte* der Radialbewegung.

Im Spezialfall $E = V_{\text{eff},\min}$ ist $V(r) = E$ für einen einzigen Wert von r . Dazu gibt Gl. (II.54) $\dot{r}(t) = 0$, d.h. r bleibt konstant: die entsprechende Bahnkurve ist ein Kreis.

- Für $E \geq 0$ hat die Gleichung $V_{\text{eff}}(r) = E$ eine einzige Lösung r'_{\min} , und die rechte Seite von Gl. (II.54) bleibt positiv für $r > r'_{\min}$. Somit geht die Bahnkurve des fiktiven Teilchens unendlich weit weg vom Kraftzentrum: das Teilchen ist „ungebunden“.

Bemerkung: Für kleine Abstände r vom Kraftzentrum dominiert der Beitrag $\ell^2/2\mu r^2$ im effektiven Potential gegenüber dem Gravitationspotential. Da dieser Term positiv ist, führt er zu einer Abstoßung vom Kraftzentrum weg. Dementsprechend wird der Term $\ell^2/2\mu r^2$ oft als *Zentrifugalpotential* bezeichnet.

Für die obige Fallunterscheidung war die genaue Form des effektiven Potentials, und daher von $V(r)$, nicht nötig, um die Existenz oder nicht-Existenz von Lösungen festzustellen. Die exakten Bahnkurven für die verschiedenen Fälle hängen aber natürlich von $V(r)$ ab.

Ausgehend aus Gl. (II.54) kann man den Verlauf der Bahnkurve bestimmen.⁽²²⁾ Einerseits gilt unter Verwendung der Kettenregel

$$\frac{dr(t)}{dt} = \frac{dr(\theta)}{d\theta} \frac{d\theta(t)}{dt} = \frac{dr(\theta)}{d\theta} \frac{\ell}{\mu r^2},$$

wobei Gl. (II.49) benutzt wurde. Andererseits ist $\dot{r}(t)$ auch durch Gl. (II.54) gegeben. Betrachtet man eine Lösung mit positiver Radialgeschwindigkeit, so kann man schreiben

$$\frac{dr(\theta)}{d\theta} = \frac{\mu r^2}{\ell} \dot{r}(t) = \frac{r^2}{\ell} \sqrt{2\mu[E - V_{\text{eff}}(r)]}.$$

Das heißt, nach Trennung der Variablen

$$d\theta = \frac{\ell dr}{r^2 \sqrt{2\mu[E - V_{\text{eff}}(r)]}} = \frac{\ell dr}{r^2 \sqrt{2\mu[E - V(r)] - \ell^2/r^2}} \quad (\text{II.55})$$

und nach formeller Integration beider Seiten dieser Gleichung

$$\theta = \int^r \frac{\ell dr'}{r'^2 \sqrt{2\mu[E - V(r')] - \ell^2/r'^2}}, \quad (\text{II.56})$$

wobei $r_{\min} < r < r_{\max}$ gelten muss.

Im Fall des newtonschen Potentials $V(r) = -\alpha/r$ mit $\alpha = G_N m_1 m_2$ kann die Integration explizit durchgeführt werden. Dabei ist

$$\theta = \int^r \frac{\ell dr'}{r'^2 \sqrt{2\mu(E + \alpha/r') - \ell^2/r'^2}}.$$

Eine erste Substitution $u' = 1/r'$, d.h. $du' = -dr'/r'^2$, gibt zunächst

$$\theta = - \int^u \frac{\ell du'}{\sqrt{2\mu(E + \alpha u') - \ell^2 u'^2}} = - \int^u \frac{du'}{\sqrt{2\mu E/\ell^2 + (2\mu\alpha/\ell^2)u' - u'^2}}.$$

Der Term unter der Wurzel im Nenner ist der Form $a + bu' - u'^2 = a + b^2/4 - (u' - b/2)^2$, mit $a = 2\mu E/\ell^2$ und $b = 2\mu\alpha/\ell^2$. Eine weitere Substitution $v' = (u' - b/2)/\sqrt{a + b^2/4}$ liefert dann

⁽²²⁾Eine alternative Herleitung wird im Anhang II.A zu diesem Kapitel vorgestellt.

$$\int^u \frac{du'}{\sqrt{a + b^2/4 - (u' - b/2)^2}} = \int^v \frac{dv'}{\sqrt{1 - v'^2}} = -\arccos(v) = -\arccos\left(\frac{u - b/2}{\sqrt{a + b^2/4}}\right).$$

Somit erhält man

$$\theta = \arccos\left(\frac{1/r - \mu\alpha/\ell^2}{\sqrt{2\mu E/\ell^2 + \mu^2\alpha^2/\ell^4}}\right) + \text{Konstante} \quad (\text{II.57})$$

Die Wahl der Integrationskonstanten ist äquivalent zur Wahl der Bezugsrichtung $\theta = 0$, so dass die Konstante problemlos gleich Null gesetzt werden kann. Damit ergibt sich

$$\cos\theta = \frac{1/r - \mu\alpha/\ell^2}{\sqrt{2\mu E/\ell^2 + \mu^2\alpha^2/\ell^4}} = \frac{1/r - \mu\alpha/\ell^2}{(\mu\alpha/\ell^2)\sqrt{2E\ell^2/\mu\alpha^2 + 1}}.$$

Mit den Definitionen

$$\epsilon \equiv \sqrt{1 + \frac{2E\ell^2}{\mu\alpha^2}} \quad \text{und} \quad p \equiv \frac{\ell^2}{\mu\alpha} \quad (\text{II.58a})$$

wird die θ - r -Abhängigkeit zu $\epsilon \cos\theta = p/r - 1$, d.h.

$$r = \frac{p}{1 + \epsilon \cos\theta}. \quad (\text{II.58b})$$

Das ist die Gleichung in Polarkoordinaten eines Kegelschnitts mit Exzentrizität ϵ und Parameter p .

Bemerkung: Man prüft einfach, dass die Wahl einer Lösung (genauer, eines Zweigs der Lösung) mit negativer Radialgeschwindigkeit $\dot{r}(t)$, entsprechend der negativen Wurzel der Gl. (II.54), führt zur gleichen Bahnkurve: $dr(\theta)/d\theta > 0$ bzw. $dr(\theta)/d\theta < 0$ entspricht dem Teil der Kurve mit $0 < \theta < \pi$ bzw. $\pi < \theta < 2\pi$.

Wie in der Diskussion der Abb. II.4 gibt es unterschiedliche Fälle je nach dem Wert der Energie E . Dafür kann man schon bemerken, dass das effektive Potential $-\alpha/r + \ell^2/2\mu r^2$ minimal für $r_0 = \ell^2/\mu\alpha$ ist, wo V_{eff} den Wert $-\mu\alpha^2/2\ell^2 \equiv V_{\text{eff,min}}$ annimmt.

- Für $V_{\text{eff,min}} \leq E < 0$ gilt laut Gl. (II.58a) $0 \leq \epsilon < 1$; d.h., die Bahnkurve ist eine Ellipse — ein Kreis falls $\epsilon = 0$, entsprechend $E = V_{\text{eff,min}}$. Dies bildet das *erste Keplersche Gesetz*:

Die Planeten bewegen sich auf elliptischen Bahnen (mit der Sonne in einem gemeinsamen Brennpunkt). (II.59)

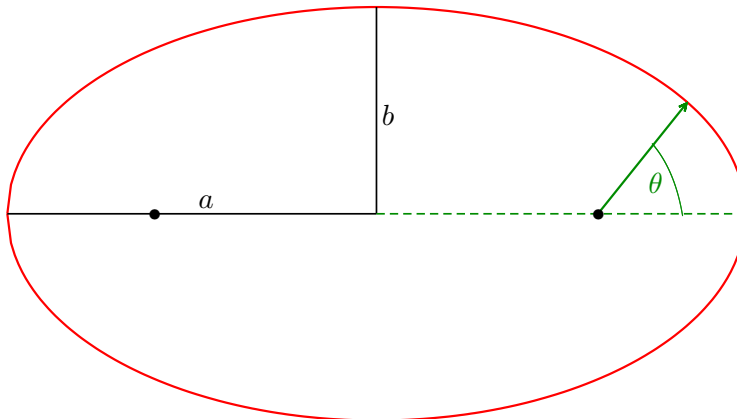


Abbildung II.5 – Ellipse mit Exzentrizität $\epsilon = 0,6$; die Punkte sind die Brennpunkte.

In einem Doppelsternsystem beschreibt jeder Stern eine elliptische Bahnkurve mit dem anderen Stern in einem Brennpunkt, während die Position ihres gemeinsamen Schwerpunkts unverändert bleibt.

Die große Halbachse der Ellipse ist $a = p/(1 - \epsilon^2)$, die kleine Halbachse $b = p/\sqrt{1 - \epsilon^2}$. Die durch die Ellipse abgeschlossene Fläche ist $\mathcal{S} = \pi ab$. Andererseits ist diese Fläche über den Flächensatz (II.50) mit der Umlaufzeit \mathcal{T} verknüpft: $\mathcal{S} = (\ell/2\mu)\mathcal{T}$. Somit gilt

$$\mathcal{T}^2 = \frac{4\mu^2}{\ell^2} \mathcal{S}^2 = \frac{4\pi^2\mu^2}{\ell^2} a^2 b^2 = \frac{4\pi^2\mu^2 p}{\ell^2} a^3 = \frac{4\pi^2\mu}{\alpha} a^3.$$

Mit den expliziten Ausdrücken der reduzierten Masse und der Konstanten $\alpha = G_N m_1 m_2$ ergibt sich schließlich

$$\mathcal{T}^2 = \frac{4\pi^2}{G_N(m_1 + m_2)} a^3.$$

Dies entspricht dem *dritten Kepler-Gesetz*:

Die Quadrate der Umlaufbahnen der Planeten sind proportional zur dritten Potenz der großen Bahnhalbachse.

(II.60)

Auf der Parametrisierung (II.58b) kann man sofort den bei $\theta = 0$ bzw. $\theta = \pi$ stattfindenden minimalen bzw. maximalen Abstand $r_{\min} = p/(1 + \epsilon)$ bzw. $r_{\max} = p/(1 - \epsilon)$ zwischen Planeten und Stern lesen.

- Für $E = 0$ ist die Exzentrizität der Bahnkurve gleich 1, so dass diese eine Parabel ist.
- Für $E > 0$ ist gemäß Gl. (II.58a) die Exzentrizität größer als 1, $\epsilon > 1$, d.h. die Bahnkurve ist eine Hyperbel (Abb. II.6).

Wie aus der Diskussion des effektiven Potentials erwartet war, ist die Bewegung räumlich unbegrenzt, und $r(\theta)$ wächst unendlich groß wenn $\cos \theta$ von oben gegen $-1/\epsilon$ geht. Dagegen wird für $\theta = 0$ der minimale Abstand zum Brennpunkt $r'_{\min} = p/(1 + \epsilon)$ erreicht.

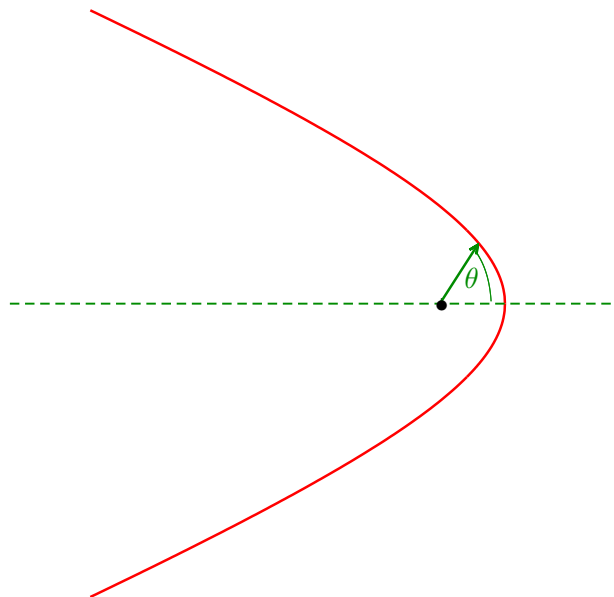


Abbildung II.6 – Hyperbel mit Exzentrizität $\epsilon = 1,6$ für ein anziehendes Potential.

Bisher wurde nur das anziehende newtonsche Potential (II.43b) betrachtet, mit einem positiven Zähler $\alpha = G_N m_1 m_2$ im Bruch. Man kann auch die Bahnkurven für das abstoßende Potential

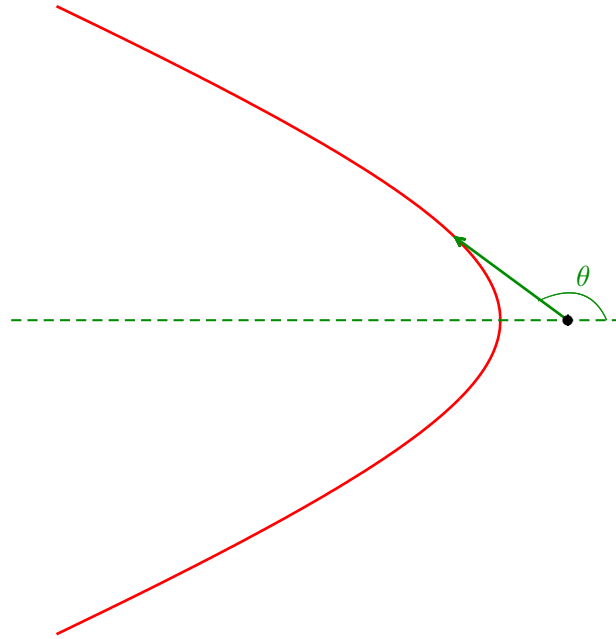


Abbildung II.7 – Hyperbel mit Exzentrizität $\epsilon = 1,6$ für ein abstoßendes Potential.

$V(r) = -\alpha/r$ mit $\alpha < 0$ betrachten, was insbesondere dem elektrostatischen Coulomb-Potential zwischen zwei positiven oder zwei negativen Punktladungen entspricht.

Wiederholt man die Herleitung zwischen Gl. (II.56) und (II.58) für den Fall $\alpha < 0$, so findet man, dass die möglichen Bahnkurven wieder Kegelschnitte sind. Dabei treten die abgeschlossenen Trajektorien (Ellipse, Kreis) aber nicht mehr auf, da es keine gebundenen Zustände des Zwei-Körper-Systems gibt — das effektive Potential $V_{\text{eff}}(r)$ ist immer positiv und nimmt monoton mit r ab —, sondern nur ungebundene. Da $V_{\text{eff}}(r) > 0$ für alle Abstände r , gibt es auch keine Bahnkurve mit $E = 0$, d.h. keine Parabel.

Dagegen gibt es Hyperbeln ($\epsilon > 1$), wobei man aufpassen soll, dass der Parameter p jetzt negativ ist. Der minimale Abstand ist somit $r_{\text{min}} = |p|/(\epsilon - 1)$ und wird für $\theta = \pi$ erreicht (Abb. II.7). Wiederum wird r unendlich groß, wenn $\cos \theta \rightarrow (-1/\epsilon)$ von unten geht.

II.2.4 Streuung

Die am Ende des § II.2.3 gefundenen unbegrenzten Bahnkurven sind solche von (fiktiven) Massenpunkten, die zur Zeit $t \rightarrow -\infty$ unendlich weit weg vom Kraftzentrum sind, sich daran annähern, und dann weg fliegen, so dass sie für $t \rightarrow \infty$ wieder unendlich entfernt vom Kraftzentrum sind. Ein solcher Prozess heißt auch *Stoß* oder *Streuung*.

Hiernach wird die Streuung eines Massenpunkts mit reduzierten Masse μ an einem *Streuzentrum* betrachtet. Diese Fragestellung entspricht den Bewegungsgleichungen, die aus der in § II.2.1 b dargestellten Separation folgen, und die in der Tat den Stoß eines physikalischen Massenpunkts an einem anderen modellieren. Der Prozess wird in einem Bezugssystem beschrieben, in welchem das Streuzentrum ruht, und das als inertial angenommen wird. Dazu wird ein Koordinatensystem gewählt, in dessen Nullpunkt O sich das Streuzentrum befindet.

In § II.2.4 a werden zunächst ein paar Größen eingeführt, die zur Kennzeichnung einer Streuung dienen. Der Zusammenhang zwischen einigen dieser Größen wird durch die Dynamik des Stoßes bestimmt, d.h. durch die für die Streuung relevante Wechselwirkung, wie in § II.2.4 b erklärt und an einem Beispiel illustriert wird. Falls der Streuprozess mehrmals stattfindet, jedoch mit (leicht) unterschiedlichen Anfangsbedingungen, werden die Wahrscheinlichkeiten für die möglichen Endzuständen durch Wirkungsquerschnitte charakterisiert (§ II.2.4 c).

II.2.4 a Geometrische Grundbegriffe der klassischen Streutheorie

Es wird angenommen, dass der Einfluss des Streuzentrums auf den gestreuten Massenpunkt vernachlässigbar für $t \rightarrow \mp\infty$ ist, d.h. lange vor oder nach dem Stoß. Das heißt wiederum, dass die Kraft zwischen den zwei physikalischen Massenpunkten für unendlich große Abstände verschwinden muss. Mathematisch muss das zugehörige Potential⁽²³⁾ $V(\vec{r})$ ebenfalls für große Abstände Null werden:

$$\lim_{|\vec{r}| \rightarrow \infty} V(\vec{r}) = 0. \tag{II.61}$$

Lange vor oder nach dem Stoß ist der gestreute Massenpunkt annähernd „frei“, d.h. er unterliegt keiner Kraft. Dann bewegt er sich laut dem Trägheitsgesetz (I.13) gleichförmig und geradlinig: sei \vec{p}_{ein} bzw. \vec{p}_{aus} sein Impuls für $t \rightarrow -\infty$ bzw. $t \rightarrow +\infty$, wobei das Kürzel *ein* bzw. *aus* für „einfallend“ bzw. „auslaufend“ steht. Wegen der Wechselwirkung mit dem Streuzentrum wird sich die Flugrichtung des gestreuten Massenpunkts im Allgemeinen ändern. Der Winkel zwischen den asymptotischen Flugrichtungen heißt *Streuwinkel* und wird hiernach mit ϑ bezeichnet, wie in Abb. II.8 zu sehen ist.⁽²⁴⁾

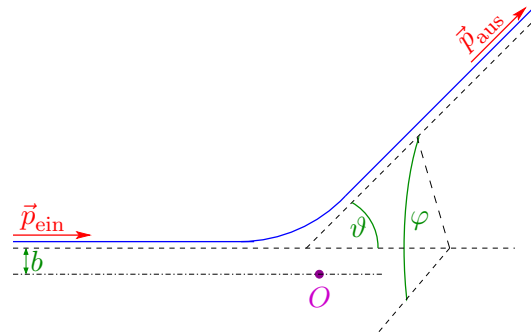


Abbildung II.8 – Charakteristische Größen zur Beschreibung einer Streuung.

Eine weitere geometrische Größe, die in klassischer Streutheorie einen bestimmten Streuprozess charakterisiert, ist der *Stoßparameter* b . Dabei handelt es sich um den minimalen Abstand des einfallenden Massenpunkts zum Streuzentrum für den fiktiven Fall, wo es sich ständig geradlinig bewegen würden, statt abgelenkt zu werden, d.h. bei ausgeschalteter Wechselwirkung, wie in Abb. II.8 illustriert wird.

Sei angenommen, dass die Flugrichtung des einfallenden Teilchens bekannt ist — was in einem Streuexperiment der Fall ist. In der *transversalen Ebene* senkrecht zu dieser Richtung kann man eine Bezugsrichtung festlegen, relativ zu dem man kann einen Winkel φ definieren. Somit wird die auslaufende Flugrichtung neben ϑ auch durch einen Azimutwinkel φ_{aus} charakterisiert. Wiederum kann der einfallenden Flugrichtung einen Winkel φ_{ein} assoziiert werden; stattdessen wird der entsprechende Azimutwinkel dem Stoßparameter zugeordnet, der somit zu einem (zweidimensionalen) Vektor \vec{b} , mit Betrag b und Winkel φ_{ein} , in der transversalen Ebene wird.

Bezeichnet man dann den Azimutwinkel der auslaufenden Flugrichtung mit φ anstatt φ_{aus} , so besteht der Zweck der klassischen Streutheorie darin, die Beziehung zwischen den Winkeln (ϑ, φ) und dem Stoßparameter \vec{b} festzulegen. Dabei charakterisieren die ersteren bzw. der letztere das auslaufende Teilchen, d.h. den Endzustand, bzw. den Anfangszustand, und zwar das einlaufende Teilchen. Somit werden im Allgemeinen Funktionen $(\vartheta(\vec{b}), \varphi(\vec{b}))$ gesucht.

Bemerkung: Im quantenmechanischen Rahmen verliert der Begriff des Stoßparameters, genau wie jener der Bahnkurve, an Bedeutung.

⁽²³⁾ ... unter der Annahme, dass die Kraft konservativ ist!

⁽²⁴⁾ Die übliche Bezeichnung des Streuwinkels ist eher θ ; hier wurde darauf verzichtet, um Verwechslung mit dem Polarwinkel zu vermeiden.

II.2.4 b Streuung an einem Zentralpotential

Im Fall der Streuung an einem Zentralpotential — d.h. falls die Kraft zwischen den zwei wechselwirkenden Massenpunkten zentral ist — liegt die ganze Bahnkurve des gestreuten Massenpunkts in einer Ebene, der *Streuebene*. Dann ist der Azimutwinkel des auslaufenden Flugrichtung automatisch gleich jenem des Stoßparameters, und die einzige festzulegende Abhängigkeit ist die des Streuwinkels ϑ , der eigentlich nur vom Betrag b abhängt.⁽²⁵⁾

Die Winkelhalbierende der zwei Asymptoten zur Bahnkurve des gestreuten Massenpunkts ist eine Symmetrieachse der Bahnkurve. Sei α der Winkel zwischen dieser Winkelhalbierenden und einer der Asymptoten. Dann gilt automatisch

$$2\alpha + \vartheta = \pi, \quad (\text{II.62a})$$

vgl. Abb. II.9.

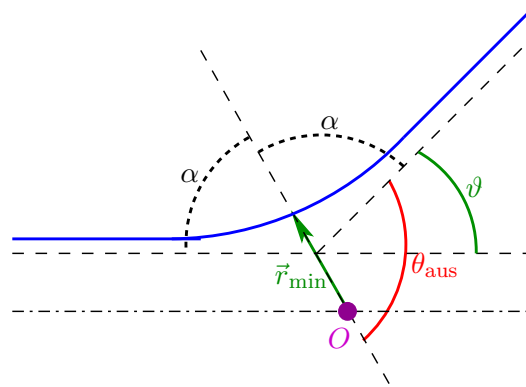


Abbildung II.9 – Definition einiger Winkel.

Wiederum ist der Schnittpunkt zwischen der Trajektorie und der Winkelhalbierenden genau der Punkt des minimalen Abstands r_{\min} zwischen Bahnkurve und Streuzentrum. Unter Verwendung der gleichen Parametrisierung $r(\theta)$ der Bahnkurve wie in § II.2.3 c wird dieser Punkt für $\theta = \pi$ erreicht. Wiederum entspricht die Asymptote für $t \rightarrow -\infty$ bzw. $t \rightarrow +\infty$ einem Winkel $\theta = \theta_{\text{ein}}$ bzw. $\theta = \theta_{\text{aus}}$, wobei (Abb. II.9)

$$\theta_{\text{ein}} = \pi + \alpha \quad \text{bzw.} \quad \theta_{\text{aus}} = \pi - \alpha. \quad (\text{II.62b})$$

Für diese Werte des Polarwinkels ist der gestreute Massenpunkt unendlich weit vom Streuzentrum. Integriert man Gl. (II.55) zwischen π und θ_{ein} , so kommt⁽²⁶⁾

$$\theta_{\text{ein}} - \pi = \int_{r_{\min}}^{\infty} \frac{\ell \, dr'}{r'^2 \sqrt{2\mu[E - V(r')] - \ell^2/r'^2}}.$$

Die linke Seite dieser Gleichung ist gleich α , d.h., dank Gl. (II.62a), $(\pi - \vartheta)/2$. Daher gilt

$$\vartheta = \pi - 2 \int_{r_{\min}}^{\infty} \frac{\ell \, dr'}{r'^2 \sqrt{2\mu[E - V(r')] - \ell^2/r'^2}}. \quad (\text{II.63})$$

Die Werte der Gesamtenergie E und des Bahndrehimpulses ℓ in Gl. (II.63) können durch die charakteristischen Größen des Anfangszustands ausgedrückt werden. Somit ist das Potential für $t \rightarrow -\infty$, d.h. $r \rightarrow \infty$, Null, woraus

$$E = \frac{\vec{p}_{\text{ein}}^2}{2\mu} = \frac{\mu v_{\infty}^2}{2}$$

folgt, wobei $v_{\infty} = |\vec{p}_{\text{ein}}|/\mu$ die Geschwindigkeit des gestreuten Massenpunkts im Unendlichen ist.

⁽²⁵⁾ Dies folgt aus der Kugelsymmetrie des Zentralpotentials, das nur vom Abstand $r = |\vec{r}|$ vom Streuzentrum abhängt.

⁽²⁶⁾ ... mit einem kleinen Dreh: Gleichung (II.56) wurde für eine Lösung mit $dr/d\varphi > 0$ hergeleitet, vgl. ein paar Zeilen oben, während hier $dr/d\varphi < 0$ gilt, weshalb das Negative der linken Seite betrachtet wird.

Wiederum gilt

$$\ell = |\vec{L}| = \mu b v_\infty.$$

Mit diesen Werten wird Gl. (II.63) zu

$$\pi - \varphi_\vartheta = \int_{r_{\min}}^{\infty} \frac{b}{\sqrt{1 - \frac{2V(r')}{\mu v_\infty^2} - \frac{b^2}{r'^2}}} \frac{dr'}{r'^2}. \quad (\text{II.64a})$$

Schließlich ist der dabei auftretende minimale Abstand r_{\min} Lösung der impliziten Gleichung

$$\frac{2V(r_{\min})}{\mu v_\infty^2} + \frac{b^2}{r_{\min}^2} = 1, \quad (\text{II.64b})$$

d.h. bei r_{\min} wird der Nenner des Integranden in Gl. (II.64a) gleich Null.

Beweis der Beziehung (II.64b): Für $\varphi = \pi$, d.h. $r = r_{\min}$, ist die Geschwindigkeit $\vec{v}(r_{\min})$, mit Betrag v_{\min} , senkrecht zum Abstandsvektor \vec{r}_{\min} . Demzufolge ist der Betrag des Drehimpulses dort $\mu v_{\min} r_{\min}$ mit $v_{\min} \equiv |\vec{v}_{\min}|$. Wegen der Drehimpulserhaltung ist dies auch gleich ℓ , d.h.

$$v_{\min} = \frac{b}{r_{\min}} v_\infty.$$

Die gesamte Energie im Punkt der minimalen Abstand lautet dann

$$E = \frac{1}{2} \mu v_{\min}^2 + V(r_{\min}) = \frac{1}{2} \mu \frac{b^2}{r_{\min}^2} v_\infty^2 + V(r_{\min}).$$

Nach Division durch E und unter Verwendung der Beziehung zwischen E , μ und v_∞ ergibt sich genau Gl. (II.64b). \square

Beispiel: Streuung harter Kugel

Zur Illustration kann man die Streuung eines Massenpunkts mit Masse μ am Zentralpotential

$$V(r) = \begin{cases} \infty & \text{für } r < R \\ 0 & \text{für } r \geq R \end{cases} \quad (\text{II.65})$$

betrachten, wobei R eine positive Zahl ist. Dieses Modell entspricht dem Stoß zweier undurchdringbaren, nicht-deformierbaren Kugeln mit Masse $m = 2\mu$ und Radius $R/2$, d.h. deren Zentren sich minimal um R annähern können.

Für $b > R$ gibt es offensichtlich keine Streuung, so dass wir $b < R$ annehmen können, um einen nicht-trivialen Prozess zu erhalten. Physikalische Intuition sagt, dass $r_{\min} = R$ gilt; dies lässt sich aber auch aus der Gleichung (II.64b) beweisen. Für $r_{\min} > R$ lautet diese nämlich $b^2/r_{\min}^2 = 1$, was mit $r_{\min} > R > b$ unmöglich ist; für $r_{\min} < R$ gibt Gl. (II.64b) mit dem Potential (II.65)

$$\frac{2\infty}{\mu v_\infty^2} + \frac{b^2}{r_{\min}^2} = 1,$$

was wieder unmöglich ist. Somit bleibt nur die Möglichkeit $r_{\min} = R$ übrig.

Dabei ist Gl. (II.64b) noch nicht erfüllt, was auf erster Sicht problematisch aussehen kann. Eigentlich ist die Geschwindigkeit des gestreuten Massenpunkts unstetig für $r = R$: dort ist sie, und mit ihr der Drehimpuls, nicht wohldefiniert, so dass der Beweis der Beziehung (II.64b), der auf dem Wert des Drehimpulses in \vec{r}_{\min} basiert, nicht mehr gilt.

Gemäß der Gl. (II.64a) ist der Streuwinkel ϑ durch

$$\vartheta = 2\varphi_\vartheta - \pi = \pi - 2 \int_R^{\infty} \frac{b}{\sqrt{1 - b^2/r'^2}} \frac{dr'}{r'^2}$$

gegeben, wobei $V(r') = 0$ für $r' \geq R$ benutzt wurde. Mit der Substitution $u = b/r'$, entsprechend $du = -b dr'/r'^2$, kommt

$$\vartheta = \pi + 2 \int_{b/R}^0 \frac{du}{\sqrt{1 - u^2}} = \pi + 2 \left[\arccos u \right]_{b/R}^0 = 2 \arccos \frac{b}{R},$$

d.h. $b(\vartheta) = R \cos \frac{\vartheta}{2}$. Dieses Ergebnis kann auch geometrisch gefunden werden: mit dem in Abb. II.10 definierten Winkel α gilt nämlich $b = R \sin \alpha$ (Trigonometrie!). Dann ist $\vartheta = \pi - 2\alpha$, woraus das Resultat folgt.

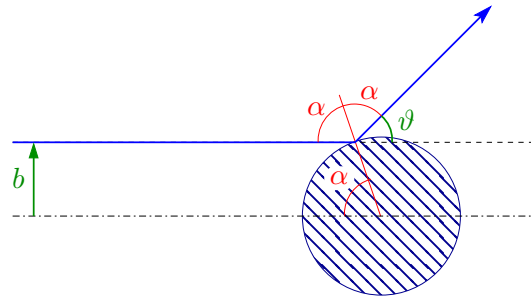


Abbildung II.10 – Streuung harter Kugel.

II.2.4 c Wirkungsquerschnitt

to do

Literatur zum Kapitel II

- Fließbach, *Mechanik* [2] Teil I, Kap. 1–6.
- Greiner, *Klassische Mechanik I* [6] Kap. II und *Klassische Mechanik II* [7] Kap. I & II.
- Nolting, *Klassische Mechanik* [15] Kap. 2 & 3.
- Scheck, *Mechanik* [18] Kap. 1.

Appendix zum Kapitel II

II.A Alternative Herleitung der Keplerschen Bahnkurven

Die im § II.2.3 dargelegte Lösung des Kepler-Problems beruht auf dem Energiesatz für die konservative Bewegung des Massenpunktes in einem Zentralkraftpotential. Insbesondere wurde dabei weder das zweite newtonsche Gesetz explizit benutzt, noch die Beschleunigung des Massenpunktes oder die darauf wirkenden Kräfte geschrieben.

In diesem Anhang wird eine alternative Herleitung der Form der Bahnkurven angegeben, in der die newtonsche Bewegungsgleichung ausdrücklich verwendet wird. Somit kommt man eigentlich schneller ans Ergebnis als im § II.2.3 — dafür verliert man aber die qualitative Diskussion zu den Bahnkurven anhand der energetischen Argumenten.

II.A.1 Vorbereitung: Binetsche Gleichung

Wie in § II.2.3 a kann man argumentieren, dass die Bahnkurve eines Massenpunktes im Zentralkraftpotential $V(\vec{r}) = -\alpha/r$ des Kepler-Problems in einer Ebene bleibt: sei $r(\theta)$ die Parametrisierung dieser Bahnkurve durch Polarkoordinaten. Hiernach wird die gleiche mitbewegte Basis $(\vec{e}_r(t), \vec{e}_\theta(t))$ wie im § II.2.3 verwendet.

Anstatt einer Gleichung für $r(\theta)$, werden wir eine für

$$u(\theta) \equiv \frac{1}{r(\theta)} \quad (\text{II.66})$$

bestimmen. Hiernach werden die Ableitungen von u mit Strichen bezeichnet: $u'(\theta)$, $u''(\theta)$. Wegen der Drehimpulserhaltung ist

$$C \equiv r^2 \dot{\theta} = \frac{\dot{\theta}}{u(\theta)^2} \quad (\text{II.67})$$

eine Konstante: für einen Massenpunkt mit (reduzierter) Masse μ ist $C = \ell/\mu$.

Anhand dieser Zahl lassen sich Zeitableitungen umschreiben. Somit gilt

$$\frac{du(\theta(t))}{dt} = \frac{du(\theta)}{d\theta} \frac{d\theta}{dt} = u'(\theta) \dot{\theta} = C u(\theta)^2 u'(\theta).$$

Auf ähnliche Weise findet man

$$\frac{d^2 u}{dt^2} = C u(\theta)^2 u''(\theta)$$

sowie, unter Verwendung von $\frac{d\vec{e}_r}{d\theta} = \vec{e}_\theta$ und $\frac{d\vec{e}_\theta}{d\theta} = -\vec{e}_r$ [vgl. Gl. (II.46a)–(II.46b)]

$$\frac{d\vec{e}_r}{dt} = C u(\theta)^2 \vec{e}_\theta, \quad \frac{d\vec{e}_\theta}{dt} = C u(\theta)^2 \vec{e}_r.$$

Sei nun

$$\vec{x} = r \vec{e}_r = \frac{1}{u} \vec{e}_r. \quad (\text{II.68})$$

die Position des Massenpunktes entlang seiner Bahnkurve zu einem Zeitpunkt t , der hiernach nicht geschrieben wird. Sukzessive Zeitableitungen mithilfe der oben gefundenen Ergebnisse liefern

$$\dot{\vec{x}} = C [-u'(\theta) \vec{e}_r + u(\theta) \vec{e}_\theta] \quad (\text{II.69})$$

für die Geschwindigkeit und

$$\dot{\vec{x}} = C(-Cu^2u''\vec{e}_r - Cu^2u'\vec{e}_\theta + Cu^2u'\vec{e}_\theta - Cu^3\vec{e}_r)$$

für die Beschleunigung. Nach einfacher Umschreibung lautet die letztere noch

$$\ddot{\vec{x}} = -C^2u(\theta)^2[u''(\theta) + u(\theta)]\vec{e}_r. \quad (\text{II.70})$$

Dieses Ergebnis wird manchmal *Binetsche^(m) Gleichung* für die Beschleunigung genannt.

Bemerkung: Da die Basisvektoren \vec{e}_r , \vec{e}_θ senkrecht aufeinander und auf 1 normiert sind, folgt aus Gl. (II.69) für das Quadrat der Geschwindigkeit

$$\dot{\vec{x}}^2 = C^2[u(\theta)^2 + u'(\theta)^2]. \quad (\text{II.71})$$

Mit dieser Formel lässt sich die kinetische Energie leicht ausdrücken.

II.A.2 Kepler-Problem

Im Kepler-Problem ist die Kraft auf einen Massenpunkt die newtonsche Gravitationskraft

$$\vec{F} = -\frac{\alpha}{r^2} \vec{e}_r \quad (\text{II.72a})$$

mit $\alpha > 0$. Ausgedrückt durch u gilt

$$\vec{F} = -\alpha u^2 \vec{e}_r. \quad (\text{II.72b})$$

Unter Verwendung der Binetschen Gleichung (II.70) lautet das zweite newtonsche Gesetz (I.14) für einen Massenpunkt mit Masse μ

$$\mu\ddot{\vec{x}} = -\mu C^2 u(\theta)^2 [u''(\theta) + u(\theta)] \vec{e}_r = -\alpha u(\theta)^2 \vec{e}_r. \quad (\text{II.73})$$

Nach trivialen Vereinfachungen — unter der Annahme, dass $u(\theta) \neq 0$ — ergibt sich somit die lineare gewöhnliche Differentialgleichung zweiter Ordnung

$$u''(\theta) + u(\theta) = \frac{\alpha}{\mu C^2}. \quad (\text{II.74})$$

Die allgemeine Lösung dieser Gleichung ist die Summe aus der allgemeinen Lösung der homogenen Differentialgleichung $u''(\theta) + u(\theta) = 0$, und zwar $A \cos(\theta - \theta_0)$, und einer speziellen Lösung der inhomogenen Gleichung, z.B. $u(\theta) = \alpha/\mu C^2$:

$$u(\theta) = \frac{\alpha}{\mu C^2} + A \cos(\theta - \theta_0). \quad (\text{II.75})$$

Die Funktion $u(\theta)$ erreicht ihre Extrema für $\theta = \theta_0$ (Maximum) und $\theta = \theta_0 + \pi/2$ (Minimum). Wählt man $\theta_0 = 0$ und definiert man

$$p \equiv \frac{\alpha}{\mu C^2}, \quad \epsilon \equiv Ap,$$

so lautet die Lösung

$$u(\theta) = p(1 + \epsilon \cos \theta) = \frac{1}{r}.$$

Somit findet man wieder die Parametrisierung [vgl. Gl. (II.58b)]

$$r = \frac{p}{1 + \epsilon \cos \theta} \quad (\text{II.76})$$

eines Kegelschnitts mit Exzentrizität ϵ und Parameter p .

Bemerkung: Wenn die Kraft die allgemeinere Form $\vec{F} = F(u) \vec{e}_r$ mit einer beliebigen Funktion F annimmt, so lautet das zweite newtonsche Gesetz $-\mu C^2 u(\theta)^2 [u''(\theta) + u(\theta)] \vec{e}_r = F(u(\theta)) \vec{e}_r$ anstelle

^(m)J. BINET, 1786–1856

der Gl. (II.73). Dies führt zur Differentialgleichung

$$u''(\theta) + u(\theta) = -\frac{1}{\mu C^2} \frac{F(u(\theta))}{u(\theta)^2},$$

die im speziellen Fall $F(u) = -\alpha u^2$ die Gl. (II.74) wiedergibt. Z.B. ergibt sich für $F(u) = -\alpha u^3$ (d.h. $F(r) \propto 1/r^3$) eine analytisch exakt lösbare Gleichung — deren Lösungen nichtperiodische Bahnkurven beschreiben.

KAPITEL III

Lagrange-Formalismus: Grundlagen

III.1	Ein Resultat aus der Variationsrechnung	65
III.1.1	Funktional	65
III.1.2	Extremierung eines Funktionals	66
III.2	Hamilton-Prinzip	68
III.2.1	Definitionen	68
III.2.2	Hamilton-Prinzip. Euler–Lagrange-Gleichungen	70
III.2.3	Erste Beispiele	72
III.2.4	Systeme mit Zwangsbedingungen	75
III.3	Symmetrien und Erhaltungsgrößen	80
III.3.1	Invarianz unter Raumzeit-Transformationen	80
III.3.2	Noether-Theorem	84

Die im Kapitel I dargelegte Formulierung der Mechanik nach Newton ist zwar sehr intuitiv: man zählt alle auf das zu studierende mechanische System wirkenden Kräfte (einschließlich Scheinkräfte, falls das System in einem nicht-inertialen Bezugssystem untersucht wird) auf, schreibt ihre Resultierende auf die rechte Seite des zweiten newtonschen Gesetzes (I.14) auf, und löst die sich daraus ergebenden Differentialgleichungen — möglicherweise anhand eines numerischen Verfahrens.

In der Praxis kann sich der hier skizzierte Vorgang als nicht so direkt herausstellen. Zum einen können die Kräfte nicht einfach formulierbar sein, insbesondere wenn es sich um eine *Zwangskraft* handelt, die die Bewegung des Systems einschränkt, von dem Bewegungszustand aber abhängt. Dies passiert z.B. im Fall eines Pendels, in dem die Stange eine Zwangskraft auf die Masse am Ende des Pendels ausübt, die von der Position und Geschwindigkeit der Masse abhängt. Zum anderen können die „natürlichen“ Variablen der newtonschen Beschreibung, d.h. die zeitabhängigen Koordinaten von Ortsvektoren, nicht die geeignetsten sein.

Zur Beseitigung der besagten Schwierigkeiten kann man zunächst die newtonschen Bewegungsgleichungen (I.14) verallgemeinern, um Zwangskräfte in einer geeigneten Form zu berücksichtigen: diese erste Erweiterung des Formalismus führt zu den sog. *Lagrange⁽ⁿ⁾-Gleichungen erster Art*. In einem zweiten Schritt werden *verallgemeinerte Koordinaten* eingeführt und die Lagrange-Gleichungen erster Art so manipuliert, dass die Zwangskräfte in den resultierenden Gleichungen nicht mehr auftreten. Auf diese Weise erhält man die *Lagrange-Gleichungen zweiter Art*.

Anstatt dieser Vorgehensweise zu folgen,⁽²⁷⁾ kann man die Lagrange-Gleichungen zweiter Art aus einem neu postulierten Prinzip folgern, dem *Hamilton^(o)-Prinzip*. Letzteres lässt sich sowohl in Systemen mit als ohne Zwangskräfte anwenden, und stellt sich oft als einfacher heraus, wenn das Interesse an den Bewegungsgleichungen, nicht an den Kräften, liegt.

Um dieses Prinzip mathematisch formulieren zu können, ist es notwendig, zuerst ein Ergebnis aus der Variationsrechnung einzuführen (Abschn. III.1). Danach werden die Definitionen einiger

⁽²⁷⁾Die interessierte Leserin kann diesen Zugang z.B. in Fließbach [2], Kap. 7–9 finden.

⁽ⁿ⁾J.-L. LAGRANGE, 1736–1813 ^(o)W. R. HAMILTON, 1805–1865

Größen — insbesondere verallgemeinerter Koordinaten, Lagrange-Funktion oder Wirkung — vorgestellt, die eine zentrale Rolle im Hamilton-Prinzip spielen. Das letztere wird dann formuliert und die daraus folgenden Lagrange-Gleichungen zweiter Art hergeleitet, die in diesem Rahmen öfter als *Euler-Lagrange-Gleichungen* bezeichnet werden (Abschn. III.2).

Ein Nebenprodukt des entwickelten Formalismus ist einerseits die genauere Definition des intuitiven Begriffs einer physikalischen Symmetrietransformation, andererseits die Entdeckung eines tiefen Zusammenhangs zwischen Symmetrien und Erhaltungsgrößen (Abschn. III.3).

III.1 Ein Resultat aus der Variationsrechnung

In diesem Abschnitt wird zunächst die Definition eines Funktionals, d.h. einer Funktion von Funktionen, gegeben (§ III.1.1). Dann wird ein nützliches Ergebnis dargelegt, um die Funktionen, die eine bestimmte Funktional extremal machen, zu charakterisieren (§ III.1.2).

III.1.1 Funktional

Im mathematischen Sinne ist eine Funktion oder Abbildung f eine Beziehung, die jedem Element x einer (Definitions-)Menge genau ein Element $f(x)$ einer anderen Menge (Zielmenge) zuordnet. Im physikalischen Kontext ist die Definitionsmenge oft eine Untermenge — ein Gebiet — eines endlich-dimensionalen (Vektor)Raums \mathcal{V} , insbesondere \mathbb{R} , \mathbb{C} , \mathbb{R}^n , oder \mathbb{C}^n , wobei $n \in \mathbb{N}$. Das gleiche gilt für die Zielmenge \mathcal{V}' .

Als *Funktional* bezeichnen Physiker eine Abbildung F , deren Definitionsmenge \mathcal{V} ein unendlich-dimensionaler Funktionenraum ist, während die Zielmenge \mathcal{V}' entweder \mathbb{R} oder \mathbb{C} ist. Der Wert eines Funktionals F für eine Funktion $f \in \mathcal{V}$ wird mit $F[f]$, oder manchmal mit $F[f(x)]$, bezeichnet.

Im Folgenden wird die Zielmenge von Funktionalen immer \mathbb{R} sein. Wiederum werden die Funktionen f der Definitionsmenge bestimmte Eigenschaften besitzen: meistens sind das (mindestens einmal) stetig differenzierbare Funktionen.

Beispiel 1: In diesem und den folgenden Beispielen ist die Definitionsmenge der jeweiligen Funktionale der Raum der stetig differenzierbaren Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$, die automatisch stetig und integrierbar auf jedem endlichen Intervall sind.

Seien zwei reelle Zahlen $x_1 < x_2$; die Integration über das Intervall $[x_1, x_2]$ definiert ein Funktional:⁽²⁸⁾

$$f \mapsto F[f] \equiv \int_{x_1}^{x_2} f(x) dx.$$

Beispiel 2: Sei $x_0 \in \mathbb{R}$. Die Abbildung, die jeder Funktion f ihren Wert im Punkt x_0 zuordnet, ist ein Funktional:

$$f \mapsto F[f] \equiv f(x_0) = \int_{-\infty}^{\infty} \delta(x - x_0) f(x) dx,$$

wobei in der zweiten Gleichung die Dirac^(p)- δ -Distribution (s. Anhang C) eingeführt wurde.

Beispiel 3: Eine reelle Funktion $x \mapsto y(x)$ lässt sich günstig als Kurve in der (x, y) -Ebene darstellen. Wenn $x_1 < x_2$ zwei reelle Zahlen sind, dann ist die Länge $\ell[y]$ der Kurve $y(x)$ zwischen den Punkten $P_1 \equiv (x_1, y(x_1))$ und $P_2 \equiv (x_2, y(x_2))$ ein Funktional

$$\ell[y] \equiv \int_{P_1}^{P_2} dl = \int_{P_1}^{P_2} \sqrt{(dx)^2 + (dy)^2} = \int_{x_1}^{x_2} \sqrt{1 + [y'(x)]^2} dx \quad (\text{III.1})$$

wobei y' die Ableitung von y nach x bezeichnet.

⁽²⁸⁾Mathematisch wird eher das Integral als Funktional mit bestimmten Eigenschaften (wie Linearität) definiert...

^(p)P. A. M. DIRAC, 1902–1984

Beispiel 4: Sei Φ eine reelle Funktion von drei reellen Argumenten und $x_1 < x_2$ zwei reelle Zahlen. Dann definiert die Formel

$$F[f] \equiv \int_{x_1}^{x_2} \Phi(x, f(x), f'(x)) \, dx$$

ein Funktional, das wir im nächsten Paragraphen wieder diskutieren werden.

III.1.2 Extremierung eines Funktionals

III.1.2a Euler-Gleichung

Sei Φ eine stetig differenzierbare Funktion von drei reellen Variablen und $x_1 < x_2$ gegebene reelle Zahlen. Wir betrachten das Funktional

$$f \mapsto F[f] \equiv \int_{x_1}^{x_2} \Phi(x, f(x), f'(x)) \, dx. \quad (\text{III.2})$$

Dabei ist f eine stetig differenzierbare reelle Funktion f — mit Ableitung f' — auf dem Intervall $[x_1, x_2]$, deren Werte an den Grenzen gegeben sind: $f(x_1) = y_1$ und $f(x_2) = y_2$ mit vorgegebenen y_1 und y_2 . Im Folgenden werden die partiellen Ableitungen von Φ nach ihrer Variablen mit $\partial\Phi/\partial x$, $\partial\Phi/\partial f$ und $\partial\Phi/\partial f'$ bezeichnet.

Wir wollen eine notwendige Bedingung dafür herleiten, dass das Funktional ein Extremum — Minimum oder Maximum — für eine Funktion f_0 hat. Das heißt, für alle Funktionen $f(x)$, die $f(x_1) = y_1$ und $f(x_2) = y_2$ erfüllen, gilt $F[f] \geq F[f_0]$ (falls F minimal für f_0) bzw. $F[f] \leq F[f_0]$ (Maximum für f_0).

Sei δf eine beliebige stetig differenzierbare Funktion $[x_1, x_2] \rightarrow \mathbb{R}$, mit $\delta f(x_1) = \delta f(x_2) = 0$. Dann ist $\lambda \mapsto F[f_0 + \lambda \delta f]$ eine Funktion von \mathbb{R} nach \mathbb{R} . Damit F extremal in f_0 sei, muss die Ableitung dieser Funktion nach λ Null für $\lambda = 0$ sein:

$$\left. \frac{dF[f_0 + \lambda \delta f]}{d\lambda} \right|_{\lambda=0} = 0.$$

Das Einsetzen des Ausdrucks von $F[f_0 + \lambda \delta f]$ als Integral [Gl. (III.2)] und das Vertauschen von Ableitung nach λ und Integration über x geben

$$\begin{aligned} \frac{dF[f_0 + \lambda \delta f]}{d\lambda} &= \frac{d}{d\lambda} \left[\int_{x_1}^{x_2} \Phi(x, f_0(x) + \lambda \delta f(x), f_0'(x) + \lambda \delta f'(x)) \, dx \right] \\ &= \int_{x_1}^{x_2} \left[\frac{\partial\Phi}{\partial f} \delta f(x) + \frac{\partial\Phi}{\partial f'} \delta f'(x) \right] dx, \end{aligned}$$

wobei die partiellen Ableitungen in der zweiten Zeile im Punkt $(x, f_0(x) + \lambda \delta f(x), f_0'(x) + \lambda \delta f'(x))$ zu berechnen sind.

Der zweite Summand im Integranden des letzten Integrals lässt sich mithilfe einer partiellen Integration umschreiben:

$$\int_{x_1}^{x_2} \frac{\partial\Phi}{\partial f'} \delta f'(x) \, dx = \left[\frac{\partial\Phi}{\partial f'} \delta f(x) \right]_{x_1}^{x_2} - \int_{x_1}^{x_2} \frac{d}{dx} \left(\frac{\partial\Phi}{\partial f'} \right) \delta f(x) \, dx.$$

Der integrierte Term verschwindet dank der Bedingung $\delta f(x_1) = \delta f(x_2) = 0$. Somit gilt

$$\frac{dF[f_0 + \lambda \delta f]}{d\lambda} = \int_{x_1}^{x_2} \left[\frac{\partial\Phi}{\partial f} - \frac{d}{dx} \left(\frac{\partial\Phi}{\partial f'} \right) \right] \delta f(x) \, dx.$$

Dies muss für $\lambda = 0$ verschwinden, d.h. wenn die partiellen Ableitungen im Integranden im Punkt $(x, f_0(x), f_0'(x))$ ausgewertet sind. Da die Wahl der „Variation“ δf bis auf ihre Werte an den Intervallgrenzen beliebig ist, findet man einfach, dass die Differenz in den eckigen Klammern selbst

identisch verschwinden muss,⁽²⁹⁾ d.h.

$$\frac{\partial \Phi(x, f_0(x), f_0'(x))}{\partial f} - \frac{d}{dx} \left(\frac{\partial \Phi(x, f_0(x), f_0'(x))}{\partial f'} \right) = 0 \quad \forall x \in [x_1, x_2].$$

Somit haben wir bewiesen das

Theorem: Sei Φ eine in allen ihren drei reellen Variablen stetig differenzierbare Funktion und x_1, x_2, y_1, y_2 gegebene reelle Zahlen mit $x_1 < x_2$. Für stetig differenzierbare Funktionen $f : [x_1, x_2] \rightarrow \mathbb{R}$, die $f(x_1) = y_1$ und $f(x_2) = y_2$ erfüllen, wird das Funktional

$$F[f] = \int_{x_1}^{x_2} \Phi(x, f(x), f'(x)) dx \quad (\text{III.3a})$$

definiert. Wenn F ein Extremum für eine Funktion f_0 hat, dann erfüllt f_0 die *Euler-Gleichung*

$$\frac{\partial \Phi(x, f_0(x), f_0'(x))}{\partial f} = \frac{d}{dx} \left(\frac{\partial \Phi(x, f_0(x), f_0'(x))}{\partial f'} \right) \quad \forall x \in [x_1, x_2]. \quad (\text{III.3b})$$

Bemerkungen:

* Bei gegebener Φ ist die zugehörige Euler-Gleichung eine gewöhnliche Differentialgleichung zweiter Ordnung in f_0 .

* Wenn die Funktion Φ bestimmte Bedingungen erfüllt, kann man auch zeigen, dass eine Lösung der Euler-Gleichung (III.3b) auch Lösung des Extremierungsproblems für F ist.

* Die Anwesenheit eines Extremums für das Funktional F wird (durch Physiker) oft mit der kurzen Schreibweise $\delta F[f] = 0$ bezeichnet.

Das obige Theorem lässt sich auf den Fall eines Funktionals

$$F[f_1, \dots, f_s] = \int_{x_1}^{x_2} \Phi(x, f_1(x), f_1'(x), \dots, f_s(x), f_s'(x)) dx \quad (\text{III.4a})$$

von s skalaren Funktionen $f_\alpha, \alpha = 1, \dots, s$ verallgemeinern, wobei Φ jetzt eine stetig differenzierbare Funktion von $2s + 1$ reellen Variablen ist. Wenn ein Satz (f_1, \dots, f_s) das Funktional F minimiert oder maximiert, dann genügen die zugehörigen Funktionen den s Euler-Gleichungen

$$\frac{\partial \Phi}{\partial f_\alpha} = \frac{d}{dx} \left(\frac{\partial \Phi}{\partial f'_\alpha} \right) \quad \forall x \in [x_1, x_2] \quad (\text{III.4b})$$

für jedes $\alpha \in \{1, \dots, s\}$.

Das Theorem kann auch verallgemeinert werden, um den Fall eines Funktionals zu berücksichtigen, dessen Argumente Funktionen mehrerer Variablen x_1, \dots, x_p sind:

$$F[f] = \int_{\mathcal{D}} \Phi \left(x_1, \dots, x_p, f(x_1, \dots, x_p), \frac{\partial f(x_1, \dots, x_p)}{\partial x_1}, \dots, \frac{\partial f(x_1, \dots, x_p)}{\partial x_p} \right) dx_1 \dots dx_p, \quad (\text{III.5a})$$

mit einem p -dimensionalen Integrationsgebiet \mathcal{D} . In dieser Definition ist Φ eine stetig differenzierbare Funktion von $2p + 1$ reellen Variablen. Wenn eine Funktion f das Funktional F minimiert oder maximiert, dann erfüllt sie die Euler-Gleichung

$$\frac{\partial \Phi}{\partial f} = \sum_{j=1}^p \frac{\partial}{\partial x_j} \left(\frac{\partial \Phi}{\partial (\partial f / \partial x_j)} \right) \quad \forall x \in \mathcal{D}. \quad (\text{III.5b})$$

Schließlich lässt sich auch die Extremierung von Funktionalen von N Funktionen von p Variablen behandeln, indem die zwei obigen Verallgemeinerungen kombiniert werden.

⁽²⁹⁾Dies bildet das *Fundamentallema der Variationsrechnung*.

III.1.2b Beispiel

Als Anwendung der Euler-Gleichung können wir den (bekannteren!) kürzesten Weg zwischen zwei Punkten (x_1, y_1) und (x_2, y_2) in der (x, y) -Ebene suchen. Das heißt, wir möchten die Kurve $f(x)$ finden, die $f(x_1) = y_1$ und $f(x_2) = y_2$ erfüllt und das Funktional

$$\ell[f] = \int_{x_1}^{x_2} \sqrt{1 + [f'(x)]^2} dx \quad (\text{III.1})$$

minimiert. Dieses Funktional ist der Form (III.3a) mit $\Phi(x, f(x), f'(x)) = \sqrt{1 + [f'(x)]^2}$. Die zugehörigen partiellen Ableitungen sind

$$\frac{\partial \Phi}{\partial x} = 0, \quad \frac{\partial \Phi}{\partial f} = 0, \quad \frac{\partial \Phi}{\partial f'} = \frac{f'(x)}{\sqrt{1 + [f'(x)]^2}}.$$

Leitet man die letztere nach x ab, so ergibt sich

$$\frac{d}{dx} \left(\frac{\partial \Phi}{\partial f'} \right) = \frac{f''(x)}{(1 + [f'(x)]^2)^{3/2}},$$

wobei f'' die zweite Ableitung von f bezeichnet. Die Euler-Gleichung (III.3b) lautet dann

$$\frac{\partial \Phi}{\partial f} = \frac{d}{dx} \left(\frac{\partial \Phi}{\partial f'} \right) \quad \Leftrightarrow \quad 0 = \frac{f''(x)}{(1 + [f'(x)]^2)^{3/2}}.$$

Dies gibt sofort $f''(x) = 0$: nach doppelter Integration unter Berücksichtigung der Randbedingungen in x_1 und x_2 kommt

$$f(x) = y_1 + \frac{y_2 - y_1}{x_2 - x_1}(x - x_1),$$

d.h. die Gleichung der Geraden zwischen den zwei Endpunkten, was zu erwarten war.

Bemerkung: Zur Berechnung der totalen Ableitung nach x von $\partial \Phi / \partial f'$ kann man entweder den Ausdruck der partiellen Ableitung direkt ableiten, oder die Kettenregel anwenden und dafür die drei zweiten Ableitungen $\partial^2 \Phi / \partial x \partial f'$, $\partial^2 \Phi / \partial f \partial f'$ und $\partial^2 \Phi / \partial f'^2$ berechnen und mit jeweils $x' = 1$, $f'(x)$ und $f''(x)$ multiplizieren. Natürlich führen beide Wege zum gleichen Ergebnis.

III.2 Hamilton-Prinzip

In diesem Abschnitt wird eine Methode zur Herleitung der Bewegungsgleichungen für ein (mechanisches) System beschrieben, die sich auch für Systeme eignet, in denen sich die genaue Form der Kräfte nicht einfach präzisieren lässt. Dabei wird eine Funktion der relevanten Freiheitsgrade des Systems „postuliert“, die Lagrange-Funktion, welche die ganze Information über die Dynamik des Systems enthält (§ III.2.1); die gesuchten Bewegungsgleichungen lassen sich dann durch Ableitungen der Lagrange-Funktion ausdrücken (§ III.2.2). Diese Konstruktion wird in § III.2.3 anhand ein paar erster einfacher Beispiele illustriert; insbesondere wird eine allgemeine Form der Lagrange-Funktion für Systeme mit konservativen Kräften gefunden. Der Formalismus wird dann auf den Fall von Systemen mit Zwangsbedingungen erweitert (§ III.2.4).

III.2.1 Definitionen

III.2.1 a Verallgemeinerte Koordinaten

Sei ein System Σ aus N Massenpunkten. Die N Ortsvektoren $\vec{x}_a(t)$ mit $a \in \{1, \dots, N\}$ sind äquivalent zu $3N$ Koordinaten. Wenn die Massenpunkte sich unabhängig voneinander bewegen können, dann besitzt das System genau $s = 3N$ Freiheitsgrade.

In der Praxis untersucht man oft aber Probleme, in denen es (feste) Zusammenhänge zwischen den $3N$ Koordinaten gibt: z.B. ist der Abstand zwischen zwei Massenpunkten festgelegt, oder die

Massenpunkte müssen auf irgendeiner Fläche bleiben. In solchen Fällen sind die $3N$ Koordinaten der Positionen nicht unabhängig von einander, d.h. die Anzahl s der Freiheitsgrade des Systems ist kleiner als $3N$. Es ist daher in solchen Situationen sinnvoll, den Zustand des Systems mit nur s skalaren Größen zu charakterisieren, anstatt mit $3N$.

Somit führt man zur Charakterisierung eines mechanischen Systems s *verallgemeinerte* bzw. *generalisierte Koordinaten* $q^1(t), q^2(t), \dots, q^s(t)$ ein, welche die folgenden Bedingungen erfüllen:

- i. Sie legen den Zustand des Systems eindeutig fest, d.h. jede Position $\vec{x}_a(t)$ mit $a \in \{1, \dots, N\}$ lässt sich als eine bestimmte Funktion — die auch mit \vec{x}_a bezeichnet wird — der Zeit und der s verallgemeinerten Koordinaten schreiben:

$$\vec{x}_a(t) = \vec{x}_a(t, q^1(t), \dots, q^s(t)) \quad \forall a \in \{1, \dots, N\}. \quad (\text{III.6})$$

- ii. Sie sind alle unabhängig voneinander, d.h. es existiert keine Funktion f von $s+1$ Argumenten, für die $f(t, q^1(t), \dots, q^s(t)) = 0$ für jede Zeit t gilt.

Die newtonschen Bewegungsgleichungen sind Differentialgleichungen zweiter Ordnung, so dass man neben den Positionen $\vec{x}_a(t)$ noch die N Geschwindigkeiten $\dot{\vec{x}}_a(t)$ betrachten muss, um die Zeitentwicklung für $t' > t$ zu bestimmen. Laut Gl. (III.6) sind diese Geschwindigkeiten Funktion von der Zeit t , den verallgemeinerten Koordinaten $q^\alpha(t)$ für $\alpha \in \{1, \dots, s\}$, und von deren Zeitableitungen. Dementsprechend muss man auch die s *verallgemeinerten Geschwindigkeiten* $\{\dot{q}^\alpha(t)\}$ in Betracht ziehen.

Genau wie die Geschwindigkeiten $\dot{\vec{x}}_a(t)$ im Allgemeinen unabhängig von den Positionen $\vec{x}_a(t)$ sind, gilt dies auch für die verallgemeinerten Geschwindigkeiten und Koordinaten.

Der Kürze halber werden die s generalisierten Koordinaten $q^\alpha(t)$ bzw. Geschwindigkeiten $\dot{q}^\alpha(t)$ kollektiv als ein s -dimensionaler Vektor $\mathbf{q}(t)$ bzw. $\dot{\mathbf{q}}(t)$ bezeichnet.

Im allgemeinen Fall wird jedoch $\mathbf{q}(t)$ kein Element eines physikalisch bedeutenden Vektorraums sein. Beispielsweise ist die Menge der möglichen Kugelkoordinaten (r, θ, z) nicht \mathbb{R}^3 , denn r ist immer nicht-negativ, $r \geq 0$. Die „Vektoren“ $\mathbf{q}(t)$ sind eher Punkte einer *Mannigfaltigkeit*⁽³⁰⁾ der Dimension s , die *Konfigurationsraum* genannt wird.

Bemerkungen:

* Während die Anzahl s der Freiheitsgrade für ein bestimmtes Problem eindeutig festgelegt ist, gilt das im Allgemeinen nicht für die verallgemeinerten Koordinaten $q^\alpha(t)$. Dies entspricht z.B. der Freiheit bei der Wahl von Koordinaten.

* Wie es schon der Fall bei „üblichen“ Koordinaten ist, z.B. bei Kugelkoordinaten, haben verallgemeinerte Koordinaten nicht unbedingt die Dimension einer Länge.

III.2.1 b Lagrange-Funktion. Wirkung

Jedem System aus s Freiheitsgraden kann eine reelle Funktion von $2s + 1$ reellen Variablen zugeordnet werden, die *Lagrange-Funktion* \mathcal{L} , welche die Bewegung des Systems völlig bestimmt, und die physikalische Dimension einer Energie hat, $[\mathcal{L}] = \text{M L}^2 \text{T}^{-2}$. Dabei nimmt die Funktion als Argumente die Zeit t , die s verallgemeinerten Koordinaten $\{q^\alpha\}$ und die s verallgemeinerten Geschwindigkeiten $\{\dot{q}^\alpha\}$:

$$\mathcal{L}(t, \mathbf{q}, \dot{\mathbf{q}}). \quad (\text{III.7})$$

Bemerkung: Die Form der Lagrange-Funktion wird später angegeben, nachdem ihre Rolle in der Bestimmung der Bewegungsgleichungen eines Systems genauer präzisiert worden ist (vgl. § III.2.3).

⁽³⁰⁾Eine (reelle) Mannigfaltigkeit der Dimension n ist ein Raum, der lokal dem euklidischen Raum \mathbb{R}^n ähnelt.

Es sei aber schon hier erwähnt, dass diese Form nicht eindeutig festgelegt ist, auch wenn die verallgemeinerten Koordinaten gewählt wurden. Dementsprechend hat die Lagrange-Funktion keine physikalische Bedeutung: sie stellt nur ein Hilfsmittel dar, um die Bewegungsgleichungen herzuleiten (§ III.2.2) oder um Konstanten der Bewegung zu finden (Abschn III.3).

Seien jetzt zwei Zeitpunkte $t_1 < t_2$. In der Zeitentwicklung des Systems sind seine verallgemeinerten Koordinaten und Geschwindigkeiten Funktionen der Zeit: $\{q^\alpha(t)\}$, $\{\dot{q}^\alpha(t)\}$. Werden diese in die Lagrange-Funktion (III.7) eingesetzt, so kann sie als eine Funktion der Variablen t allein betrachtet werden:

$$\mathcal{L}(t, \mathbf{q}(t), \dot{\mathbf{q}}(t)). \quad (\text{III.8})$$

Das Integral dieser Funktion über das Zeitintervall $[t_1, t_2]$ definiert die *Wirkung*, die auch *Wirkungsintegral* genannt wird:

$$S[\mathbf{q}] \equiv \int_{t_1}^{t_2} \mathcal{L}(t, \mathbf{q}(t), \dot{\mathbf{q}}(t)) dt. \quad (\text{III.9})$$

In der Schreibweise wurde schon berücksichtigt, dass die Wirkung von den verallgemeinerten Koordinaten und Geschwindigkeiten $\{q^\alpha(t)\}$, $\{\dot{q}^\alpha(t)\}$ abhängt, so dass es sich um ein Funktional handelt — weshalb $S[\mathbf{q}]$ auch als *Wirkungsfunktional* bezeichnet wird.

Die Wirkung hat die Dimension des Produkts von Energie und Zeit, $[S] = \text{M L}^2 \text{T}^{-1}$; die zugehörige Einheit im SI-System ist das $\text{J} \cdot \text{s} = \text{kg} \cdot \text{m}^2 \cdot \text{s}^{-1}$.

III.2.2 Hamilton-Prinzip. Euler–Lagrange-Gleichungen

Es seien $t_1 < t_2$. Bei dem *Hamilton-Prinzip* handelt es sich um ein *Extremalprinzip*, laut dem die Zeitentwicklung eines Systems im Intervall $[t_1, t_2]$ so erfolgt, dass die physikalisch realisierten verallgemeinerten Koordinaten $\mathbf{q}(t)$ und Geschwindigkeiten $\dot{\mathbf{q}}(t)$ das Wirkungsfunktional (III.9) extremal (oder „stationär“) machen. Dies wird oft kurz als

$$\text{Hamilton-Prinzip: } \delta S[\mathbf{q}] = 0 \quad (\text{III.10})$$

ausgedrückt.

Gemäß der Verallgemeinerung des in § III.1.2 a angegebenen Theorems genügt dann das Integrand des Wirkungsintegrals, d.h. die Lagrange-Funktion \mathcal{L} , den Euler-Gleichungen (III.4b), die in diesem Kontext als *Euler–Lagrange-Gleichungen* bezeichnet werden:

$$\frac{\partial \mathcal{L}(t, \mathbf{q}(t), \dot{\mathbf{q}}(t))}{\partial q^\alpha} = \frac{d}{dt} \left(\frac{\partial \mathcal{L}(t, \mathbf{q}(t), \dot{\mathbf{q}}(t))}{\partial \dot{q}^\alpha} \right) \quad \forall \alpha \in \{1, \dots, s\}. \quad (\text{III.11})$$

Diese s Differentialgleichungen zweiter Ordnung für die s verallgemeinerten Koordinaten stellen die Bewegungsgleichungen des Systems dar.

Bemerkungen:

* Das Hamilton-Prinzip wird auch *Wirkungsprinzip*, *Prinzip der stationären Wirkung* oder, fehlerhaft, *Prinzip der kleinsten Wirkung* genannt.

Wiederum werden die Euler–Lagrange-Gleichungen (III.11) noch *Lagrange-Gleichungen zweiter Art* genannt.

* Die Lagrange-Funktion für ein gegebenes System ist nicht eindeutig: wenn $\mathcal{L}(t, \mathbf{q}(t), \dot{\mathbf{q}}(t))$ eine mögliche Lagrange-Funktion bezeichnet, dann ist

$$\mathcal{L}'(t, \mathbf{q}(t), \dot{\mathbf{q}}(t)) \equiv \mathcal{L}(t, \mathbf{q}(t), \dot{\mathbf{q}}(t)) + \frac{d}{dt} \mathcal{M}(t, \mathbf{q}(t)) \quad (\text{III.12})$$

eine äquivalente Lagrange-Funktion, die zu denselben Euler–Lagrange-Gleichungen führt, wobei \mathcal{M} eine beliebige stetig differenzierbare Funktion der Zeit und der generalisierten Koordinaten ist.

Beweis: Das mit \mathcal{L}' berechnete Wirkungsintegral lautet

$$S'[\mathbf{q}] \equiv \int_{t_1}^{t_2} \mathcal{L}'(t, \mathbf{q}(t), \dot{\mathbf{q}}(t)) dt = S[\mathbf{q}] + \int_{t_1}^{t_2} \frac{d\mathcal{M}(t, \mathbf{q}(t))}{dt} dt = S[\mathbf{q}] + \mathcal{M}(t_2, \mathbf{q}(t_2)) - \mathcal{M}(t_1, \mathbf{q}(t_1)).$$

Unter Variationen $\mathbf{q} \rightarrow \mathbf{q} + \delta\mathbf{q}$, mit $\delta\mathbf{q}(t_1) = \delta\mathbf{q}(t_2) = \mathbf{0}$ bleiben die zwei letzten Terme unverändert, d.h. sie haben keinen Einfluss auf die Position des Extremums bzw. auf die Euler–Lagrange-Gleichungen. \square

Alternativ kann man die totale Ableitung nach der Zeit $d\mathcal{M}(t, \mathbf{q}(t))/dt$ durch die partiellen Ableitungen (nach t und nach den q^α) ausdrücken. Dann findet man, dass der zusätzliche Term in der Lagrange-Funktion Beiträge zu den beiden Seiten der Euler–Lagrange-Gleichungen liefert, die sich gegenseitig kompensieren.

Die Invarianz der Bewegungsgleichungen unter Transformationen (III.12) wird manchmal als *Eichinvarianz* bezeichnet, und eine derartige Transformation als *Eichtransformation*.

* Die Euler–Lagrange-Gleichungen (III.11) sind *Form-invariant* oder *kovariant* unter den gleichzeitigen Transformationen

$$q^\alpha \rightarrow q^{\alpha'} = q^{\alpha'}(t, \mathbf{q}) \quad \text{für } \alpha = 1, \dots, s, \quad (\text{III.13a})$$

$$\mathcal{L}(t, \mathbf{q}, \dot{\mathbf{q}}) \rightarrow \mathcal{L}'(t, \mathbf{q}', \dot{\mathbf{q}}') \equiv \mathcal{L}(t, \mathbf{q}(t, \mathbf{q}'), \dot{\mathbf{q}}(t, \mathbf{q}')), \quad (\text{III.13b})$$

wobei die Transformation ein *Diffeomorphismus* ist, d.h. eine beliebig oft differenzierbare (\mathcal{C}^∞) bijektive Abbildung, deren Umkehrfunktion

$$q^\alpha = q^\alpha(t, \mathbf{q}') \quad \text{für } \alpha = 1, \dots, s \quad (\text{III.14})$$

auch \mathcal{C}^∞ ist. Die Transformation (III.13a) stellt ein Koordinatenwechsel dar.

Beweis: Unter Verwendung der Kettenregel lautet die verallgemeinerte Geschwindigkeit $\dot{q}^{\alpha'}$

$$\dot{q}^{\alpha'}(t, \mathbf{q}, \dot{\mathbf{q}}) = \frac{d}{dt} q^{\alpha'}(t, \mathbf{q}) = \frac{\partial q^{\alpha'}}{\partial t} + \sum_{\beta} \frac{\partial q^{\alpha'}}{\partial q^{\beta}} \frac{dq^{\beta}}{dt} = \frac{\partial q^{\alpha'}}{\partial t} + \sum_{\beta} \frac{\partial q^{\alpha'}}{\partial q^{\beta}} \dot{q}^{\beta}. \quad (\text{III.15a})$$

Somit ist jede $\dot{q}^{\alpha'}$ eine lineare Funktion der verallgemeinerten Geschwindigkeiten $\dot{\mathbf{q}}$, obwohl die Koordinate $q^{\alpha'}$ unabhängig davon ist. Ähnlich folgt aus Gl. (III.14) und der Kettenregel

$$\dot{q}^{\alpha}(t, \mathbf{q}', \dot{\mathbf{q}}') = \frac{d}{dt} q^{\alpha}(t, \mathbf{q}') = \frac{\partial q^{\alpha}}{\partial t} + \sum_{\beta'} \frac{\partial q^{\alpha}}{\partial q^{\beta'}} \frac{dq^{\beta'}}{dt} = \frac{\partial q^{\alpha}}{\partial t} + \sum_{\beta'} \frac{\partial q^{\alpha}}{\partial q^{\beta'}} \dot{q}^{\beta'}. \quad (\text{III.15b})$$

Seien $\alpha, \beta \in \{1, \dots, s\}$. Auf Gl. (III.15a) bzw. (III.15b) liest man die partielle Ableitung von $\dot{q}^{\alpha'}$ nach \dot{q}^{β} bzw. von \dot{q}^{α} nach $\dot{q}^{\beta'}$, und zwar

$$\frac{\partial \dot{q}^{\alpha'}}{\partial \dot{q}^{\beta}} = \frac{\partial q^{\alpha'}}{\partial q^{\beta}} \quad \text{bzw.} \quad \frac{\partial \dot{q}^{\alpha}}{\partial \dot{q}^{\beta'}} = \frac{\partial q^{\alpha}}{\partial q^{\beta'}}. \quad (\text{III.16})$$

Sei $\mathcal{L}'(t, \mathbf{q}', \dot{\mathbf{q}}') \equiv \mathcal{L}(t, \mathbf{q}(t, \mathbf{q}'), \dot{\mathbf{q}}(t, \mathbf{q}', \dot{\mathbf{q}}'))$. Dank der Kettenregel lautet die partielle Ableitung dieser Funktion nach $\dot{q}^{\beta'}$ mit $\beta \in \{1, \dots, s\}$

$$\frac{\partial \mathcal{L}'(t, \mathbf{q}', \dot{\mathbf{q}}')}{\partial \dot{q}^{\beta'}} = \frac{\partial}{\partial \dot{q}^{\beta'}} \mathcal{L}(t, \mathbf{q}(t, \mathbf{q}'), \dot{\mathbf{q}}(t, \mathbf{q}', \dot{\mathbf{q}}')) = \sum_{\beta} \frac{\partial \mathcal{L}}{\partial \dot{q}^{\alpha}} \frac{\partial \dot{q}^{\alpha}}{\partial \dot{q}^{\beta'}},$$

denn die $\{q^{\alpha}\}$ sind unabhängig von $\dot{q}^{\beta'}$. Die totale Zeitableitung dieser partiellen Ableitung folgt dann aus der Produktregel:

$$\frac{d}{dt} \frac{\partial \mathcal{L}'(t, \mathbf{q}', \dot{\mathbf{q}}')}{\partial \dot{q}^{\beta'}} = \sum_{\beta} \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}^{\alpha}} \frac{\partial \dot{q}^{\alpha}}{\partial \dot{q}^{\beta'}} \right) = \sum_{\beta} \left[\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}^{\alpha}} \right) \frac{\partial \dot{q}^{\alpha}}{\partial \dot{q}^{\beta'}} + \frac{\partial \mathcal{L}}{\partial \dot{q}^{\alpha}} \frac{d}{dt} \left(\frac{\partial \dot{q}^{\alpha}}{\partial \dot{q}^{\beta'}} \right) \right].$$

Der erste Term in den eckigen Klammern lässt sich mit den Euler–Lagrange-Gleichungen für \mathcal{L}

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}^\alpha} \right) = \frac{\partial \mathcal{L}}{\partial q^\alpha}$$

umschreiben. Dank Gl. (III.16) kann $\partial \dot{q}^\alpha / \partial \dot{q}^{\beta'}$ durch $\partial q^\alpha / \partial q^{\beta'}$ ersetzt werden. Das Austausch der Ableitungen nach t und $q^{\beta'}$ gibt dann

$$\frac{d}{dt} \left(\frac{\partial q^\alpha}{\partial q^{\beta'}} \right) = \frac{\partial}{\partial q^{\beta'}} \frac{dq^\alpha}{dt} = \frac{\partial \dot{q}^\alpha}{\partial q^{\beta'}}.$$

Insgesamt gilt somit

$$\frac{d}{dt} \frac{\partial \mathcal{L}'(t, \mathbf{q}', \dot{\mathbf{q}}')}{\partial \dot{q}^{\beta'}} = \sum_{\beta} \left(\frac{\partial \mathcal{L}}{\partial q^\alpha} \frac{\partial q^\alpha}{\partial q^{\beta'}} + \frac{\partial \mathcal{L}}{\partial \dot{q}^\alpha} \frac{\partial \dot{q}^\alpha}{\partial q^{\beta'}} \right) = \frac{\partial}{\partial q^{\beta'}} \mathcal{L}(t, \mathbf{q}(t, \mathbf{q}'), \dot{\mathbf{q}}(t, \mathbf{q}', \dot{\mathbf{q}}')) = \frac{\partial \mathcal{L}'(t, \mathbf{q}', \dot{\mathbf{q}}')}{\partial q^{\beta'}}$$

wobei die zweite Gleichheit aus der Kettenregel folgt. \square

Diese Form-Invarianz ist wichtig, denn sie bedeutet, dass die Euler–Lagrange-Gleichungen die gleiche Form annehmen, wenn man die verallgemeinerten Koordinaten (diffeomorph) transformiert. Somit ist die Wahl der Koordinaten „beliebig“.

Definition: Der *verallgemeinerte Impuls* oder (zu q^α) *kanonisch konjugierte Impuls* wird als

$$p_\alpha(t, \mathbf{q}(t), \dot{\mathbf{q}}(t)) \equiv \frac{\partial \mathcal{L}(t, \mathbf{q}(t), \dot{\mathbf{q}}(t))}{\partial \dot{q}^\alpha} \quad (\text{III.17})$$

definiert.

Falls die Lagrange-Funktion nicht explizit von einer verallgemeinerten Koordinaten q^α abhängt, dann folgt sofort aus der entsprechenden Euler–Lagrange-Gleichung, dass der zugehörige generalisierte Impuls p_α eine Erhaltungsgröße ist (vgl. auch § III.3.1 b).

III.2.3 Erste Beispiele

III.2.3 a Freier Massenpunkt

Als mögliche Lagrange-Funktion für einen freien Massenpunkt mit Masse m und Position $\vec{x}(t)$ kann man folgende Funktion annehmen:

$$\mathcal{L}(t, \vec{x}, \dot{\vec{x}}) = \frac{m}{2} \dot{\vec{x}}^2. \quad (\text{III.18})$$

Dabei sind die generalisierten Koordinaten q^α die kartesischen Koordinaten x^i mit $\alpha = i = 1, \dots, 3$, so dass die zugehörigen verallgemeinerten Geschwindigkeiten einfach $\dot{q}^\alpha = \dot{x}^i$ sind.

Wegen der Gleichheit

$$\frac{\partial \mathcal{L}(t, \vec{x}, \dot{\vec{x}})}{\partial \dot{x}^i} = m \dot{x}^i$$

ist der mit x^i assoziierte verallgemeinerte Impuls einfach die i -te Komponente des kinetischen Impulses $\vec{p} = m \dot{\vec{x}}$. Wiederum gilt $\partial \mathcal{L} / \partial x^i = 0$. Dann drücken die assoziierten Euler–Lagrange-Gleichungen

$$\frac{d}{dt} \frac{\partial \mathcal{L}(t, \vec{x}(t), \dot{\vec{x}}(t))}{\partial \dot{x}^i} = m \ddot{x}^i(t) = 0 = \frac{\partial \mathcal{L}(t, \vec{x}(t), \dot{\vec{x}}(t))}{\partial x^i}$$

die Erhaltung des kinetischen Impulses in der Bewegung aus, d.h. man findet das erste newtonsche Gesetz (I.13) wieder.

III.2.3 b System aus Massenpunkten mit konservativen Kräften

Für ein System aus N Massenpunkten, die miteinander über konservative Kräfte wechselwirken, kann man als generalisierte Koordinaten wieder die kartesischen Koordinaten der Ortsvektoren

betrachten, d.h. $q^\alpha = x_a^j$, wobei x_a^i die i -te Koordinate der Position des Massenpunkts a bezeichnet. Dann lautet eine mögliche Lagrange-Funktion

$$\mathcal{L}(t, \{\vec{x}_a\}, \{\dot{\vec{x}}_a\}) = \sum_{a=1}^N \frac{m}{2} \dot{\vec{x}}_a^2 - V(\{\vec{x}_a\}) = T - V \quad (\text{III.19})$$

mit der kinetischen Energie T und dem Potential V .

Schreibt man nämlich die entsprechenden Euler-Lagrange-Gleichungen

$$\frac{\partial \mathcal{L}}{\partial x_a^i} = \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}_a^i} \right) \quad \text{für } i \in \{1, \dots, 3\} \text{ und } a \in \{1, \dots, N\},$$

so gilt einerseits, genau wie im obigen Fall des freien Massenpunkts, $\partial \mathcal{L} / \partial \dot{x}_a^i = m \dot{x}_a^i$. Daraus ergibt sich sofort

$$\frac{d}{dt} \frac{\partial \mathcal{L}(t, \{\vec{x}_a(t)\}, \{\dot{\vec{x}}_a(t)\})}{\partial \dot{x}_a^i} = m \ddot{x}_a^i(t).$$

Andererseits beträgt die Ableitung nach x_a^i auf der linken Seite der Euler-Lagrange-Gleichung

$$\frac{\partial \mathcal{L}}{\partial x_a^i} = - \frac{\partial V}{\partial x_a^i} = F_a^i,$$

wobei F_a^i die i -te Komponente der Kraft auf den a -ten Massenpunkt bezeichnet. Insgesamt findet man somit

$$m \ddot{x}_a^i(t) = F_a^i(t), \quad \text{für } i \in \{1, \dots, 3\} \text{ und } a \in \{1, \dots, N\}, \quad (\text{III.20})$$

d.h. genau die i -te Komponente der aus dem zweiten newtonschen Gesetz bekannten Bewegungsgleichung für den a -ten Massenpunkt.

III.2.3 c System aus Massenpunkten mit nicht-konservativen Kräften

Falls die Massenpunkte eines System miteinander über nicht-konservative Kräfte wechselwirken, kann man generell keine allgemein geltende zugehörige Lagrange-Funktion schreiben. In zwei Fällen sind jedoch Erweiterungen der Lagrange-Funktion (III.19) möglich, und zwar einerseits für Kräfte, die aus einem verallgemeinerten, Zeit- und Geschwindigkeitsabhängigen Potential abgeleitet werden können, andererseits für Reibungskräfte proportional zur Geschwindigkeit.

Kräfte aus einem generalisierten Potential

Betrachten wir eine nicht-konservative Kraft $\vec{F}_{\text{n.-k.}}$ auf einen Massenpunkt derart, dass ihre kartesischen Komponenten in der Form

$$F_{\text{n.-k.}}^i(t) = - \frac{\partial U(t, \vec{x}(t), \dot{\vec{x}}(t))}{\partial x^i} + \frac{d}{dt} \frac{\partial U(t, \vec{x}(t), \dot{\vec{x}}(t))}{\partial \dot{x}^i} \quad \text{für } i = 1, 2, 3 \quad (\text{III.21a})$$

geschrieben werden können, wobei $x^i(t)$ bzw. $\dot{x}^i(t)$ die i -te Komponente der Position bzw. der Geschwindigkeit des Massenpunkts bezeichnet. Dabei ist U ein *verallgemeinertes Potential*, das nicht nur von den Positionen, sondern auch von den Geschwindigkeiten abhängt.

Vektoriell kann man schreiben

$$\vec{F}_{\text{n.-k.}}(t) = - \vec{\nabla}_{\vec{r}} U(t, \vec{x}(t), \dot{\vec{x}}(t)) + \frac{d}{dt} \vec{\nabla}_{\vec{v}} U(t, \vec{x}(t), \dot{\vec{x}}(t)), \quad (\text{III.21b})$$

wobei $\vec{\nabla}_{\vec{r}}$ bzw. $\vec{\nabla}_{\vec{v}}$ den Gradienten bezüglich der Orts- bzw. Geschwindigkeitskoordinaten bezeichnet.

Sei zudem V das Potential, aus dem die konservativen Kräfte \vec{F} auf den Massenpunkt folgen.

Dann stellt

$$\mathcal{L}(t, \vec{x}, \dot{\vec{x}}) = T - V - U = \frac{m}{2} \dot{\vec{x}}^2 - V(\vec{x}) - U(t, \vec{x}, \dot{\vec{x}}) \quad (\text{III.22})$$

eine geeignete Lagrange-Funktion für den Massenpunkt dar. Die daraus folgenden Euler–Lagrange-Gleichungen nehmen nämlich die Form:

$$m\ddot{x}^i(t) = F^i(t) + F_{\text{n.-k.}}^i(t) \quad \text{für } i = 1, 2, 3$$

an.

Beispiel: Sei Φ bzw. \vec{A} eine reelle bzw. vektorielle Funktion von Zeit und Ort. Man kann eine Lagrange-Funktion für einen Massenpunkt mit Position \vec{x} und Geschwindigkeit $\dot{\vec{x}}$ durch

$$U(t, \vec{x}, \dot{\vec{x}}) = q\Phi(t, \vec{x}) - q\dot{\vec{x}} \cdot \vec{A}(t, \vec{x}) \quad (\text{III.23a})$$

definieren, mit einer für den Massenpunkt charakteristischen Zahl q . Über die Beziehung (III.21b) führt dieses verallgemeinerte Potential zur Kraft

$$\vec{F}_{\text{n.-k.}}(t) = q \left(-\vec{\nabla}\Phi(t, \vec{x}(t)) - \frac{\partial \vec{A}(t, \vec{x}(t))}{\partial t} + \dot{\vec{x}}(t) \times [\vec{\nabla} \times \vec{A}(t, \vec{x}(t))] \right). \quad (\text{III.23b})$$

Definiert man $\vec{E} \equiv -\vec{\nabla}\Phi - \partial\vec{A}/\partial t$ und $\vec{B} \equiv \vec{\nabla} \times \vec{A}$, dann ist dies genau die Lorentz-Kraft, die eine Punktladung q in einem elektromagnetischen Feld (\vec{E}, \vec{B}) erfährt.⁽³¹⁾

Bemerkung: Wie in § I.1.3 b schon bemerkt wurde leistet der „magnetische Anteil“ $q\dot{\vec{x}}(t) \times \vec{B}$ der Lorentz-Kraft keine Arbeit. Diese Aussage entspricht genau der Tatsache, dass die Lorentz-Kraft sich aus einem verallgemeinerten Potential über Gl. (III.21a) ableiten lässt.

Reibungskräfte

Im Gegensatz zu konservativen Kräften oder zur Lorentz-Kraft hängt die Arbeit von Reibungskräften zwischen zwei Punkten von der genauen Trajektorie ab. Dementsprechend kann man kein generalisiertes Potential finden, mit dessen Hilfe Reibungskräfte über Gl. (III.21a) abgeleitet werden können.

Für die Stokessche Reibungskraft (vgl. Beispiel in § I.2.2), die proportional zur Geschwindigkeit ist, kann man eine Funktion einführen, deren Ableitung die Kraft ist. Im Fall der Kraft auf einen Massenpunkt mit Position bzw. Geschwindigkeit \vec{x} bzw. $\dot{\vec{x}}$ definiert man die *Rayleigh*^(q)-*Dissipationsfunktion*

$$\mathcal{D}(\dot{\vec{x}}) \equiv \sum_{i=1}^3 \gamma_i (\dot{x}^i)^2 \quad (\text{III.24})$$

mit positiven Koeffizienten γ_i . Nach Ableitung nach der Komponente \dot{x}^i der Geschwindigkeit folgt

$$-\frac{\partial \mathcal{D}}{\partial \dot{x}^i} = -\gamma_i \dot{x}^i, \quad (\text{III.25})$$

hier ohne Summe über den doppelt auftretenden Index i , d.h. eine Stokessche Reibungskraft mit unterschiedlichen Reibungskoeffizienten in den drei Richtungen.

Führt man dann generalisierte Koordinaten $\mathbf{q}(t)$ ein, so lässt sich die Dissipationsfunktion durch diese über $\mathcal{D}(t, \mathbf{q}, \dot{\mathbf{q}}) = \mathcal{D}(\{\dot{\vec{x}}(t, \mathbf{q}, \dot{\mathbf{q}})\})$ ausdrücken. Um die Bewegungsgleichungen zu erhalten werden dann „modifizierte Lagrange-Gleichungen“ postuliert:

$$\frac{\partial \mathcal{L}(t, \mathbf{q}(t), \dot{\mathbf{q}}(t))}{\partial q^\alpha} = \frac{d}{dt} \left(\frac{\partial \mathcal{L}(t, \mathbf{q}(t), \dot{\mathbf{q}}(t))}{\partial \dot{q}^\alpha} \right) + \frac{\partial \mathcal{D}(t, \mathbf{q}(t), \dot{\mathbf{q}}(t))}{\partial \dot{q}^\alpha}, \quad (\text{III.26})$$

die das richtige Ergebnis liefern. Es muss aber betont werden, dass sich die Gleichungen (III.26) im Gegensatz zu den üblichen Euler–Lagrange-Gleichungen nicht aus einem Extremalprinzip herleiten lassen. In diesem Sinne ist die obige Konstruktion ziemlich künstlich.

⁽³¹⁾Die Kräfte auf elektrische Ladungs- und Stromverteilungen in einem nicht-dynamischen elektromagnetischen Feld werden in Abschn. XIII.1 detaillierter diskutiert.

^(q)J. W. Strutt, Lord RAYLEIGH, 1842–1919

Bemerkung: Das Interessante bei der Rayleigh-Dissipationsfunktion liegt daran, dass die instantane Leistung, welche der Massenpunkt gegen die Reibungskraft verrichten muss, gleich $2\mathcal{D}$ ist.

III.2.4 Systeme mit Zwangsbedingungen

Oft können sich die individuellen Punkte eines Systems nicht uneingeschränkt unabhängig voneinander bewegen. Stattdessen müssen die Positionen oder die Geschwindigkeiten der Punkte gewissen Bedingungen genügen (§ III.2.4 a): beispielsweise müssen Abstände konstant bleiben, oder die Punkte müssen auf einer gegebenen Fläche, oder innerhalb eines gegebenen Volumens bleiben.

Diese Einschränkungen der Bewegung werden durch Zwangskräfte erzwungen, deren Form nicht immer einfach auszudrücken ist. Im Rahmen des Lagrange-Formalismus kann man die Bewegungsgleichungen formulieren, ohne die Zwangskräfte genau zu kennen, indem man geeignete verallgemeinerte Koordinaten verwendet (§ III.2.4 b).

III.2.4 a Zwangsbedingungen

A priori besitzt ein System aus N Massenpunkten $3N$ Freiheitsgrade für seine Bewegung, und zwar für jeden Massenpunkt drei (orthogonal zueinander) Bewegungsrichtungen, die sich als die insgesamt $3N$ Komponenten $x_a^i(t)$, $a \in \{1, \dots, N\}$, $i = 1, 2, 3$ der Ortsvektoren visualisieren lassen.

Sei angenommen, dass diese Massenpunkte Zwangskräften unterliegen. Diese schränken die Entwicklung des Systems ein, was sich mathematisch durch Gleichungen der Form

$$f_j(t, \vec{x}_1(t), \dots, \vec{x}_N(t), \dot{\vec{x}}_1(t), \dots, \dot{\vec{x}}_N(t)) = 0 \quad (\text{III.27})$$

ausdrücken lässt, wobei jede f_j eine genügend reguläre Funktion von $6N + 1$ Variablen ist. Wir werden annehmen, dass es r unabhängige solche Beziehungen gibt, d.h. $j \in \{1, \dots, r\}$.

Definition: Eine Gleichung der Form (III.27), in welcher nur die Zeit und die Positionen der Massenpunkte auftreten, heißt *holonome Zwangsbedingung* (oder *holonome Nebenbedingung*):

$$f_j(t, \vec{x}_1(t), \dots, \vec{x}_N(t)) = 0, \quad (\text{III.28})$$

wobei f_j jetzt eine Funktion von $3N + 1$ Variablen ist.

Dank jeder holonomen Zwangsbedingung lässt sich eine der $3N$ räumlichen Koordinaten durch alle anderen (und die Zeit) ausdrücken, zumindest implizit. Ein System aus N Massenpunkten mit r holonomen Nebenbedingungen hat somit nur noch $s = 3N - r$ unabhängige Freiheitsgrade.

Bemerkungen:

* Einschränkungen der Bewegung können auch durch Ungleichungen ausgedrückt werden. Zum Beispiel lässt das Einsperren von Teilchen in einem gegebenen Volumen, z.B. in einem Behälter, mit Ungleichungen formulieren. Dabei handelt es sich um eine *nichtholonome Zwangsbedingung*.

* Man unterscheidet noch zwischen *skleronomen*, d.h. zeitunabhängigen, und *rheonomen*, d.h. zeitabhängigen, Zwangsbedingungen.

Beispiel 1: Gleitende Masse auf einer schiefen Ebene

Sei ein durch einen Massenpunkt modellierter Körper mit Masse m , der reibungslos auf einer Ebene unter dem Einfluss des Schwerfeldes $\vec{g} = -g\vec{e}_z$ gleitet.

Wegen einer ersten Zwangsbedingung $f_1(t, \vec{x}(t)) \equiv x(t) = 0$ bleibt der Körper in der (y, z) -Ebene. Dazu bewegt er sich in der schiefen Ebene mit Neigungswinkel ϕ , so dass die kartesischen Koordinaten seiner Bahnkurve $z(t) = -y(t) \tan \phi$ erfüllen. Dementsprechend gibt es eine zweite Zwangsbedingung $f_2(t, \vec{x}(t)) \equiv z(t) - y(t) \tan \phi = 0$.

Da es $r = 2$ holonome Zwangsbedingungen gibt, bleibt nur $s = 3N - r = 1$ Freiheitsgrad übrig.

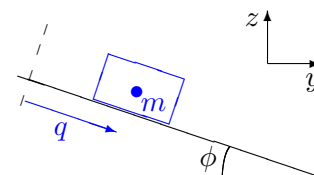


Abbildung III.1

Beispiel 2: Einfaches ebenes Pendel⁽³²⁾

Wir betrachten ein ebenes Pendel bestehend aus einer Punktmasse m am Ende eines masselosen Stabs mit fester Länge l ; das ganze System liegt im Schwerfeld $\vec{g} = -g\vec{e}_z$. Sei $\vec{x}(t) = (x(t), y(t), z(t))$ die Trajektorie der Masse, wobei der Nullpunkt des kartesischen Koordinatensystems im Aufhängepunkt des Pendels genommen wird.

Das Pendel schwingt in der (y, z) -Ebene, d.h. $x(t) = 0$: dies lässt sich durch eine erste Zwangsbedingung $f_1(t, \vec{x}(t)) \equiv x(t) = 0$ ausdrücken.

Wiederum entspricht die Forderung einer festen Länge l einer zweiten Zwangsbedingung $f_2(t, \vec{x}(t)) \equiv y(t)^2 + z(t)^2 - l^2 = 0$.

Wie bei der gleitenden Masse bleibt somit nur ein Freiheitsgrad übrig, d.h. ein einziger Parameter soll reichen, um die Position des Pendels völlig zu charakterisieren.

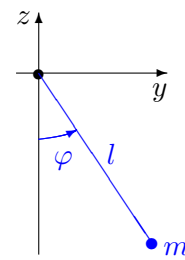


Abbildung III.2

III.2.4 b Lagrange-Formalismus für Systeme mit Zwangsbedingungen

Einer der Vorteile des Lagrange-Formalismus besteht darin, dass man Kräfte und insbesondere Zwangskräfte nicht explizit ausdrücken soll, um die Bewegungsgleichungen zu erhalten. Zu diesem Zweck kann man einem relativ einfachen Rezept folgen:⁽³³⁾

- Als erster Schritt muss man $s = 3N - r$ verallgemeinerte Koordinaten $\{q_1, \dots, q_s\} \equiv \mathbf{q}$ einführen. Diese sollen so gewählt werden, dass sie die Konfigurationen des Systems parametrisieren, welche die Zwangsbedingungen erfüllen. Dementsprechend sollen die $\{q^\alpha\}$ durch die Nebenbedingungen uneingeschränkt sein.
- Danach soll man die Lagrange-Funktion $\mathcal{L} = T - V$, die man als Funktion von den kartesischen Ortskoordinaten kennt, durch die verallgemeinerten Koordinaten \mathbf{q} und Geschwindigkeiten $\dot{\mathbf{q}}$ ausdrücken. Dafür müssen die kartesischen Koordinaten als Funktionen der \mathbf{q} geschrieben werden.
In diesem Falle enthält das Potential V eigentlich nur die Beiträge, die nicht Zwangskräfte verursachen.
- Schließlich kann man die Euler-Lagrange-Gleichungen aufstellen und lösen — möglicherweise numerisch.

Beispiel 1: Gleitende Masse auf einer schiefen Ebene

Für den gleitenden Körper der Abb. III.1 ist die Bewegung eingeschränkt in die Richtung senkrecht zur schiefen Ebene und in die x -Richtung; entlang der Neigung der Ebene bleibt die Bewegung noch frei. Daher stellt die dargestellte Entfernung (aus einem beliebigen Referenzpunkt in der Ebene, der als Nullpunkt des kartesischen Koordinatensystems dient) q eine geeignete generalisierte Koordinate dar. Dann gelten

$$x(t) = 0, \quad y(t) = q(t) \cos \phi, \quad z(t) = -q(t) \sin \phi.$$

Die Zeitableitungen sind einfach und führen sofort zur kinetischen Energie

$$T = \frac{m}{2} (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) = \frac{m}{2} \dot{q}^2.$$

Wiederum lautet das Potential im Schwerfeld

$$V = mgz = -mgq \sin \phi.$$

⁽³²⁾Dieses System wird auch wegen der vielen Vereinfachungen (Annahmen einer punktförmigen Masse, eines masselosen Stabs) als „mathematisches Pendel“ bezeichnet.

⁽³³⁾Die „Begründung“ des Rezepts wird später hinzugefügt.

Dies gibt die Lagrange-Funktion

$$\mathcal{L} = T - V = \frac{m}{2} \dot{q}^2 + mgq \sin \phi$$

mit partiellen Ableitungen

$$\frac{\partial \mathcal{L}}{\partial q} = mg \sin \phi \quad \text{und} \quad \frac{\partial \mathcal{L}}{\partial \dot{q}} = m\dot{q},$$

so dass die Euler–Lagrange-Gleichung (III.11) zur Bewegungsgleichung

$$\ddot{q}(t) = g \sin \phi$$

führt, d.h. zur Gleichung einer gleichförmig beschleunigten geradlinigen Bewegung.

Beispiel 2: Einfaches ebenes Pendel

Wir betrachten wieder das in Abb. III.2 dargestellte ebene Pendel. Die zwei Zwangsbedingungen schränken die (kartesische) z -Komponente und den Abstand vom Aufhängepunkt ein. Dagegen ist der Ablenkwinkel φ noch uneingeschränkt, und jede Konfiguration des System, welche die Nebenbedingungen erfüllt, kann mit diesem Winkel φ allein charakterisiert werden. Somit ist $q = \varphi$ eine gute Wahl für die (hier ist $s = 1$) verallgemeinerte Koordinate.

Dann lassen sich die kartesischen Koordinaten der Position $(x(t), y(t), z(t))$ einfach durch $\varphi(t)$ ausdrücken:

$$x(t) = 0, \quad y(t) = l \sin \varphi(t), \quad z(t) = -l \cos \varphi(t),$$

woraus dank der Kettenregel

$$\dot{x}(t) = 0, \quad \dot{y}(t) = l\dot{\varphi}(t) \cos \varphi(t), \quad \dot{z}(t) = l\dot{\varphi}(t) \sin \varphi(t)$$

folgt. Ersetzt man diese kartesischen Koordinaten in den Ausdrücken der kinetischen Energie

$$T = \frac{m}{2} \dot{x}^2 = \frac{m}{2} (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) = \frac{m}{2} l^2 \dot{\varphi}^2 (\cos^2 \varphi + \sin^2 \varphi) = \frac{ml^2}{2} \dot{\varphi}^2$$

und des Potentials

$$V = mgz = -mgl \cos \varphi,$$

so ergibt sich für die Lagrange-Funktion:

$$\mathcal{L}(t, \varphi, \dot{\varphi}) = T - V = \frac{ml^2}{2} \dot{\varphi}^2 + mgl \cos \varphi. \quad (\text{III.29})$$

Die partiellen Ableitungen der Lagrange-Funktion lauten dann

$$\frac{\partial \mathcal{L}}{\partial \varphi} = -mgl \sin \varphi \quad \text{und} \quad \frac{\partial \mathcal{L}}{\partial \dot{\varphi}} = ml^2 \dot{\varphi}.$$

Diese Ableitungen sollen nun im Punkt $(t, \varphi(t), \dot{\varphi}(t))$ ausgewertet werden. Die zweite führt zu

$$\frac{d}{dt} \frac{\partial \mathcal{L}(t, \varphi(t), \dot{\varphi}(t))}{\partial \dot{\varphi}} = ml^2 \ddot{\varphi}(t),$$

so dass die Euler–Lagrange-Gleichung (III.11) zur Bewegungsgleichung

$$\ddot{\varphi}(t) = -\frac{g}{l} \sin \varphi(t) \quad (\text{III.30})$$

führt.

Der Vollständigkeit halber wird noch die Lösung der Gl. (III.30) dargelegt.

Für Schwingungen kleiner Amplitude, d.h. für $|\varphi(t)| \ll 1$, liefert eine Taylor-Entwicklung zur ersten Ordnung die Näherung $\sin \varphi(t) \sim \varphi(t)$, d.h. die Differentialgleichung (III.30) vereinfacht sich zur Bewegungsgleichung

$$\ddot{\varphi}(t) = -\omega^2 \varphi(t) \quad \text{mit } \omega \equiv \sqrt{\frac{g}{l}}$$

eines harmonischen Oszillators mit Kreisfrequenz ω bzw. Periodendauer $\mathcal{T} = 2\pi/\omega = 2\pi\sqrt{l/g}$, deren allgemeine Lösung der Form

$$\varphi(t) = A \cos \omega t + B \sin \omega t$$

mit $A, B \in \mathbb{R}$ ist. Dabei sollen $|A|$ und $|B|$ viel kleiner als 1 sein, damit die Amplitude der Bewegung „klein“ bleibt.

Für Schwingungen beliebiger Amplitude lässt sich kein einfacher Ausdruck für die Funktion $\varphi(t)$ finden. Offensichtlich ist die Bewegung noch periodisch: sei \mathcal{T} die Periodendauer. Multipliziert man Gl. (III.30) mit $\dot{\varphi}(t)$ und integriert man über die Zeit, so kommt

$$\int_{t_0}^t \ddot{\varphi}(t') \dot{\varphi}(t') dt' = \frac{1}{2} [\dot{\varphi}(t)^2 - \dot{\varphi}(t_0)^2] = -\frac{g}{l} \int_{t_0}^t \sin \varphi(t') \dot{\varphi}(t') dt' = \frac{g}{l} [\cos \varphi(t) - \cos \varphi(t_0)],$$

wobei t_0 einen gegebenen Referenzzeitpunkt bezeichnet. Mit Anfangsbedingungen $\varphi(t_0) \equiv \varphi_0$ und $\dot{\varphi}(t_0) = 0$ stellt $|\varphi_0|$ die maximale Auslenkung des Pendels dar, d.h. für jede Zeit $t \geq t_0$ gilt $\varphi(t) \leq |\varphi_0|$ bzw. $\cos \varphi(t) \geq \cos \varphi_0$. Dann liefert die Separation der Variablen

$$dt = \pm \sqrt{\frac{l}{2g}} \frac{d\varphi}{\sqrt{\cos \varphi - \cos \varphi_0}}.$$

Wählt man z.B. $\varphi_0 > 0$, so nimmt $\varphi(t)$ während der ersten Halbperiode ab, und man soll die Gleichung mit dem $-$ Vorzeichen betrachten. Für $t = t_0 + \mathcal{T}/4$ erreicht das Pendel zum ersten Mal die Position $\varphi = 0$; nach Integration zwischen t_0 und $t \leq t_0 + \mathcal{T}/4$ ergibt sich

$$t - t_0 = -\sqrt{\frac{l}{2g}} \int_{\varphi_0}^{\varphi(t)} \frac{d\varphi}{\sqrt{\cos \varphi - \cos \varphi_0}} = \sqrt{\frac{l}{2g}} \int_{\varphi(t)}^{\varphi_0} \frac{d\varphi}{\sqrt{\cos \varphi - \cos \varphi_0}}. \quad (\text{III.31})$$

Dabei kann der Term auf der rechten Seite auch als

$$\sqrt{\frac{l}{2g}} \int_{\varphi(t)}^{\varphi_0} \frac{d\varphi}{\sqrt{\cos \varphi - \cos \varphi_0}} = \sqrt{\frac{l}{2g}} \int_0^{\varphi_0} \frac{d\varphi}{\sqrt{\cos \varphi - \cos \varphi_0}} - \sqrt{\frac{l}{2g}} \int_0^{\varphi(t)} \frac{d\varphi}{\sqrt{\cos \varphi - \cos \varphi_0}}$$

geschrieben werden: dann gilt⁽³⁴⁾

$$\begin{aligned} \sqrt{\frac{l}{2g}} \int_0^{\varphi(t)} \frac{d\varphi}{\sqrt{\cos \varphi - \cos \varphi_0}} &= \sqrt{\frac{l}{g}} F \left(\arcsin \sqrt{\frac{1 - \cos \varphi(t)}{1 - \cos \varphi_0}}, \sqrt{\frac{1 - \cos \varphi_0}{2}} \right) \\ &= \sqrt{\frac{l}{g}} F \left(\arcsin \frac{\sin(\varphi(t)/2)}{\sin(\varphi_0/2)}, \sin \frac{\varphi_0}{2} \right), \end{aligned} \quad (\text{III.32})$$

wobei das *unvollständige elliptische Integral erster Art* F durch

$$F(\vartheta, k) \equiv \int_0^{\vartheta} \frac{d\varphi}{\sqrt{1 - k^2 \sin^2 \varphi}} = \int_0^{\sin \vartheta} \frac{dt}{\sqrt{(1 - t^2)(1 - k^2 t^2)}}$$

definiert ist.⁽³⁵⁾ Der Position $\varphi(t = t_0 + \mathcal{T}/4) = 0$ entspricht das Argument $\vartheta = \frac{\pi}{2}$, d.h. die Periode einer Schwingung wird durch

$$\frac{\mathcal{T}}{4} = \sqrt{\frac{l}{g}} F \left(\frac{\pi}{2}, \sin \frac{\varphi_0}{2} \right) = \sqrt{\frac{l}{g}} K \left(\sin \frac{\varphi_0}{2} \right) \quad (\text{III.33})$$

⁽³⁴⁾ Dies folgt aus Gl. 2.571-4 in Ref. [20].

⁽³⁵⁾ Die Funktionen F und K werden im Kap. 19 vom *NIST Handbook of Mathematical Functions* [21] diskutiert, dessen online Version unten <http://dlmf.nist.gov/> verfügbar ist.

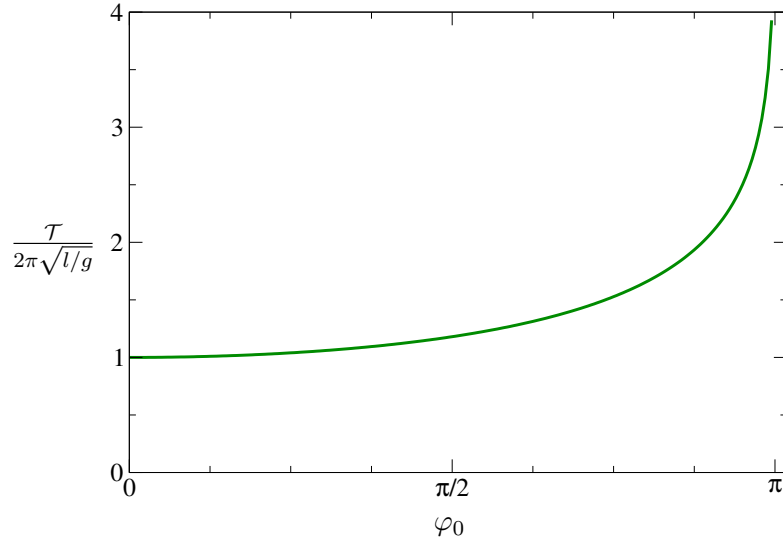


Abbildung III.3 – Periode (III.33) des einfachen Pendels, geteilt durch den Wert für kleine Schwingungen, in Abhängigkeit der maximalen Amplitude φ_0 .

gegeben, wobei $K(k) \equiv F(\frac{\pi}{2}, k)$ das *elliptische Integral erster Art* ist.⁽³⁵⁾ Für kleinen „elliptischen Modul“ $k \ll 1$ gilt $K(k) \sim \frac{\pi}{2}(1 + k^2/4)$, d.h. man findet $\mathcal{T} \sim 2\pi\sqrt{l/g}$ für Schwingungen mit kleiner Amplitude. Dagegen geht $K(k)$ im Limes $k \rightarrow 1^-$ gegen $+\infty$. Dementsprechend braucht das Pendel, ausgehend von der Anfangsposition $\varphi_0 \rightarrow \pi^-$, eine unendliche lange Zeit, um $\varphi = 0$ zu erreichen: die Position $\varphi_0 = \pi$ ist eine (instabile) Gleichgewichtsposition.

Aus Gl. (III.31), (III.32) und (III.33) folgt (für $t - t_0 \leq \mathcal{T}/4$)

$$t - t_0 = \frac{\mathcal{T}}{4} - \sqrt{\frac{l}{g}} F\left(\arcsin \frac{\sin(\varphi(t)/2)}{\sin(\varphi_0/2)}, \sin \frac{\varphi_0}{2}\right).$$

Diese Beziehung kann invertiert werden mithilfe der durch $\operatorname{am}(F(\vartheta, k), k) \equiv \vartheta$ definierten *Jacobi-schen Amplitude*^(r), die somit eine Umkehrfunktion des unvollständigen elliptischen Integrals erster Art ist.⁽³⁶⁾

$$\arcsin \frac{\sin(\varphi(t)/2)}{\sin(\varphi_0/2)} = \operatorname{am}\left(\sqrt{\frac{g}{l}}\left(\frac{\mathcal{T}}{4} - t + t_0\right), \sin \frac{\varphi_0}{2}\right),$$

und daher

$$\sin \frac{\varphi(t)}{2} = \sin\left(\frac{\varphi_0}{2}\right) \sin\left[\operatorname{am}\left(\sqrt{\frac{g}{l}}\left(\frac{\mathcal{T}}{4} - t + t_0\right), \sin \frac{\varphi_0}{2}\right)\right].$$

Dabei definiert der Sinus der Jacobi-Amplitude den *sinus amplitudinis* $\operatorname{sn}(z, k) \equiv \sin \operatorname{am}(z, k)$ für $z \in \mathbb{C}$.⁽³⁶⁾ Somit kann die Position $\varphi(t)$ des Pendels für die Anfangsbedingungen $\varphi(t_0) \equiv \varphi_0$ und $\dot{\varphi}(t_0) = 0$ durch den letzteren ausgedrückt werden:

$$\varphi(t) = 2 \arcsin\left[\sin\left(\frac{\varphi_0}{2}\right) \operatorname{sn}\left(\sqrt{\frac{g}{l}}\left(\frac{\mathcal{T}}{4} - t + t_0\right)\right)\right] \quad \text{für } 0 \leq t - t_0 \leq \frac{\mathcal{T}}{4}. \quad (\text{III.34})$$

Der Verlauf von $\varphi(t)$ für verschiedene Werte der Amplitude φ_0 der Schwingung wird gezeigt in Abb. III.4. Für den Vergleich wird $\varphi(t)$ durch φ_0 geteilt, während die Zeitvariable t durch die Periode (III.33) normiert wird — für $\varphi_0 = 1,99\pi/2$ ist \mathcal{T} etwa 4 mal größer als für kleine Schwingungen. Für die Letzteren mit $\varphi_0 \ll 1$ findet man das harmonische Verhalten wieder. Dagegen bleibt im Limes $\varphi_0 \rightarrow \pi$ der Winkel $\varphi(t)$ viel länger in der Nähe von φ_0 , bevor es schnell durch 0 geht, um wieder länger in der Nähe von $-\varphi_0$ zu bleiben.

⁽³⁶⁾Die Jacobischen elliptischen Funktionen, insbesondere am und sn , werden im Kap. 22 vom *NIST Handbook of Mathematical Functions* vorgestellt [21].

^(r)C. G. JACOBI, 1804–1851

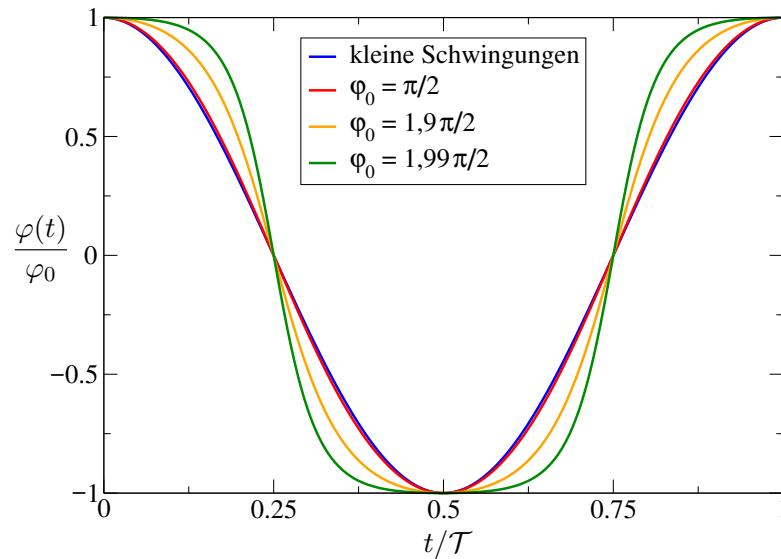


Abbildung III.4 – Zeitabhängigkeit (III.34) des Winkels des einfachen Pendels, geteilt durch die maximale Amplitude φ_0 . Die Zeitvariable wird durch die Periode (III.33) der Schwingung normiert.

III.3 Symmetrien und Erhaltungsgrößen

Im Rahmen des newtonschen Formalismus sind Erhaltungsgrößen bzw. Erhaltungssätze schöne Konstrukte, deren Ursprung aber unklar bleibt: Gleichungen werden manipuliert, Größen werden eingeführt, auf erster Sicht gibt es aber keine tiefere unterliegende, gemeinsame Begründung für die Existenz der verschiedenen Konstanten der Bewegung.

Im Lagrange-Formalismus lässt sich eine solche globale Erklärung für die Erhaltungssätze einfacher erkennen. Somit spiegelt eine Erhaltungsgröße die Invarianz der theoretischen Beschreibung, und der Physik, unter bestimmten Transformationen der Koordinaten wider. Diese Invarianz entspricht einer *Symmetrie* des physikalischen Systems.

Dieser Zusammenhang zwischen Invarianz der Physik unter einigen Transformationen und Erhaltungsgrößen wird zunächst am Beispiel verschiedener Arten von Transformationen der Raumzeit-Koordinaten — Translationen in der Zeit oder im Raum und Drehungen — illustriert (§ III.3.1). Die Verallgemeinerung auf den Fall mehr abstrakter Transformationen wird dann in § III.3.2 skizziert, wobei der Begriff von Symmetrien im physikalischen Sinne genauer definiert wird.

III.3.1 Invarianz unter Raumzeit-Transformationen

Per Annahme ist die Raumzeit der klassischen Mechanik *zeitlich und räumlich homogen* und *räumlich isotrop*. Betrachte man ein idealisiertes Experiment — einschließlich der ganzen Umgebung, die dabei einen Einfluss hat —, das man beliebig verschieben, sowohl im Raum als in der Zeit, oder drehen kann. Dann soll das Ergebnis des Experiments weder von der Anfangszeit, noch von der Position oder der Orientierung im Raum abhängen. Die Modellierung dieses Experiments soll dann invariant unter zeitlichen und räumlichen Translationen und unter Drehungen sein. Daraus ergibt sich die Konstanz in der Bewegung der drei „klassischen“ Erhaltungsgrößen der newtonschen Mechanik, und zwar Energie, Impuls und Drehimpuls.

III.3.1 a Homogenität der Zeit und Energieerhaltung

Als erstes können wir die Folgerung der Invarianz unter Zeittranslationen untersuchen. Wir betrachten dementsprechend eine Theorie mit einer Lagrange-Funktion \mathcal{L} , die nicht explizit von der

Zeit abhängt:

$$\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}),$$

wobei die $\mathbf{q} = \{q^\alpha\}_{\alpha=1,\dots,s}$ verallgemeinerte Koordinaten sind.

Definiert man dann eine Größe, die *Energie* des Systems, durch

$$E \equiv \sum_{\alpha=1}^s \frac{\partial \mathcal{L}(\mathbf{q}(t), \dot{\mathbf{q}}(t))}{\partial \dot{q}^\alpha} \dot{q}^\alpha(t) - \mathcal{L}(\mathbf{q}(t), \dot{\mathbf{q}}(t)), \quad (\text{III.35a})$$

so ist diese eine Erhaltungsgröße

$$\frac{dE}{dt} = 0. \quad (\text{III.35b})$$

Anders gesagt folgt Energieerhaltung in einem physikalischen System aus der Invarianz des Systems unter Zeitverschiebungen.

Beweis: das einfache Ableiten der Definition (III.35a) nach der Zeit gibt

$$\frac{dE}{dt} = \frac{d}{dt} \left(\sum_{\alpha=1}^s \frac{\partial \mathcal{L}}{\partial \dot{q}^\alpha} \dot{q}^\alpha \right) - \frac{d\mathcal{L}}{dt} = \sum_{\alpha=1}^s \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}^\alpha} \dot{q}^\alpha \right) - \frac{d\mathcal{L}}{dt},$$

wobei die Lagrange-Funktion \mathcal{L} und ihre partielle Ableitungen im Punkt $(\mathbf{q}(t), \dot{\mathbf{q}}(t))$ auszuwerten sind, bevor die totale Zeitableitung gebildet wird. Der erste Term auf der rechten Seite kann mithilfe der Produktregel und der Euler-Lagrange-Gleichung (III.11) transformiert werden:

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}^\alpha} \dot{q}^\alpha \right) = \left(\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}^\alpha} \right) \dot{q}^\alpha + \frac{\partial \mathcal{L}}{\partial \dot{q}^\alpha} \frac{d\dot{q}^\alpha}{dt} = \frac{\partial \mathcal{L}}{\partial q^\alpha} \dot{q}^\alpha + \frac{\partial \mathcal{L}}{\partial \dot{q}^\alpha} \ddot{q}^\alpha \quad \text{für jedes } \alpha.$$

Wiederum lässt sich der zweite Term mithilfe der Kettenregel umschreiben:

$$\frac{d\mathcal{L}}{dt} = \sum_{\alpha=1}^s \left(\frac{\partial \mathcal{L}}{\partial q^\alpha} \frac{dq^\alpha}{dt} + \frac{\partial \mathcal{L}}{\partial \dot{q}^\alpha} \frac{d\dot{q}^\alpha}{dt} \right) = \sum_{\alpha=1}^s \left(\frac{\partial \mathcal{L}}{\partial q^\alpha} \dot{q}^\alpha + \frac{\partial \mathcal{L}}{\partial \dot{q}^\alpha} \ddot{q}^\alpha \right).$$

Somit ist die Differenz der beiden Terme genau Null. \square

Bemerkung: Unter Verwendung der verallgemeinerten Impulse (III.17) lässt sich die Energie noch als

$$E \equiv \sum_{\alpha=1}^s p_\alpha(\mathbf{q}(t), \dot{\mathbf{q}}(t)) \dot{q}^\alpha(t) - \mathcal{L}(\mathbf{q}(t), \dot{\mathbf{q}}(t)) \quad (\text{III.35c})$$

schreiben.

Beispiel: Sei ein System mit der Lagrange-Funktion

$$\mathcal{L} = T - V = \frac{1}{2} \sum_{\beta,\gamma=1}^s \lambda_{\beta\gamma}(\mathbf{q}) \dot{q}^\beta \dot{q}^\gamma - V(\mathbf{q}), \quad (\text{III.36})$$

d.h. mit einer kinetischen Energie, die quadratisch in den verallgemeinerten Geschwindigkeiten $\{\dot{q}^\alpha\}$ ist. Dabei sollen die möglicherweise \mathbf{q} -abhängigen Koeffizienten $\lambda_{\beta\gamma}$ symmetrisch unter dem Austausch von β und γ sein: $\lambda_{\beta\gamma} = \lambda_{\gamma\beta}$ für alle $\beta, \gamma \in \{1, \dots, s\}$.

Der Leserin sind schon zwei Beispiele von kinetischen Termen dieser Form bekannt. Wenn die verallgemeinerten Koordinaten $\{q^\beta\}$ die kartesischen Komponenten $\{x_a^i\}$ von Massenpunkten sind, dann lautet die kinetische Energie

$$\sum_a \frac{1}{2} m_a \dot{x}_a^2 = \sum_a \sum_i \frac{1}{2} m_a (x_a^i)^2 = \frac{1}{2} \sum_\beta m_\beta (\dot{q}^\beta)^2,$$

was ein Spezialfall der Gl. (III.36) mit $\lambda_{\beta\gamma} = \frac{1}{2} m_\beta \delta_{\beta\gamma}$ darstellt, wobei $\delta_{\beta\gamma}$ das Kronecker-Symbol ist.

Beschreibt man einen Massenpunkt m , dessen Bewegung in einer Ebene stattfindet, mit Polarkoordinaten (r, θ) , so lautet seine kinetische Energie (vgl. § II.2.3 c)

$$T = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2).$$

Setzt man $q^1 \equiv r$, $q^2 \equiv \theta$, so ist dies noch gleich $\frac{1}{2}[m(\dot{q}^1)^2 + m(q^1)^2(\dot{q}^2)^2]$, d.h. der Form (III.36) mit $\lambda_{11} = m$, $\lambda_{22} = m(q^1)^2$ — in diesem Fall also \mathbf{q} -abhängig — und $\lambda_{12} = \lambda_{21} = 0$.

Das Einsetzen der Lagrange-Funktion (III.36) in die Definition (III.35a) gibt

$$E = \sum_{\alpha=1}^s \dot{q}^\alpha \frac{\partial}{\partial \dot{q}^\alpha} \left(\frac{1}{2} \sum_{\beta,\gamma=1}^s \lambda_{\beta\gamma} \dot{q}^\beta \dot{q}^\gamma - V \right) - \frac{1}{2} \sum_{\beta,\gamma=1}^s \lambda_{\beta\gamma} \dot{q}^\beta \dot{q}^\gamma + V.$$

Einerseits hängt das Potential nicht von der generalisierten Geschwindigkeiten ab, $\partial V / \partial \dot{q}^\alpha = 0$ für alle $\alpha \in \{1, \dots, s\}$, so dass der zweite Term in den Klammern keinen Beitrag liefert. Andererseits gilt unter Verwendung der Produktregel

$$\sum_{\alpha=1}^s \dot{q}^\alpha \frac{\partial}{\partial \dot{q}^\alpha} \left(\frac{1}{2} \sum_{\beta,\gamma=1}^s \lambda_{\beta\gamma} \dot{q}^\beta \dot{q}^\gamma \right) = \frac{1}{2} \sum_{\alpha=1}^s \dot{q}^\alpha \sum_{\beta,\gamma=1}^s \lambda_{\beta\gamma} (\delta_{\alpha\beta} \dot{q}^\gamma + \dot{q}^\beta \delta_{\alpha\gamma}) = \frac{1}{2} \sum_{\beta,\gamma=1}^s \lambda_{\beta\gamma} (\dot{q}^\beta \dot{q}^\gamma + \dot{q}^\beta \dot{q}^\gamma) = 2T.$$

Insgesamt ergibt sich daher

$$E = \frac{1}{2} \sum_{\beta,\gamma=1}^s \lambda_{\beta\gamma} \dot{q}^\beta \dot{q}^\gamma + V = T + V.$$

Man erkennt die bekannte Definition der Gesamtenergie.

III.3.1 b Homogenität des Raums und Impulserhaltung

In diesem Paragraphen wird der Zusammenhang zwischen Invarianz unter räumlichen Translationen und Impulserhaltung diskutiert.

Definition: Eine verallgemeinerte Koordinate q^α heißt *zyklische Koordinate*, wenn die Lagrange-Funktion nicht explizit davon abhängt, d.h.

$$\frac{\partial \mathcal{L}(t, \mathbf{q}, \dot{\mathbf{q}})}{\partial q^\alpha} = 0. \quad (\text{III.37})$$

Falls eine generalisierte Koordinate q^α zyklisch ist, dann ist die Lagrange-Funktion invariant unter den Translationen $q^\alpha \rightarrow q^\alpha + c$ mit $c \in \mathbb{R}$. Dann ist der dazu kanonisch konjugierte Impuls

$$p_\alpha \equiv \frac{\partial \mathcal{L}(t, \mathbf{q}, \dot{\mathbf{q}})}{\partial \dot{q}^\alpha} \quad (\text{III.17})$$

ausgewertet im Punkt $(t, \mathbf{q}(t), \dot{\mathbf{q}}(t))$ eine Konstante der Bewegung, d.h.

$$\frac{dp_\alpha}{dt} = 0. \quad (\text{III.38})$$

Beispiel: Sei ein System aus N Massenpunkten mit nur inneren Zentralkräften, die aus dem Potential

$$V(\vec{x}_1, \dots, \vec{x}_N) = \sum_{1 \leq a < b \leq N} V_{ab}(|\vec{x}_a - \vec{x}_b|)$$

abgeleitet werden können [vgl. Gl. (II.18)].

Wählt man als verallgemeinerte Koordinaten zum einen (die drei kartesischen Koordinaten von) \vec{x}_1 , und zum anderen die Relativkoordinaten $\vec{x}_{a1} \equiv \vec{x}_a - \vec{x}_1$, so lautet die Lagrange-Funktion des Systems

$$\mathcal{L} = T - V = \frac{m_1}{2} \dot{\vec{x}}_1^2 + \sum_{a=2}^N \frac{m_a}{2} [\dot{\vec{x}}_{a1} + \dot{\vec{x}}_1]^2 - \sum_{1 \leq a < b \leq N} V_{ab}(|\vec{x}_{a1} - \vec{x}_{b1}|).$$

Diese Lagrange-Funktion ist unabhängig von jeder der drei kartesischen Komponenten x_1^i von \vec{x}_1 , die somit zyklische Koordinaten sind. Demzufolge sind die zugehörigen verallgemeinerten Impulse p_i , $i = 1, 2, 3$ Erhaltungsgrößen:

$$p_i \equiv \frac{\partial \mathcal{L}}{\partial \dot{x}_1^i} = m_1 \dot{x}_1^i + \sum_{a=2}^N m_a (\dot{x}_{a1}^i + \dot{x}_1^i) = m_1 \dot{x}_1^i + \sum_{a=2}^N m_a \dot{x}_a^i,$$

wobei die Unabhängigkeit der Variablen \dot{x}_1^i und \dot{x}_{a1}^i benutzt wurde. Das heißt, die Erhaltungsgröße ist hier die i -te Komponente des Gesamtimpulses des Systems.

III.3.1 c Isotropie des Raums und Drehimpulserhaltung

Wir betrachten jetzt ein System aus Massenpunkten mit Positionen $\vec{x}_a(t)$, dessen Lagrange-Funktion \mathcal{L} invariant unter Drehungen ist. Eine solche Invarianz folgt aus der Isotropie des Raums in Abwesenheit eines externen Feldes — z.B. eines Magnet- oder Schwerefeldes —, das eine bevorzugte Richtung auswählen würde.

Sei \vec{p}_a der zu \vec{x}_a kanonisch konjugierte Impuls. Dann ist der Gesamtdrehimpuls

$$\vec{L} \equiv \sum_a \vec{x}_a(t) \times \vec{p}_a(t) \quad (\text{III.39})$$

eine Erhaltungsgröße, d.h.

$$\frac{d\vec{L}}{dt} = 0. \quad (\text{III.40})$$

Zum Beweis dieses Ergebnisses, betrachte man eine beliebige infinitesimale Drehung mit Winkel $\delta\varphi$ um eine Achse mit Einheitsvektor $\vec{e}_{\delta\varphi}$. Sei $\vec{\delta\varphi} \equiv \delta\varphi \vec{e}_{\delta\varphi}$. Die Änderung der Position eines Massenpunkts, die aus der Anwendung der Drehung resultiert, lautet dann

$$\delta\vec{x}_a = \vec{\delta\varphi} \times \vec{x}_a. \quad (\text{III.41})$$

Unter Nutzung der Kettenregel lautet die Änderung der Lagrange-Funktion für eine Variation $\delta\vec{x}_a$, $\delta\dot{\vec{x}}_a = d(\delta\vec{x}_a)/dt$ ihrer Variablen

$$\delta\mathcal{L} = \sum_a \sum_{i=1}^3 \left(\frac{\partial \mathcal{L}}{\partial x_a^i} \delta x_a^i + \frac{\partial \mathcal{L}}{\partial \dot{x}_a^i} \delta \dot{x}_a^i \right).$$

Der erste Term in den Klammern lässt sich mithilfe der Euler-Lagrange-Gleichung (III.11) umschreiben:

$$\delta\mathcal{L} = \sum_a \sum_{i=1}^3 \left[\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}_a^i} \right) \delta x_a^i + \frac{\partial \mathcal{L}}{\partial x_a^i} \delta x_a^i \right] = \sum_a \sum_{i=1}^3 \left(\frac{dp_{a,i}}{dt} \delta x_a^i + p_{a,i} \delta \dot{x}_a^i \right),$$

wobei in der zweiten Gleichung die verallgemeinerten Impulse (III.17) eingeführt wurden. Dabei erkennt man in den Klammern die Zeitableitung des Produkts $p_{a,i} \delta x_a^i$, so dass die Summe über i einfach einem Skalarprodukt entspricht:

$$\delta\mathcal{L} = \sum_a \frac{d}{dt} (\vec{p}_a \cdot \delta\vec{x}_a) = \frac{d}{dt} \left(\sum_a \vec{p}_a \cdot \delta\vec{x}_a \right).$$

Fordert man dann, dass die Lagrange-Funktion invariant unter Drehungen bleibt, so ist dieser Variation $\delta\mathcal{L}$ Null für Variationen $\delta\vec{x}_a$ der Form (III.41), d.h.

$$\frac{d}{dt} \left[\sum_a \vec{p}_a \cdot (\vec{\delta\varphi} \times \vec{x}_a) \right] = \frac{d}{dt} \left[\sum_a \vec{\delta\varphi} \cdot (\vec{x}_a \times \vec{p}_a) \right] = 0,$$

wobei die erste Gleichheit aus der Invarianz des Spatprodukts unter zyklischen Permutationen folgt.

Da die Drehung zeitunabhängig ist, kann $\vec{\delta\varphi}$ aus der Ableitung herausgezogen werden. Daher soll die Gleichung

$$\vec{\delta\varphi} \cdot \frac{d}{dt} \left[\sum_a (\vec{x}_a \times \vec{p}_a) \right] = 0$$

für eine beliebige infinitesimale Drehung, d.h. ein beliebiges $\vec{\delta\varphi}$, gelten: Dies ist nur dann möglich, wenn das zweite Multiplikand des Skalarprodukts verschwindet, was genau Gl. (III.41) entspricht.

Bemerkung: Falls ein System nicht invariant unter allen Drehungen ist, sondern nur unter den Drehungen um eine gegebene Achse, dann ist nur die Komponente des Drehimpulses entlang dieser Richtung erhalten.

III.3.2 Noether-Theorem

Die oben gefundenen Zusammenhänge zwischen Invarianz unter Zeittranslationen und Energieerhaltung, Invarianz unter räumlichen Translationen und Impulserhaltung, oder Invarianz unter Drehungen und Drehimpulserhaltung, sind alle Beispiele eines allgemeineren Theorems, die durch Emmy Noether^(s) erfunden wurde. Hiernach wird eine mathematisch nicht-rigoreuse Version des Satzes dargestellt, ohne genaue Angabe der Annahmen und ohne Beweis.

III.3.2 a Symmetrien

Definition: Ein physikalisches System sei beschrieben durch verallgemeinerte Koordinaten \mathbf{q} . Ist seine Wirkung $S[\mathbf{q}]$ invariant unter einer bestimmten Koordinatentransformation

$$t \rightarrow t'(t, \mathbf{q}, \dot{\mathbf{q}}) \quad , \quad \mathbf{q} \rightarrow \mathbf{q}'(t, \mathbf{q}, \dot{\mathbf{q}}), \quad (\text{III.42})$$

so heißt die letztere *Symmetrietransformation*. Dann wird gesagt, dass das System die entsprechende *Symmetrie* besitzt.

Bemerkung: Dabei wird die Invarianz der Wirkung erfordert, nicht jene der Lagrange-Funktion, die eine stärkere Forderung bildet.

Beispielsweise wurden in § III.3.1 verschiedene Symmetrietransformationen erwähnt, die die Wirkung physikalischer Systeme oft invariant lassen: Zeit- und Raumtranslationen, Drehungen. Dabei handelt es sich um sog. *kontinuierliche Symmetrien*, denn jede davon lässt sich durch einen Parameter beschreiben — Translationsparameter, Winkel der Drehung —, der seine Werte in einer „kontinuierlichen Menge“⁽³⁷⁾ annehmen kann.

Im Gegensatz dazu ist eine Punktspiegelung — wie z.B. die *Zeitumkehr* $t \rightarrow t' \equiv -t$ oder die *Raumspiegelung* $\vec{r} \rightarrow \vec{r}' \equiv -\vec{r}$ — keine kontinuierliche, sondern eine *diskrete Symmetrie* (-Transformation).

Im Fall einer kontinuierlichen Symmetrie existieren auch infinitesimale Transformationen

$$q^\alpha \rightarrow q^{\alpha'} = q^\alpha + \varepsilon Q^\alpha(t, \mathbf{q}, \dot{\mathbf{q}}) \quad \text{mit } |\varepsilon| \ll 1 \quad (\text{III.43})$$

der Koordinaten, die um den kleinen Parameter ε von der Identität abweichen. Dabei heißen die Funktionen Q^α die *Generatoren* der Koordinatentransformationen.

III.3.2 b Noether-Theorem

Ausgehend von kontinuierlichen Symmetrietransformationen kann man einen wichtigen Satz der mathematischen Physik beweisen, dessen Inhalt wir jetzt nur relativ informell formulieren.

Theorem (Noether-Theorem):

Mit jeder kontinuierlichen Symmetrie eines physikalischen Systems kann eine Erhaltungsgröße, die Noether-Ladung, assoziiert werden. (III.44)

⁽³⁷⁾D.h. eine Menge, die homöomorph zu einem zusammenhängenden Intervall bzw. Gebiet von \mathbb{R} bzw. \mathbb{R}^n ist.

^(s)E. NOETHER, 1882–1935

Dazu erklärt noch der genauere Satz, wie die Erhaltungsgröße für eine bestimmte Symmetrie aussieht bzw. wie sie sich berechnen lässt. Die zugehörige Form werden wir jetzt finden, indem wir plausible Elemente eines Beweises darstellen.⁽³⁸⁾

Dafür soll man eine infinitesimale Symmetrietransformation betrachten, hier der Einfachheit halber mit $t' = t$. Neben Gl. (III.43) gilt noch für die generalisierten Geschwindigkeiten

$$\dot{q}^\alpha \rightarrow \dot{q}^{\alpha'} = \dot{q}^\alpha + \varepsilon \dot{Q}^\alpha(t, \mathbf{q}, \dot{\mathbf{q}}). \quad (\text{III.45})$$

Die Änderung der Lagrange-Funktion

$$\delta \mathcal{L}(t) \equiv \mathcal{L}(t, \mathbf{q}'(t), \dot{\mathbf{q}}'(t)) - \mathcal{L}(t, \mathbf{q}(t), \dot{\mathbf{q}}(t)),$$

die aus den Transformationen (III.43), (III.45) folgt, kann auf zwei alternative Weisen geschrieben werden. Einerseits gilt

$$\delta \mathcal{L}(t) = \varepsilon \frac{df(t, \mathbf{q}(t), \dot{\mathbf{q}}(t))}{dt} \quad (\text{III.46})$$

mit einer Funktion f , deren genaue Form von jener der Lagrange-Funktion \mathcal{L} und von der Symmetrie abhängt. Der Faktor ε — es könnte theoretisch auch eine ganzzahlige Potenz davon sein — folgt daraus, dass die Variation $\delta \mathcal{L}$ verschwinden muss, wenn ε gegen Null geht. Dann multipliziert dieser Faktor eine Funktion der verallgemeinerten Koordinaten und Geschwindigkeiten, die sich in der Form einer totalen Zeitableitung schreiben lässt, denn ein solcher Ansatz ist erlaubt, weil er die Wirkung invariant lässt: $S[\mathbf{q}] = S[\mathbf{q}']$, entsprechend der Forderung einer Symmetrietransformation.

Andererseits liefert die Kettenregel für die Variation von \mathcal{L} unter kleinen Variationen $\{\delta q^\alpha\}$, $\{\delta \dot{q}^\alpha\}$ ihrer Argumente die Gleichung

$$\delta \mathcal{L} = \sum_{\alpha=1}^s \left(\frac{\partial \mathcal{L}}{\partial q^\alpha} \delta q^\alpha + \frac{\partial \mathcal{L}}{\partial \dot{q}^\alpha} \delta \dot{q}^\alpha \right) = \varepsilon \sum_{\alpha=1}^s \left(\frac{\partial \mathcal{L}}{\partial q^\alpha} Q^\alpha + \frac{\partial \mathcal{L}}{\partial \dot{q}^\alpha} \dot{Q}^\alpha \right),$$

wobei in der zweiten Gleichung die aus den Transformationen (III.43), (III.45) folgenden Variationen δq^α , $\delta \dot{q}^\alpha$ benutzt wurden. Die Euler–Lagrange-Gleichungen (III.11) und die Produktregel geben dann

$$\delta \mathcal{L} = \varepsilon \sum_{\alpha=1}^s \left[\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}^\alpha} \right) Q^\alpha + \frac{\partial \mathcal{L}}{\partial \dot{q}^\alpha} \dot{Q}^\alpha \right] = \varepsilon \sum_{\alpha=1}^s \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}^\alpha} Q^\alpha \right),$$

d.h. noch

$$\delta \mathcal{L} = \varepsilon \frac{d}{dt} \left(\sum_{\alpha=1}^s \frac{\partial \mathcal{L}}{\partial \dot{q}^\alpha} Q^\alpha \right). \quad (\text{III.47})$$

Subtrahiert man schließlich Gl. (III.46) von Gl. (III.47), so ergibt sich

$$\varepsilon \frac{dQ}{dt} = 0 \quad \text{mit} \quad Q \equiv \sum_{\alpha=1}^s \frac{\partial \mathcal{L}}{\partial \dot{q}^\alpha} Q^\alpha - f. \quad (\text{III.48})$$

Da die linke Gleichung für jede infinitesimale Transformation, d.h. für ein beliebiges ε , gilt, muss die hier definierte Noether-Ladung Q eine Konstante der Bewegung sein. \square

Beispiel 1: Räumliche Translation

Sei q^β eine zyklische Koordinate (vgl. § III.3.1 b). Dann ist die Lagrange-Funktion des Systems invariant unter Translationen $q^\beta \rightarrow q^{\beta'} = q^\beta + c$ dieser Koordinate, wobei $c = \varepsilon$ für eine infinitesimale Transformation. Für die anderen Koordinaten q^α mit $\alpha \neq \beta$ betrachten wir die Identitätstransformation $q^\alpha \rightarrow q^{\alpha'} = q^\alpha$.

Identifiziert man diese Transformationen mit Gl. (III.43), so ist der Generator in diesem Fall einfach $Q^\alpha = \delta_{\alpha\beta}$, woraus nach Ableitung $\dot{Q}^\alpha = 0$ für alle α folgt. Wegen der angenommenen Zyklizität

⁽³⁸⁾Ein ausführlicher Beweis „für Physiker“ befindet sich in (fast) jedem Lehrbuch zur theoretischen Mechanik. Wer mathematische Genauigkeit sehen möchte, kann sie z.B. im Kap. 4 von Arnold, *Mathematical methods of classical mechanics* [1] finden.

der Koordinate ist die Lagrange-Funktion \mathcal{L} invariant unter einer (infinitesimalen) Transformation, d.h. $f = 0$ in Gl. (III.46).

Laut Gl. (III.48) lautet die zugehörige Erhaltungsgröße

$$Q = \sum_{\alpha=1}^s \frac{\partial \mathcal{L}}{\partial \dot{q}^\alpha} \delta_{\alpha\beta} = p_\beta,$$

d.h. wir finden noch einmal, dass die Invarianz unter räumlichen Transformationen die Erhaltung des Impulses impliziert.

Beispiel 2: Zeitliche Translation

Als zweites Beispiel betrachten wir ein System, dessen Wirkung invariant unter den gleichzeitigen Transformationen

$$t \rightarrow t' = t + \tau \quad , \quad q^\alpha \rightarrow q^{\alpha'}(t') = q^\alpha(t)$$

ist, wobei $\tau \in \mathbb{R}$. Dank einer Taylor-Entwicklung gilt $q^{\alpha'}(t') = q^\alpha(t + \varepsilon) = q^\alpha + \varepsilon \dot{q}^\alpha + \mathcal{O}(\varepsilon^2)$, so dass die Generatoren der (infinitesimalen) Transformationen [Gl. (III.43)] durch $Q^\alpha = \dot{q}^\alpha$ gegeben sind.

Da $t' = t + \varepsilon$ gibt eine zweite Taylor-Entwicklung

$$\delta \mathcal{L} = \mathcal{L}(t', \mathbf{q}'(t), \dot{\mathbf{q}}'(t)) - \mathcal{L}(t, \mathbf{q}(t), \dot{\mathbf{q}}(t)) = \varepsilon \frac{d\mathcal{L}(t, \mathbf{q}(t), \dot{\mathbf{q}}(t))}{dt} + \mathcal{O}(\varepsilon^2),$$

d.h. $f = \mathcal{L}$ in Gl. (III.47). Insgesamt ist die erhaltene Noether-Ladung gegeben durch

$$Q = \sum_{\alpha=1}^s \frac{\partial \mathcal{L}}{\partial \dot{q}^\alpha} \dot{q}^\alpha - \mathcal{L} = E.$$

Somit wird das Ergebnis des § III.3.1 a wieder gefunden.

Literatur zum Kapitel III

- Arnold, *Mathematical methods of classical mechanics* [1] Teil II.
- Feynman, *Vorlesungen über Physik. Band 2* [22] = *Lectures on Physics. Vol. II* [23], Kap. 19.
- Fließbach, *Mechanik* [2] Teil II, Kap. 7–10 & Teil III, Kap. 12–14.
- Goldstein, *Klassische Mechanik* [4] = *Classical Mechanics* [5], Kap. 1.3–1.6, 2.1–2.3 & 2.5.
- Greiner, *Klassische Mechanik II* [7] Kap. V.
- Landau & Lifschitz, *Mechanik* [13], Kap. I.
- Nolting, *Analytische Mechanik* [16] Kap. 1.
- Scheck, *Mechanik* [18] Kap. 2.1–2.10.

KAPITEL IV

Lagrange-Formalismus: Anwendungen

IV.1 Kleine Schwingungen 87

IV.1.1 Eindimensionales Problem 87

IV.1.2 Multidimensionales Problem 90

IV.2 Starre Körper 94

IV.2.1 Beschreibung des starren Körpers 94

IV.2.2 Bewegungsgleichungen 97

In diesem Kapitel werden einige längeren Anwendungen des in Kap III eingeführten Lagrange-Formalismus vorgestellt.

IV.1 Kleine Schwingungen

In der Physik, oder in der mathematischen Modellierung eines physikalischen Systems, kommt oft das folgende Problem vor. Gegeben ein mechanisches System, beschrieben durch eine Lagrange-Funktion $\mathcal{L}(t, \mathbf{q}, \dot{\mathbf{q}})$, ist die Lösung der Bewegungsgleichungen für bestimmte Anfangsbedingungen bekannt. Sei $\mathbf{q}^{(0)}(t)$ diese Lösung. Das System — in der Praxis entweder das gleiche oder eine Kopie davon, wobei alle relevanten Nebenbedingungen gleich bleiben — wird irgendwie leicht „gestört“, entsprechend z.B. einer kleiner Änderung der Anfangsbedingungen. Die Frage ist dann, was ist die neue Lösung $\mathbf{q}(t)$ der Bewegungsgleichungen für das gestörte System?

Intuitiv kann man sich vorstellen, dass in manchen Fällen $\mathbf{q}(t)$ „in der Nähe“⁽³⁹⁾ der ungestörten Lösung $\mathbf{q}^{(0)}(t)$ bleibt. Andererseits kann eine kleine Störung auch manchmal zu einer großen Änderung führen, wie z.B. wenn eine Kugel auf dem schmalen Gipfel eines Hügels ein wenig verschoben ist, und damit anfängt, hinabzurollen.

Dieser Abschnitt befasst sich mit diesem Problem der kleinen Auslenkungen aus einer Bezugslösung, erstens für ein System mit einem einzigen Freiheitsgrad (§ IV.1.1), dann für Systeme mit mehreren Freiheitsgraden (§ IV.1.2). Dabei werden die beiden oben diskutierten Verhalten gefunden. Im letzteren Fall mit $s > 1$ Freiheitsgraden wird auch gezeigt, dass auch wenn diese unter einander gekoppelt sind, jedoch kann man immer s Linearkombinationen ihrer kleinen Variationen finden, die sich unabhängig von einander bewegen.

Der Einfachheit halber beschränken sich die Diskussionen auf Systeme mit einer zeitunabhängigen Lagrange-Funktion und mit konservativen Kräften. Darüber hinaus wird angenommen, dass die Referenzlösung eine „Gleichgewichtslösung“ der Euler-Lagrange-Gleichungen darstellt, d.h. sie ist stationär ($\mathbf{q}^{(0)}$ ist zeitunabhängig), woraus sofort $\dot{\mathbf{q}}^{(0)}(t) = 0$ folgt.

IV.1.1 Eindimensionales Problem

Wir betrachten zuerst ein physikalisches System mit einem einzigen Freiheitsgrad q . Dies kann z.B. die eindimensionale Bewegung eines Massenpunkts sein, oder eine mehrdimensionale Bewegung

⁽³⁹⁾Um diese qualitative Redensart mehr quantitativ auszudrücken, sollte man einen Abstand in einem Funktionenraum einführen, was hier nicht gemacht wird.

mit genug Zwangsbedingungen, damit am Ende nur $s = 1$ Freiheitsgrad übrig bleibt.

Es wird angenommen, dass das System invariant unter zeitlichen Translationen ist, d.h. dass seine Lagrange-Funktion \mathcal{L} nicht von der Zeit abhängt. Für diese wird die Form

$$\mathcal{L}(q, \dot{q}) = \frac{\lambda(q)}{2} \dot{q}^2 - V(q) \quad (\text{IV.1})$$

angenommen mit einer kinetischen Energie, die quadratisch von der generalisierten Geschwindigkeit abhängt. Dabei sind λ und V genug reguläre Funktionen einer reellen Variablen.

Sei $q^{(0)}(t) \equiv q_0$ mit $q_0 \in \mathbb{R}$ eine stationäre Lösung der assoziierten Euler-Lagrange-Gleichung, d.h. $\dot{q}^{(0)}(t) = 0$. Es wird angenommen, dass $\lambda(q_0)$ ungleich Null ist. Damit die konstante Funktion $q^{(0)}$ wirklich eine Gleichgewichtslösung darstellt, soll das Potential ein Extremum bei q_0 haben

$$V'(q_0) = 0, \quad (\text{IV.2})$$

wobei der Strich die Ableitung nach der verallgemeinerten Koordinate bezeichnet. D.h., die generalisierte Kraft verschwindet in diesem Punkt.

Die Euler-Lagrange-Gleichung (III.11) lautet hier

$$\lambda(q(t)) \ddot{q}(t) + \lambda'(q(t)) \dot{q}(t)^2 = \frac{\lambda'(q(t))}{2} \dot{q}(t)^2 - V'(q(t)),$$

so dass $\dot{q}^{(0)}(t) = 0$ und daher $\ddot{q}^{(0)}(t) = 0$ automatisch zur Bedingung (IV.2) führen.

Wir betrachten nun eine kleine, zeitabhängige Variation der generalisierten Koordinate um den Gleichgewichtswert

$$q(t) \equiv q^{(0)}(t) + \delta q(t) = q_0 + \delta q(t), \quad (\text{IV.3})$$

wobei $\delta q(t)$ „klein“ sein soll, d.h. höhere Potenzen $[\delta q(t)]^n$ sollen viel kleiner als die niedrigeren Ordnungen sein. Die Zeitableitung $\delta \dot{q}(t)$ wird als kleine Größe der gleichen Ordnung wie $\delta q(t)$ betrachtet.

Die Taylor-Entwicklung der Lagrange-Funktion (IV.1) um den Punkt ($q = q_0, \dot{q} = 0$) bis zur zweiten Ordnung lautet

$$\begin{aligned} \mathcal{L}(q_0 + \zeta, \eta) &= \frac{\lambda(q_0 + \zeta)}{2} \eta^2 - V(q_0 + \zeta) \\ &= \frac{\lambda(q_0)}{2} \eta^2 - V(q_0) - V'(q_0) \zeta - \frac{V''(q_0)}{2} \zeta^2 + \mathcal{O}(\zeta^3, \zeta^2 \eta, \zeta \eta^2, \eta^3). \end{aligned} \quad (\text{IV.4})$$

Dabei tragen die höheren Terme in ζ in der Taylor-Entwicklung von $\lambda(q_0 + \zeta)$ nicht bei, weil η^2 schon zweiter Ordnung ist. Mit $\zeta = \delta q(t)$, $\eta = \delta \dot{q}$ und unter Berücksichtigung der Bedingung (IV.2) ergibt sich

$$\mathcal{L}(q_0 + \delta q, \delta \dot{q}) \simeq \frac{\lambda(q_0)}{2} (\delta \dot{q})^2 - V(q_0) - \frac{V''(q_0)}{2} (\delta q)^2, \quad (\text{IV.5})$$

wobei die vernachlässigten Terme der Ordnung $\mathcal{O}((\delta q)^3)$ sind. D.h. die Lagrange-Funktion ist quadratisch in der verallgemeinerten Koordinate und der zugehörigen Geschwindigkeit, ohne Mischterm. Der zweite Term auf der rechten Seite ist eine Konstante, und trägt daher nicht zu den Bewegungsgleichungen bei.

Die Aussage „ $\delta q(t)$ soll klein sein“ ist in der Tat sehr schlampig: da $q(t)$, und somit $\delta q(t)$, im Allgemeinen eine physikalische Dimension besitzen kann, hängt sein Zahlenwert von der gewählten Einheit ab, so dass $\delta q(t)$ nicht „klein“ sein kann, sondern nur klein gegenüber einem Referenzwert. Eine mathematisch mehr präzise Formulierung lautet wie folgt.

Sei $q_1(t)$ beliebig und $|\varepsilon| \ll 1$: dann darf ε „klein“ genannt werden. Setzt man nun $\zeta = \varepsilon q_1(t)$ und $\eta = \varepsilon \dot{q}_1(t)$ in Eq. (IV.4) ein, so findet man

$$\mathcal{L}(q_0 + \varepsilon q_1, \varepsilon \dot{q}_1) \sim \frac{\lambda(q_0)}{2} (\varepsilon \dot{q}_1)^2 - V(q_0) - V'(q_0) \varepsilon q_1 - \frac{V''(q_0)}{2} (\varepsilon q_1)^2 + \mathcal{O}(\varepsilon^2).$$

Mit $\varepsilon q_1 = \delta q$ und $\varepsilon \dot{q}_1 = \delta \dot{q}$ ist dies genau Gl. (IV.5).

Ausgehend von der Lagrange-Funktion (IV.5) und unter Verwendung von $\partial/\partial q = \partial/\partial(\delta q)$ und $\partial/\partial \dot{q} = \partial/\partial(\delta \dot{q})$ lauten die partiellen Ableitungen auf beiden Seiten der Euler-Lagrange-Gleichung (III.11)

$$\frac{\partial \mathcal{L}}{\partial q} = -V''(q_0)\delta q, \quad \frac{\partial \mathcal{L}}{\partial \dot{q}} = \lambda(q_0)\delta \dot{q}.$$

Nach einer Zeitableitung des letzteren Terms lautet die Bewegungsgleichung schließlich

$$\delta \ddot{q}(t) = -\frac{V''(q_0)}{\lambda(q_0)}\delta q(t). \quad (\text{IV.6})$$

Je nach dem Zahlenwert des konstanten Verhältnisses $V''(q_0)/\lambda(q_0) > 0$ — oder eigentlich von $V''(q_0)$, weil $\lambda(q_0)$ in physikalischen Problemen immer nicht-negativ ist, wie hiernach angenommen wird — lassen sich drei Fälle unterscheiden.

- Für $V''(q_0) > 0$ kann man $\omega^2 \equiv V''(q_0)/\lambda(q_0)$ definieren. Dann ist Gl. (IV.6) die Bewegungsgleichung eines *harmonischen Oszillators* mit Kreisfrequenz ω .

Die allgemeine Lösung der Gleichung

$$\delta q(t) = A \cos(\omega t) + B \sin(\omega t)$$

mit Konstanten A und B , die von den Anfangsbedingungen abhängen, bleibt begrenzt: das System führt kleine Schwingungen um die Gleichgewichtslage q_0 durch, die somit *stabil* ist.

- Für $V''(q_0) < 0$ kann man $\kappa^2 \equiv -V''(q_0)/\lambda(q_0)$ einführen. Dann ist die allgemeine Lösung der Bewegungsgleichung der Form

$$\delta q(t) = A e^{\kappa t} + B e^{-\kappa t}.$$

Einer der beiden Terme — der erste falls $\kappa > 0$ ist — wächst exponentiell schnell mit der Zeit und ist daher unbegrenzt, d.h. bleibt nicht in der Nähe der Gleichgewichtslösung, die *instabil* ist.

Wegen des Wachstums wird $\delta q(t)$ nicht „klein“ bleiben, so dass die Näherungen, die zur Bewegungsgleichung führen, nicht mehr gelten. Anstatt unendlich zu wachsen, wird möglicherweise eine andere (Gleichgewichts)Lösung in endlichem Abstand von q_0 gefunden.

- Für $V''(q_0) = 0$ führt die approximative Bewegungsgleichung (IV.6) zu einer Bewegung mit konstanter Geschwindigkeit, d.h. wieder unbegrenzt. Jedoch die höheren Terme in der Taylor-Entwicklung könnten dieses Verhalten ändern.

Beispiel: Einfaches ebenes Pendel

In § III.2.4 b wurde das ebene mathematische Pendel (Abb. III.2) schon ausführlich diskutiert. Zur Erinnerung lautet dessen (zeitunabhängige) Standard-Lagrange-Funktion [Gl. (III.29)]

$$\mathcal{L}(\varphi, \dot{\varphi}) = T(\dot{\varphi}) - V(\varphi) = \frac{ml^2}{2}\dot{\varphi}^2 + mgl \cos \varphi, \quad (\text{IV.7})$$

wobei φ der Ablenkwinkel aus der z -Richtung ist. Stationäre Lösungen der zugehörigen Bewegungsgleichung (III.30) sind $\varphi^{(0)}(t) \equiv 0$ und $\varphi^{(0)}(t) \equiv \pi$, entsprechend den Nullstellen der „verallgemeinerten Kraft“ $-dV(\varphi)/d\varphi = -mgl \sin \varphi$ im Intervall $-\pi < \varphi \leq \pi$.

Das Verhalten des Pendels für kleine Abweichungen aus der Gleichgewichtsposition $\varphi^{(0)} = 0$ sind wohl bekannt: Für kleine Winkel gilt $V(\varphi) \simeq -mgl(1 - \varphi^2/2)$, d.h. die Lagrange-Funktion (IV.7) wird zu

$$\mathcal{L}(\varphi, \dot{\varphi}) \simeq \frac{ml^2}{2}\dot{\varphi}^2 - \frac{mgl}{2}\varphi^2 + mgl,$$

wobei der letzte Term auf der rechten Seite irrelevant für die Bewegungsgleichung ist. Diese (approximative) Lagrange-Funktion ist quadratisch in φ und $\dot{\varphi}$, d.h. der Form (IV.5), und führt somit zu

einer Bewegungsgleichung [vgl. (IV.6)] der Form $\ddot{\varphi} \propto \varphi$ mit einem hier negativen Proportionalitätsfaktor $-\omega^2 \equiv -g/l$, so dass die Lösungen kleine Schwingungen sind.

Betrachten wir jetzt Abweichungen aus der Gleichgewichtsposition $\varphi^{(0)} = \pi$, d.h. wir schreiben $\varphi = \pi + \delta\varphi$. Dann gilt $V(\pi + \delta\varphi) \simeq V(\pi) + V''(\pi)(\delta\varphi)^2/2 = mgl - mgl(\delta\varphi)^2/2$, und daher

$$\mathcal{L}(\pi + \varphi, \delta\dot{\varphi}) \simeq \frac{ml^2}{2}\delta\dot{\varphi}^2 + \frac{mgl}{2}(\delta\varphi)^2 - mgl.$$

Diese Lagrange-Funktion liefert die Bewegungsgleichung $\ddot{\varphi}(t) = \kappa^2\varphi(t)$ mit einem positiven Koeffizienten $\kappa^2 \equiv g/l$. Die Lösungen dieser Differentialgleichung sind nicht kleine Schwingungen, sondern wachsen exponentiell schnell mit der Zeit. Somit finden wir wieder, dass die Gleichstellungsposition $\varphi = \pi + \delta\varphi$ instabil ist, so dass kleine Abweichungen daraus sich davon (schnell) entfernen.

IV.1.2 Multidimensionales Problem

Sei jetzt ein physikalisches System mit $s > 1$ Freiheitsgraden, denen verallgemeinerte Koordinaten q^1, \dots, q^s — kollektiv mit \mathbf{q} bezeichnet — zugeordnet sind. Für die Lagrange-Funktion wird die zeitunabhängige Form

$$\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2} \sum_{\alpha, \beta=1}^s \lambda_{\alpha\beta}(\mathbf{q}) \dot{q}^\alpha \dot{q}^\beta - V(\mathbf{q}) \quad (\text{IV.8})$$

angenommen, wobei $\lambda_{\alpha\beta}$ symmetrisch unter dem Austausch von α und β ist.

Falls $\lambda_{\alpha\beta}$ nicht symmetrisch ist, trägt eigentlich nur der symmetrische Anteil $\frac{1}{2}(\lambda_{\alpha\beta} + \lambda_{\beta\alpha})$ zur kinetischen Energie bei, während der antisymmetrische Anteil $\frac{1}{2}(\lambda_{\alpha\beta} - \lambda_{\beta\alpha})$ sich in der Summe über alle Werte von α und β kürzt.

Sei $\mathbf{q}^{(0)}(t) \equiv \mathbf{q}_0 \equiv \{q_0^\alpha\}_{\alpha=1, \dots, s}$ eine Gleichgewichtslösung, d.h. eine stationäre Lösung der Bewegungsgleichungen. Dann hat das Potential V automatisch ein Extremum bei \mathbf{q}_0 , d.h.

$$\frac{\partial V(\mathbf{q}_0)}{\partial q^\alpha} = 0 \quad \forall \alpha \in \{1, \dots, s\}, \quad (\text{IV.9})$$

so dass die verallgemeinerte Kraft für die Gleichgewichtslösung Null ist.

Betrachtet man jetzt eine kleine Variation

$$q^\alpha(t) = q_0^\alpha + \delta q^\alpha(t) \quad (\text{IV.10})$$

der Trajektorie im Konfigurationsraum, so gibt eine Taylor-Entwicklung der Lagrange-Funktion bis zur zweiten Ordnung in den δq^α und $\delta \dot{q}^\alpha$

$$\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2} \sum_{\alpha, \beta=1}^s \lambda_{\alpha\beta}(\mathbf{q}_0) \delta \dot{q}^\alpha \delta \dot{q}^\beta - V(\mathbf{q}_0) - \sum_{\alpha=1}^s \frac{\partial V(\mathbf{q}_0)}{\partial q^\alpha} \delta q^\alpha - \sum_{\alpha, \beta=1}^s \frac{1}{2} \frac{\partial^2 V(\mathbf{q}_0)}{\partial q^\alpha \partial q^\beta} \delta q^\alpha \delta q^\beta.$$

Unter Verwendung der Bedingungen (IV.9) und nach Einführung der kurzen Notationen

$$m_{\alpha\beta} \equiv \lambda_{\alpha\beta}(\mathbf{q}_0), \quad k_{\alpha\beta} \equiv \frac{\partial^2 V(\mathbf{q}_0)}{\partial q^\alpha \partial q^\beta} \quad (\text{IV.11a})$$

bleibt

$$\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}) \simeq \sum_{\alpha, \beta=1}^s \frac{m_{\alpha\beta}}{2} \delta \dot{q}^\alpha \delta \dot{q}^\beta - \sum_{\alpha, \beta=1}^s \frac{k_{\alpha\beta}}{2} \delta q^\alpha \delta q^\beta - V(\mathbf{q}_0). \quad (\text{IV.11b})$$

Wie in § IV.1.1 ist die Lagrange-Funktion quadratisch in den verallgemeinerten Koordinaten und den zugehörigen Geschwindigkeiten ohne Mischterm zwischen den beiden. Dafür sind die verschiedenen Koordinaten miteinander gekoppelt und das gleiche gilt für die Geschwindigkeiten.

Die Lagrange-Funktion (IV.11b) kann noch in Matrixschreibweise geschrieben werden. Sei \mathbf{m} bzw. \mathbf{k} die $s \times s$ -Matrix mit Elementen $m_{\alpha\beta}$ bzw. $k_{\alpha\beta}$. Beide Matrizen sind symmetrisch, d.h. $\mathbf{m} = \mathbf{m}^\top$ und $\mathbf{k} = \mathbf{k}^\top$: einerseits wurde angenommen, dass die $f_{\alpha\beta}$ symmetrisch sind, woraus die Symmetrie von \mathbf{m} folgt. Andererseits sind die $k_{\alpha\beta}$ die zweiten Ableitungen einer Funktion — die wir implizit als mindestens zweimal kontinuierlich differenzierbar annehmen —, und somit ebenfalls symmetrisch.

Führt man nun die s -dimensionalen Spaltenvektoren $\delta\mathbf{q}$ bzw. $\delta\dot{\mathbf{q}}$ mit Komponenten $\{\delta q^\alpha\}$ bzw. $\{\delta\dot{q}^\alpha\}$ ein, und dementsprechend die Zeilenvektoren $\delta\mathbf{q}^\top$ und $\delta\dot{\mathbf{q}}^\top$, so lässt sich Gl. (IV.11b) noch in der kürzeren Form

$$\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}) \simeq \frac{1}{2} \delta\dot{\mathbf{q}}^\top \mathbf{m} \delta\dot{\mathbf{q}} + \frac{1}{2} \delta\mathbf{q}^\top \mathbf{k} \delta\mathbf{q} - V(\mathbf{q}_0) \quad (\text{IV.11c})$$

schreiben. Ab jetzt werden wir die Konstante $-V(\mathbf{q}_0)$, die keine Rolle in den Bewegungsgleichungen spielt, weglassen.

Als symmetrische reelle Matrizen sind \mathbf{m} und \mathbf{k} diagonalisierbar mit reellen Eigenwerten. Im Folgenden wird gezeigt, dass sie in der Tat gleichzeitig diagonalisierbar sind, und zwar auf Kosten einer geeigneten Wahl von verallgemeinerten Koordinaten.

Zuerst kann man die Matrix \mathbf{m} diagonalisieren: da \mathbf{m} symmetrisch ist, existieren eine orthogonale $s \times s$ -Matrix \mathcal{R} und eine diagonale Matrix

$$\mathbf{m}_D \equiv \begin{pmatrix} m_1 & 0 & \dots & 0 \\ 0 & m_2 & 0 & \dots \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \dots & 0 & m_{s-1} \\ 0 & \dots & \dots & 0 & m_s \end{pmatrix} \equiv \text{diag}(m_1, \dots, m_s)$$

derart, dass sie

$$\mathbf{m}_D = \mathcal{R} \mathbf{m} \mathcal{R}^\top$$

erfüllen, wobei die Eigenwerte $\{m_\alpha\}$ alle reell sind. Eigentlich werden in praktischen Fällen die $\{m_\alpha\}$ sogar positiv sein: $m_\alpha > 0$. Da \mathcal{R} eine Drehmatrix ist, gilt für ihre Inverse $\mathcal{R}^{-1} = \mathcal{R}^\top$, und man hat äquivalent

$$\mathbf{m} = \mathcal{R}^\top \mathbf{m}_D \mathcal{R},$$

so dass nach Einsetzen in Gl. (IV.11c)

$$\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2} \delta\dot{\mathbf{q}}^\top \mathcal{R}^\top \mathbf{m}_D \mathcal{R} \delta\dot{\mathbf{q}} + \frac{1}{2} \delta\mathbf{q}^\top \mathbf{k} \delta\mathbf{q}.$$

Sei $\delta\mathbf{q}' \equiv \mathcal{R} \delta\mathbf{q}$, entsprechend einem Koordinatenwechsel — hier handelt es sich um eine einfache Drehung — im Raum der verallgemeinerten Koordinaten. Da die Matrix \mathbf{m} , wie ihre Elemente, nicht von der Zeit abhängt, ist es auch der Fall der Drehmatrix \mathcal{R} . Somit gilt $\delta\dot{\mathbf{q}}' = \mathcal{R} \delta\dot{\mathbf{q}}$. Dann erhält man für die Lagrange-Funktion in den neuen Variablen

$$\mathcal{L}(\delta\mathbf{q}', \delta\dot{\mathbf{q}}') = \frac{1}{2} \delta\dot{\mathbf{q}}'^\top \mathbf{m}_D \delta\dot{\mathbf{q}}' + \frac{1}{2} \delta\mathbf{q}'^\top \mathcal{R} \mathbf{k} \mathcal{R}^\top \delta\mathbf{q}'. \quad (\text{IV.12})$$

Drückt man den kinetischen Term durch die (Zeitableitungen der) Komponenten $\{\delta q'^\alpha\}$ aus, die Linearkombinationen der ursprünglichen verallgemeinerten Koordinaten $\{\delta q^\beta\}$ sind, so sieht man explizit, dass die kinetische Energie jetzt keinen Mischterm enthält:

$$\mathcal{L}(\delta\mathbf{q}', \delta\dot{\mathbf{q}}') = \sum_{\alpha=1}^s \frac{m_\alpha}{2} (\delta\dot{q}'^\alpha)^2 + \frac{1}{2} \delta\mathbf{q}'^\top \mathcal{R} \mathbf{k} \mathcal{R}^\top \delta\mathbf{q}'.$$

Dagegen wird der Potential-Term im Allgemeinen noch nicht-diagonal sein.

Um den letzteren zu diagonalisieren, ohne dabei die Diagonalform der kinetischen Energie zu zerstören, kann man nicht direkt eine zweite Drehung durchführen. Stattdessen soll man zunächst

neue verallgemeinerte Koordinaten $Q^{\alpha'} \equiv \sqrt{m_{\alpha}} \delta q^{\alpha'}$ einführen, entsprechend einer Reskalierung der $\delta q^{\alpha'}$, wobei der Skalierungsfaktor unterschiedlich für jede Koordinate ist. Das ist ja möglich, denn die generalisierten Koordinaten keine „absolute“ Zahlenwerte von Vektorkomponenten sind, sondern parametrisieren sie die möglichen Konfigurationen der Freiheitsgrade.

Führt man die zeitunabhängigen diagonalen Matrizen

$$\mathbf{m}_D^{1/2} \equiv \text{diag}(\sqrt{m_1}, \dots, \sqrt{m_s}) \quad \text{und} \quad \mathbf{m}_D^{-1/2} \equiv \text{diag}(1/\sqrt{m_1}, \dots, 1/\sqrt{m_s})$$

ein, wobei die Exponenten $1/2$, $-1/2$ als eine Notation ohne Bedeutung zu sehen sind, so gelten

$$\mathbf{Q}' = \mathbf{m}_D^{1/2} \delta \mathbf{q}' \quad \text{und} \quad \delta \mathbf{q}' = \mathbf{m}_D^{-1/2} \mathbf{Q}',$$

wobei \mathbf{Q}' der Spaltenvektor mit Komponenten $\{Q^{\alpha'}\}$ bezeichnet. Die Lagrange-Funktion lässt sich dann als Funktion der $\{Q^{\alpha'}\}$ und ihrer Zeitableitungen schreiben

$$\mathcal{L}(\mathbf{Q}', \dot{\mathbf{Q}}') = \frac{1}{2} \dot{\mathbf{Q}}'^{\text{T}} \dot{\mathbf{Q}}' - \frac{1}{2} \mathbf{Q}'^{\text{T}} \mathbf{m}_D^{-1/2} \mathcal{R} \mathbf{k} \mathcal{R}^{\text{T}} \mathbf{m}_D^{-1/2} \mathbf{Q}'. \quad (\text{IV.13})$$

Wie sich einfach nachprüfen lässt, ist die Matrix $\mathbf{m}_D^{-1/2} \mathcal{R} \mathbf{k} \mathcal{R}^{\text{T}} \mathbf{m}_D^{-1/2}$ symmetrisch:

$$[\mathbf{m}_D^{-1/2} \mathcal{R} \mathbf{k} \mathcal{R}^{\text{T}} \mathbf{m}_D^{-1/2}]^{\text{T}} = [\mathbf{m}_D^{-1/2}]^{\text{T}} \mathcal{R} \mathbf{k} \mathcal{R}^{\text{T}} [\mathbf{m}_D^{-1/2}]^{\text{T}}.$$

Deshalb kann sie diagonalisiert werden, d.h. es existiert eine zeitunabhängige diagonale Matrix $\mathbf{K}_D \equiv \text{diag}(K_1, \dots, K_s)$ mit reellen $\{K_{\alpha}\}$ und eine s -dimensionale Drehmatrix \mathcal{R}' , die

$$\mathbf{m}_D^{-1/2} \mathcal{R} \mathbf{k} \mathcal{R}^{\text{T}} \mathbf{m}_D^{-1/2} = \mathcal{R}'^{\text{T}} \mathbf{K}_D \mathcal{R}'$$

erfüllen. Sei ein neuer Satz von s verallgemeinerten Koordinaten $\{Q^{\alpha}\}$, Linearkombinationen der $\{Q^{\beta'}\}$, definiert durch $\mathbf{Q} \equiv \mathcal{R}' \mathbf{Q}'$, woraus $\mathbf{Q}' = \mathcal{R}'^{\text{T}} \mathbf{Q}$ folgt. Diese neuen Koordinaten bringen natürlich den Potential-Term der Lagrange-Funktion in Diagonalform.

Andererseits gilt

$$\dot{\mathbf{Q}}'^{\text{T}} \dot{\mathbf{Q}}' = \dot{\mathbf{Q}}^{\text{T}} \mathcal{R}' \mathcal{R}'^{\text{T}} \dot{\mathbf{Q}} = \dot{\mathbf{Q}}^{\text{T}} \dot{\mathbf{Q}} = \sum_{\alpha=1}^s (\dot{Q}^{\alpha})^2$$

für den kinetischen Anteil der Lagrange-Funktion, die insgesamt durch

$$\mathcal{L}(\mathbf{Q}, \dot{\mathbf{Q}}) = \frac{1}{2} \dot{\mathbf{Q}}^{\text{T}} \dot{\mathbf{Q}} - \frac{1}{2} \mathbf{Q}^{\text{T}} \mathbf{K}_D \mathbf{Q} = \frac{1}{2} \sum_{\alpha=1}^s [(\dot{Q}^{\alpha})^2 - K_{\alpha} (Q^{\alpha})^2] \quad (\text{IV.14})$$

gegeben ist. Das heißt, die neuen verallgemeinerten Koordinaten $\{Q^{\alpha}\}$, deren Bewegungsgleichungen nach Anwendung der Euler-Lagrange-Gleichungen

$$\ddot{Q}_{\alpha}(t) = -K_{\alpha} Q^{\alpha}(t) \quad \text{für } \alpha = 1, \dots, s$$

lauten, sind jetzt entkoppelt von einander. Diese Variablen $\{Q^{\alpha}\}$ mit einer einfachen Bewegungsgleichung werden *Normalkoordinaten* oder auch *Normalmoden* genannt.

Für jede dieser Normalkoordinaten kann dann die Fallunterscheidung am Ende des § IV.1.1 wiederholt werden.

Somit entspricht der Fall $K_{\alpha} > 0$ einem harmonischen Oszillator mit Kreisfrequenz $\sqrt{K_{\alpha}}$, und das System ist stabil unter kleinen Variationen entlang der Richtung von Q^{α} . Dagegen ist das System instabil falls $K_{\alpha} < 0$, denn eine kleine Auslenkung kann zu einer großen Änderung führen. Schließlich ist das Potential „flach“ entlang einer Richtung mit $K_{\alpha} = 0$, und die zugehörige verallgemeinerte Koordinate Q^{α} ist zyklisch — zumindest in der Näherung, in der Gl. (IV.14) gilt.

Beispiel: Gekoppelte harmonische Oszillatoren

Zur Illustration der Zerlegung in Normalmoden, d.h. des Übergangs aus der Form (IV.11b) der Lagrange-Funktion zur einfacheren Form (IV.14), betrachten wir das in Abb. IV.1 dargestellte

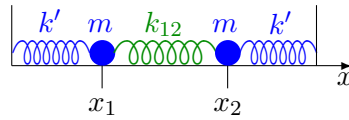


Abbildung IV.1

System aus zwei gekoppelten identischen harmonischen Oszillatoren (Masse m , Stärke k'), wobei die Schwingungen entlang der x -Richtung stattfinden.

Seien x_1 und x_2 die Auslenkungen der beiden Massen aus ihren jeweiligen Ruhelagen. Die (Standard-)Lagrange-Funktion \mathcal{L} des Systems ist

$$\mathcal{L} = T - V = \frac{1}{2}m\dot{x}_1^2 + \frac{1}{2}m\dot{x}_2^2 - \frac{1}{2}k'x_1^2 - \frac{1}{2}k'x_2^2 - \frac{1}{2}k_{12}(x_2 - x_1)^2. \quad (\text{IV.15})$$

Dabei spielen die Ablenkungen x_1, x_2 die Rolle der „kleinen Variationen“ $\{\delta q^\alpha\}$ des allgemeinen Formalismus. Das angenommene Potential ist schon quadratisch in diesen Ablenkungen, entsprechend dem Modell des harmonischen Oszillators. Das heißt, hier ist keine Taylor-Entwicklung in x_1, x_2 bis zur zweiten Ordnung nötig, weil die Lagrange-Funktion [s. auch Gl. (IV.16a) unten] schon der Form (IV.11b) ist.

Diese Lagrange-Funktion lässt sich noch als

$$\mathcal{L} = \frac{1}{2}m\dot{x}_1^2 + \frac{1}{2}m\dot{x}_2^2 - \frac{1}{2}kx_1^2 - \frac{1}{2}kx_2^2 + k_{12}x_1x_2 \quad (\text{IV.16a})$$

schreiben, wobei $k \equiv k' + k_{12}$. Dies ist der Form (IV.12): der kinetische Term ist schon in Diagonalform, die potentielle Energie noch nicht. In Matrixform lautet dies noch

$$\mathcal{L} = \frac{1}{2} \begin{pmatrix} \dot{x}_1 & \dot{x}_2 \end{pmatrix} \begin{pmatrix} m & 0 \\ 0 & m \end{pmatrix} \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} k & -k_{12} \\ -k_{12} & k \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \quad (\text{IV.16b})$$

Sei nun $Q^{1'} \equiv \sqrt{m}x_1$, $Q^{2'} \equiv \sqrt{m}x_2$. Definiert man

$$\omega_0^2 \equiv \frac{k}{m} \quad \text{und} \quad \omega_1^2 \equiv \frac{k_{12}}{m},$$

so lässt sich die Lagrange-Funktion als

$$\mathcal{L} = \frac{1}{2} \begin{pmatrix} \dot{Q}^{1'} & \dot{Q}^{2'} \end{pmatrix} \begin{pmatrix} \dot{Q}^{1'} \\ \dot{Q}^{2'} \end{pmatrix} - \frac{1}{2} \begin{pmatrix} Q^{1'} & Q^{2'} \end{pmatrix} \begin{pmatrix} \omega_0^2 & -\omega_1^2 \\ -\omega_1^2 & \omega_0^2 \end{pmatrix} \begin{pmatrix} Q^{1'} \\ Q^{2'} \end{pmatrix}$$

umschreiben, d.h. wie in Gl. (IV.13). Jetzt soll die Matrix

$$\mathbf{M} \equiv \begin{pmatrix} \omega_0^2 & -\omega_1^2 \\ -\omega_1^2 & \omega_0^2 \end{pmatrix}$$

diagonalisiert werden. Die Nullstellen des charakteristischen Polynoms

$$\det(\mathbf{M} - \lambda \mathbf{1}_2) = (\omega_0^2 - \lambda)^2 - (\omega_1^2)^2$$

sind die Eigenwerte von \mathbf{M} , und zwar $\lambda_1 = \omega_0^2 - \omega_1^2$ und $\lambda_2 = \omega_0^2 + \omega_1^2$. Die zugehörigen (hier auf 1 normierten) Eigenvektoren sind

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{und} \quad \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Dann gilt

$$\mathbf{M} = \mathcal{R}'^T \begin{pmatrix} \omega_0^2 - \omega_1^2 & 0 \\ 0 & \omega_0^2 + \omega_1^2 \end{pmatrix} \mathcal{R}' \quad \text{mit} \quad \mathcal{R}' \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

Schließlich sind die Normalmoden Q^1, Q^2 durch

$$\begin{pmatrix} Q^1 \\ Q^2 \end{pmatrix} = \mathcal{R}' \begin{pmatrix} Q^{1'} \\ Q^{2'} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} Q^{1'} + Q^{2'} \\ Q^{2'} - Q^{1'} \end{pmatrix}$$

gegeben. Somit ist Q^1 proportional zur Schwerpunktkoordinate $\frac{1}{2}(x_1 + x_2)$ und Q^2 zur Relativkoordinate $x_2 - x_1$. Die entsprechenden Bewegungsgleichungen lauten

$$\ddot{Q}^1 = -\lambda_1 Q^1 = -\frac{k'}{m} Q^1 \quad \text{und} \quad \ddot{Q}^2 = -\lambda_2 Q^2 = -\frac{k' + 2k_{12}}{m} Q^2.$$

Wie erwartet sind diese Gleichungen entkoppelt. Da k' und $k' + 2k_{12}$ beide positiv sind, sind die Bewegungen von Q^1 und Q^2 harmonische Schwingungen.

IV.2 Starre Körper

Ein physikalischer Festkörper ist ein System aus vielen Atomen — für ein Stück auf menschlicher Skala, etwa 10^{23} – 10^{26} Atomen. Dementsprechend ist die Beschreibung als ein Mehrteilchensystem wie in Abschn. II.1 unrealistisch. Stattdessen werden idealisierte Modelle des Festkörpers benutzt.

Das einfachste davon ist das Modell eines *starrten Körpers*. Dabei handelt es sich um ein System aus (vielen) Massenpunkten m_a mit konstanten Abständen zueinander. Diese Zwangsbedingungen über die Abstände folgen aus inneren Kräften. Dank den letzteren bleibt die Anzahl der eigentlichen Freiheitsgrade im Modell trotz der hohen Teilchenzahl immer klein (§ IV.2.1). Für die Beschreibung der globalen Bewegung des Körpers ist eine Kenntnis der inneren Zwangskräfte nicht nötig. Deshalb eignet sich die Lagrange'sche Vorgehensweise besser als der newtonsche Formalismus, um die Bewegungsgleichungen für den starren Körper herzuleiten (§ IV.2.2).

Anstatt den Körper als eine Menge von diskreten punktförmigen Teilchen zu beschreiben, ist es auch günstig, das Modell eines *kontinuierlichen Mediums* einzuführen, z.B. um einige Eigenschaften zu berechnen, die in der Bewegung von Relevanz sind. Im Fall des starren Körpers ist das Kontinuum definitionsgemäß nicht deformierbar. Eine weitere Modellierung des physikalischen Festkörpers, die in diesem Kapitel nicht diskutiert wird, ist die eines verformbaren Kontinuums, in dessen Rahmen Eigenschaften wie Elastizität, Plastizität... untersucht werden können.

IV.2.1 Beschreibung des starren Körpers

IV.2.1 a Anzahl der Freiheitsgrade

Das Modell des starren Körpers reduziert die a priori große ($3N$) Anzahl der Freiheitsgrade eines N -Teilchen-Systems zu einer kleinen Anzahl, und zwar meistens nur 6 für die Fälle, wo die Vereinfachung von Nutzen ist.

Im Fall von zwei Massenpunkten 1 und 2 mit der Zwangsbedingung eines zeitlich konstanten Abstands $|\vec{x}_2(t) - \vec{x}_1(t)|$ besitzt das System insgesamt $s = 3N - 1 = 5$ Freiheitsgrade. Diese können z.B. die drei Koordinaten des Schwerpunkts und zwei Winkel für die Orientierung des Verbindungsvektors zwischen 1 und 2 sein.

Für drei Massenpunkte 1, 2 und 3 gibt es drei Zwangsbedingungen über die Abstände, und zwar $|\vec{x}_2(t) - \vec{x}_1(t)|$, $|\vec{x}_3(t) - \vec{x}_1(t)|$ und $|\vec{x}_3(t) - \vec{x}_2(t)|$. Unter deren Berücksichtigung bleiben im allgemeinen Fall $s = 3N - 3 = 6$ Freiheitsgrade übrig. Beispielsweise kann man drei Schwerpunktskoordinaten betrachten, zusammen mit drei Winkeln für die Orientierung im Raum der Ebene, in welcher die drei Massenpunkte liegen.

Fügt man einen vierten Massenpunkt hinzu, so ist seine Position durch 3 Zwangsbedingungen $|\vec{x}_4(t) - \vec{x}_1(t)|$, $|\vec{x}_4(t) - \vec{x}_2(t)|$ und $|\vec{x}_4(t) - \vec{x}_3(t)|$ völlig bestimmt — bis auf eine Spiegelung bezüglich der durch die drei ersten Punkte definierten Ebene. Somit gibt es immer noch 6 Freiheitsgrade.

Das gleiche gilt für eine beliebige Anzahl $N \geq 4$. Wenn man einen N -ten Massenpunkt zu $N - 1$ schon anwesenden Punkten addiert, dann bestimmen drei Zwangsbedingungen $|\vec{x}_N(t) - \vec{x}_1(t)|$,

$|\vec{x}_N(t) - \vec{x}_2(t)|$ und $|\vec{x}_N(t) - \vec{x}_3(t)|$ seine Position: die weiteren Abstände sind dadurch automatisch festgelegt bzw. sind nicht unabhängig davon. Deshalb werden die drei möglichen Freiheitsgrade dieses zusätzlichen Massenpunkts durch genau drei Bedingungen annulliert, d.h. die Anzahl der Freiheitsgrade ändert sich nicht, und bleibt gleich 6.

Bemerkung: Im Sonderfall, wo die $N \geq 2$ Massenpunkte alle auf einer Geraden sitzen, gibt es eigentlich wie für $N = 2$ nur 5 Freiheitsgrade. Dieser Fall wird im Folgenden stillschweigend ausgeschlossen.

Kontinuuumlimes

Dank der Unabhängigkeit der Anzahl von Freiheitsgraden von der Anzahl N von Massenpunkten für $N \geq 3$ kann man problemlos den Limes $N \rightarrow \infty$ betrachten, und danach zum Modell eines *Kontinuums* übergehen. Im letzteren Fall entspricht der starre Körper einem endlichen kontinuierlichen Volumen, das sich nicht verformen kann, wie z.B. eine Vollkugel. Dementsprechend werden einige Eigenschaften des Körpers durch stückweise stetige Funktionen vom Ort beschrieben — insbesondere durch eine Massendichte $\rho(\vec{r})$ —, die genau da ungleich Null sind, wo der Körper sich befindet.

Zur genaueren Definition des Grenzfalls sollte man das nicht-deformierbare kontinuierliche Medium zuerst als Zusammensetzung aus endlich vielen (N) kleinen Massenstücken m_a beschreiben, die sich relativ zu einander nicht bewegen; seien \vec{x}_a ihre Ortsvektoren. Ersetzt man die Massenstückchen durch Massenpunkte, so hat man einen starren Vielteilchen-Körper. Sei $M = \sum_a m_a$ die Gesamtmasse des Körpers.

Im *Kontinuuumlimes* wird die Anzahl der Massenpunkte unendlich groß, $N \rightarrow \infty$, und ihre Massen werden verschwindend klein, $m_a \rightarrow 0$, bei festgehaltener Gesamtmasse M . Im Grenzfalle werden die Summen über a , d.h. über alle Massenpunkte, durch Integrale über das durch den Körper besetzte Volumen \mathcal{V} ersetzt. Genauer schreibt man für jede Funktion f der Position

$$\sum_a m_a f(\vec{x}_a) \rightarrow \int_{\mathcal{V}} \rho(\vec{r}) f(\vec{r}) d^3\vec{r} \quad (\text{IV.17})$$

mit der Massendichte ρ .

Bemerkung: Schreibt man

$$\rho(\vec{r}) = \sum_a m_a \delta^{(3)}(\vec{r} - \vec{x}_a),$$

wobei $\delta^{(3)}$ die dreidimensionale Dirac-Distribution⁽⁴⁰⁾ bezeichnet, so kann man ein System aus diskreten Massenpunkten in der Sprache des Kontinuuumlimes beschreiben.

IV.2.1 b Kinematik des starren Körpers

Zur Untersuchung der Bewegung eines starren Körpers ist es sinnvoll, zwei unterschiedliche Bezugssysteme mit zugehörigen kartesischen Koordinatensystemen einzuführen. Somit betrachtet man einerseits ein raumfestes Inertialsystem \mathcal{B}_I , entsprechend z.B. einem inertialen Beobachter, der die Bewegung des Körpers misst, ohne sich damit zu bewegen. Ortsvektoren relativ zu diesem System werden hiernach mit Großbuchstaben bezeichnet, wie z.B. $\{\vec{X}_a(t)\}$ für die Ortsvektoren der individuellen Massenpunkte des starren Körpers.

Andererseits wird auch ein körperfestes Bezugssystem \mathcal{B} eingeführt, das sich zusammen mit dem starren Körper bewegt. Insbesondere ist \mathcal{B} nicht unbedingt Inertial, weil die Bewegung des Körpers beschleunigt relativ zu \mathcal{B}_I sein kann. In \mathcal{B} wird dann ein Koordinatensystem gewählt, insbesondere ein Ursprungspunkt O , dessen Position bezüglich \mathcal{B}_I mit $\vec{X}(t)$ bezeichnet wird. Ortsvektoren relativ zu \mathcal{B} werden mit Kleinbuchstaben denotiert, z.B. \vec{x}_a für die Positionen der Massenpunkte des starren Körpers, die definitionsgemäß zeitunabhängig sind.

⁽⁴⁰⁾ $\delta^{(3)}(\vec{r}) \equiv \delta(x)\delta(y)\delta(z)$ mit den kartesischen Koordinaten (x, y, z) von \vec{r} , vgl. Anhang C.

Bemerkung: Der Nullpunkt O kann prinzipiell beliebig gewählt werden. Wie wir hiernach sehen werden [s. Diskussion unten Gl. (IV.22)], gibt es in der Praxis „bessere“ Wähle, die zu einer einfacheren Form der kinetischen Energie bzw. der Bewegungsgleichungen des Körpers führen.

Drei der für die Beschreibung des starren Körpers nötigen 6 Freiheitsgrade beschreiben die Translationsbewegung des Ursprungspunkts O von \mathcal{B} relativ zu \mathcal{B}_I . Die drei anderen entsprechen z.B. den Winkeln der Drehungen, die von den Achsen des Koordinatensystems in \mathcal{B}_I zu den Koordinatenachsen von \mathcal{B} führen (vgl. § IV.2.1 c).

Unter Verwendung der oben eingeführten Notationen kann die Position eines Massenpunkts a des starren Körpers bezüglich des Inertialsystems als

$$\vec{X}_a(t) = \vec{X}(t) + \mathcal{R}(t)\vec{x}_a \quad (\text{IV.18})$$

geschrieben werden. Dabei ist $\mathcal{R}(t)$ die zeitabhängige 3×3 -Drehmatrix, die die Koordinatenachsen von \mathcal{B} in jene von \mathcal{B}_I transformiert. Eine Ableitung nach der Zeit gibt nun die Geschwindigkeit des Massenpunkts relativ zu \mathcal{B}_I [vgl. Gl. (I.46)]

$$\dot{\vec{X}}_a(t) = \dot{\vec{X}}(t) + \mathcal{R}(t)[\vec{\omega}(t) \times \vec{x}_a] \quad (\text{IV.19})$$

wobei $\vec{x}_a = \vec{0}$ benutzt wurde. Dabei ist $\vec{\omega}(t)$ die Winkelgeschwindigkeit der Rotationsbewegung des starren Körpers. Da der letztere sich definitionsgemäß nicht deformieren kann, ist $\vec{\omega}(t)$ die gleiche für alle Punkte a des Körpers.

IV.2.1 c Euler-Winkel

Die Orientierung der Achsen (x, y, z) des körperfesten Bezugssystems relativ zu denen (X, Y, Z) des raumfesten Bezugssystems kann mithilfe von drei Winkeln φ , θ und ψ , den *Euler-Winkeln*, beschrieben werden.⁽⁴¹⁾

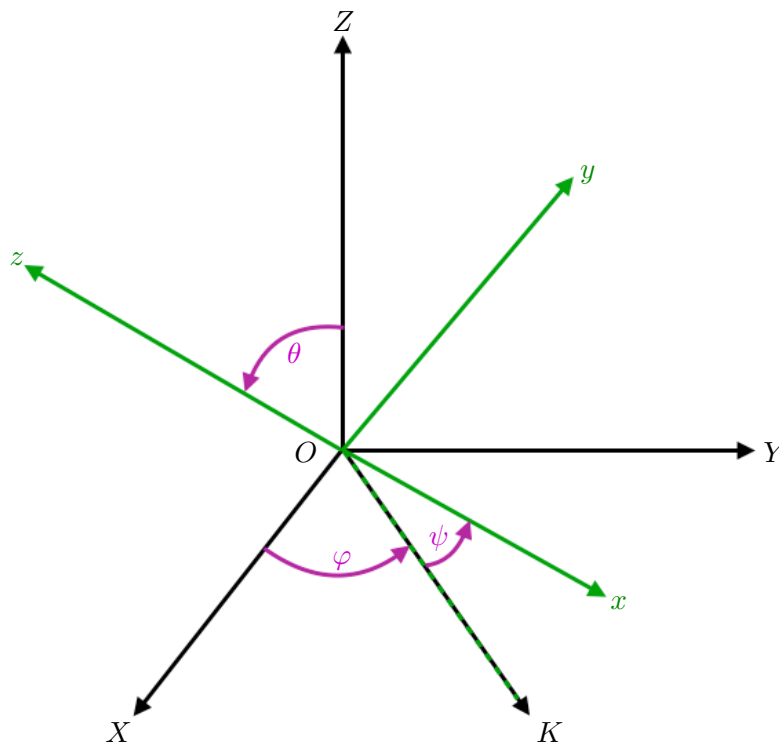


Abbildung IV.2 – Euler-Winkel

⁽⁴¹⁾ Anstatt φ, θ, ψ werden die Euler-Winkel auch oft mit α, β, γ bezeichnet, entsprechend der Reihenfolge der Drehungen.

Zur Beschreibung wird der Einfachheit angenommen, dass die zwei Systeme den gleichen Nullpunkt haben, s. Abb. IV.2. Dann schneiden sich die (x, y) -Ebene und die (X, Y) -Ebene in einer Geraden K , der sog. *Knotenlinie*. Konstruktionsgemäß ist diese orthogonal zu den z - und Z -Achsen.

Der erste Euler-Winkel φ ist der Winkel — gemessen in Richtung der Y -Achse — zwischen der X -Achse und der Knotenlinie K , entsprechend einer Drehung $\mathcal{R}_Z(\varphi)$ um die Z -Achse: in Matrixform gilt

$$\mathcal{R}_Z(\varphi) = \begin{pmatrix} \cos \varphi & \sin \varphi & 0 \\ -\sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Der zweite Euler-Winkel θ ist der Winkel zwischen der Z -Achse und der z -Achse, entsprechend einer Drehung $\mathcal{R}_K(\theta)$ um die Gerade K :

$$\mathcal{R}_K(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & -\sin \theta & \cos \theta \end{pmatrix}.$$

Dabei ist aufzupassen, dass $\mathcal{R}_K(\theta)$ nicht auf die Komponenten von Vektoren auf den (X, Y, Z) -Achsen wirkt, sondern auf die Komponenten entlang Achsen (K, K_\perp, Z) , wobei K_\perp senkrecht auf K und der Z -Achse steht.

Schließlich ist der dritte Euler-Winkel ψ der Winkel zwischen der Knotenlinie K und der x -Achse, entsprechend einer Drehung $\mathcal{R}_z(\psi)$ um die z -Achse:

$$\mathcal{R}_z(\psi) = \begin{pmatrix} \cos \psi & \sin \psi & 0 \\ -\sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Insgesamt ist die Drehung \mathcal{R} , die das (X, Y, Z) -System in das (x, y, z) -System transformiert, das Produkt aus den drei Drehungen. Unter Verwendung der kürzeren Notationen $c_\varphi \equiv \cos \varphi$, $s_\varphi \equiv \sin \varphi$, usw. findet man

$$\mathcal{R}(\varphi, \theta, \psi) \equiv \mathcal{R}_z(\psi) \mathcal{R}_K(\theta) \mathcal{R}_Z(\varphi) = \begin{pmatrix} c_\varphi c_\psi - s_\varphi c_\theta s_\psi & s_\varphi c_\psi + c_\varphi c_\theta s_\psi & s_\theta s_\psi \\ -s_\varphi c_\theta c_\psi - c_\varphi s_\psi & c_\varphi c_\theta c_\psi - s_\varphi s_\psi & s_\theta c_\psi \\ s_\varphi s_\theta & -c_\varphi s_\theta & c_\theta \end{pmatrix}. \quad (\text{IV.20})$$

Diese Drehmatrix gibt den Zusammenhang zwischen den Ortsvektoren \vec{R} bzw. \vec{r} eines Punkts relativ zu \mathcal{B}_I bzw. \mathcal{B} :

$$\vec{r} = \mathcal{R}(\varphi, \theta, \psi) \vec{R}. \quad (\text{IV.21a})$$

d.h.

$$\vec{R} = \mathcal{R}(\varphi, \theta, \psi)^\top \vec{r}. \quad (\text{IV.21b})$$

IV.2.2 Bewegungsgleichungen

IV.2.2 a Kinetische Energie

Die kinetische Energie des starren Körpers bezüglich des Inertialsystems \mathcal{B}_I ist die Summe aus der kinetischen Energien der individuellen Massenpunkten, d.h.

$$T = \sum_a \frac{m_a}{2} \dot{\vec{X}}_a(t)^2.$$

Dabei kann der Ausdruck (IV.19) der Geschwindigkeit $\dot{\vec{X}}_a(t)$ substituiert werden

$$T = \sum_a \frac{m_a}{2} \left(\dot{\vec{X}}(t) + \mathcal{R}(t) [\vec{\omega}(t) \times \vec{x}_a] \right)^2.$$

Die Berechnung des (Betrags)Quadrats liefert dann

$$T = \frac{1}{2} \sum_a m_a \dot{\vec{X}}(t)^2 + \dot{\vec{X}}(t) \cdot \left(\mathcal{R}(t) [\vec{\omega}(t) \times \sum_a m_a \vec{x}_a] \right) + \sum_a \frac{m_a}{2} \left(\mathcal{R}(t) [\vec{\omega}(t) \times \vec{x}_a] \right)^2 \quad (\text{IV.22})$$

$$\equiv T_{\text{Trans.}} + T_{\text{Misch.}} + T_{\text{Rot.}}$$

Der erste Term lässt sich mithilfe der Gesamtmasse einfach als $T_{\text{Trans.}} = \frac{1}{2} M \dot{\vec{X}}(t)^2$ umschreiben: diese *Translationsenergie* entspricht der kinetischen Energie eines Massenpunkts der Masse M , der sich im Ursprungspunkt O des körperfesten Bezugssystems befindet.

Bei dem zweiten Beitrag $T_{\text{Misch.}}$ in Gl. (IV.22) handelt es sich um einen Mischterm, der sowohl von der Translationsbewegung des starren Körpers als von seiner Rotationsbewegung abhängt. Dieser Term verschwindet in zwei Fällen: wenn der Nullpunkt O in \mathcal{B}_I ruht, d.h. falls $\dot{\vec{X}}(t) = \vec{0}$, oder wenn O genau im Schwerpunkt des starren Körpers liegt, so dass $\sum_a m_a \vec{x}_a = \vec{0}$ gilt. Hiernach wird meistens angenommen, dass einer dieser beiden Möglichkeit erfüllt ist.

Der dritte Term in Gl. (IV.22) ist die *Rotationsenergie* des starren Körpers. Deren Ausdruck lässt sich zuerst vereinfachen, indem man die Invarianz des Skalarprodukts zweier Vektoren, und daher des Betragsquadrats eines Vektors, unter Drehungen benutzt:

$$T_{\text{Rot.}} = \sum_a \frac{m_a}{2} \left(\mathcal{R}(t) [\vec{\omega}(t) \times \vec{x}_a] \right)^2 = \frac{1}{2} \sum_a m_a [\vec{\omega}(t) \times \vec{x}_a]^2. \quad (\text{IV.23})$$

Das Betragsquadrat kann noch mithilfe einiger Manipulationen umgeschrieben werden. Führt man z.B. die kartesischen Koordinaten $\omega^i(t)$, x_a^i der Vektoren $\vec{\omega}(t)$ und \vec{x}_a ein, so gilt

$$[\vec{\omega}(t) \times \vec{x}_a]^2 = [\vec{\omega}(t) \times \vec{x}_a]^k [\vec{\omega}(t) \times \vec{x}_a]^k = [\epsilon^{kij} \omega^i(t) x_a^j] [\epsilon^{klm} \omega^l(t) x_a^m],$$

wobei Summen über k, i, j, l und m von 1 bis 3 nicht geschrieben wurden. Fängt man mit der Summe über k an, so führt die Identität $\epsilon^{kij} \epsilon^{klm} = \delta^{il} \delta^{jm} - \delta^{im} \delta^{jl}$ zu

$$[\vec{\omega}(t) \times \vec{x}_a]^2 = (\delta^{il} \delta^{jm} - \delta^{im} \delta^{jl}) \omega^i(t) \omega^l(t) x_a^j x_a^m.$$

Die Summen über l und m liefern dann

$$[\vec{\omega}(t) \times \vec{x}_a]^2 = \omega^i(t) \left[(\vec{x}_a)^2 \delta^{ij} - x_a^i x_a^j \right] \omega^j(t), \quad (\text{IV.24})$$

wobei im ersten Term $\omega^i(t) \omega^i(t) = \omega^i(t) \delta^{ij} \omega^j(t)$ benutzt wurde. Dieses Resultat kann in Gl. (IV.23) eingesetzt werden. Dann lautet die Rotationsenergie

$$T_{\text{Rot.}} = \frac{1}{2} \sum_{i,j=1}^3 \omega^i(t) I^{ij} \omega^j(t) \quad (\text{IV.25})$$

mit

$$I^{ij} \equiv \sum_a m_a \left[(\vec{x}_a)^2 \delta^{ij} - x_a^i x_a^j \right] \quad \text{für } i, j = 1, 2, 3, \quad (\text{IV.26a})$$

wobei die Summe über alle Massenpunkte des starren Körpers läuft. Gemäß dem Rezept (IV.17) gilt im Grenzfall eines kontinuierlichen Körpers

$$I^{ij} \equiv \int \rho(\vec{r}) \left[\vec{r}^2 \delta^{ij} - x^i x^j \right] d^3 \vec{r} \quad \text{für } i, j = 1, 2, 3 \quad (\text{IV.26b})$$

mit den kartesischen Komponenten $\{x^i\}_{i=1,2,3}$ des Vektors \vec{r} . Der Integrationsbereich ist entweder das durch den starren Körper besetzte Volumen \mathcal{V} , oder kann der ganze Raum \mathbb{R}^3 sein, unter Verwendung der Konvention, dass $\rho(\vec{r}) = 0$ für \vec{r} außerhalb \mathcal{V} .

Alternative Herleitung der Gl. (IV.24): Das Betragsquadrat auf der linken Seite kann als Spatprodukt interpretiert werden und in der Form

$$(\vec{\omega} \times \vec{x}_a)^2 = (\vec{\omega} \times \vec{x}_a) \cdot (\vec{\omega} \times \vec{x}_a) = [\vec{x}_a \times (\vec{\omega} \times \vec{x}_a)] \cdot \vec{\omega} = [(\vec{x}_a \cdot \vec{x}_a)\vec{\omega} - (\vec{x}_a \cdot \vec{\omega})\vec{x}_a] \cdot \vec{\omega}$$

umgeschrieben werden, wobei die Zeitabhängigkeit von $\vec{\omega}$ der Kurze halber nicht geschrieben wurde. Dies vereinfacht sich noch zu

$$(\vec{\omega} \times \vec{x}_a)^2 = \vec{x}_a^2 \vec{\omega}^2 - (\vec{x}_a \cdot \vec{\omega})^2,$$

was äquivalent zu Gl. (IV.24) ist. \square

Insgesamt lautet die kinetische Energie des starren Körpers bezüglich des Inertialsystems \mathcal{B}_I

$$T = \frac{1}{2} M \dot{\vec{X}}(t)^2 + \frac{1}{2} \sum_{i,j=1}^3 \omega^i(t) I^{ij} \omega^j(t). \quad (\text{IV.27})$$

Dabei ist $\vec{X}(t)$ die Position in \mathcal{B}_I entweder des Schwerpunkts des starren Körpers oder eines Punkts davon, der sich relativ zu \mathcal{B}_I nicht bewegt. Im letzteren Fall wird die Translationsenergie in Gl. (IV.27) auch Null sein.

Im Gegensatz zur Gesamtmasse, deren Wert unabhängig vom Koordinatensystem ist, hängen die durch Gl. (IV.26) definierten Zahlen I^{ij} von der Wahl des körperfesten Koordinatensystems ab, insbesondere von dessen Nullpunkt. Wie in § IV.2.2c unten argumentiert wird, sind die I^{ij} die kartesischen Komponenten eines Tensors zweiter Stufe \mathbf{I} , der *Trägheitstensor* genannt wird.

IV.2.2b Bewegungsgleichungen

Entsprechend den 6 Freiheitsgraden eines starren Körpers wählt man als verallgemeinerte Koordinaten für die Beschreibung seiner Bewegung einerseits die 3 kartesischen Koordinaten X^i des Ortsvektors \vec{X} eines bestimmten Punkts des Körpers und andererseits drei Winkel φ^i ; die letzteren beschreiben Drehungen des starren Körpers um drei zueinander senkrechte x^i -Achsen, die relativ zum Körper fest bleiben. Die zugehörigen verallgemeinerten Geschwindigkeiten sind \dot{X}^i und $\dot{\varphi}^i$: die drei $\dot{\varphi}^i$ sind die kartesischen Komponenten der Winkelgeschwindigkeit $\vec{\omega} = \vec{\dot{\varphi}}$ des Körpers.

Die Standard-Lagrange-Funktion des starren Körpers lautet dann

$$\mathcal{L}(\{X^i\}, \{\varphi^i\}, \{\dot{X}^i\}, \{\dot{\varphi}^i\}) = \frac{M}{2} \sum_{i=1}^3 (\dot{X}^i)^2 + \frac{1}{2} \sum_{i,j=1}^3 \dot{\varphi}^i I^{ij} \dot{\varphi}^j - V(\{X^i\}, \{\varphi^i\}), \quad (\text{IV.28})$$

mit einem Potential V . Dabei sind die I^{ij} die Komponenten des Trägheitstensors bezüglich der körperfesten Drehachsen, die den drei Winkeln φ^i entsprechen. Ausgehend von dieser Lagrange-Funktion liefern die Euler-Lagrange-Gleichungen (III.11) die Bewegungsgleichungen des starren Körpers.

Translationsbewegung des starren Körpers

Betrachten wir zuerst die Bewegungsgleichung für die kartesischen Koordinaten $\{X^i\}$. Laut der Definition (III.17) ist der zu X^i kanonisch konjugierte Impuls durch

$$p^i \equiv \frac{\partial \mathcal{L}}{\partial \dot{X}^i} = M \dot{X}^i \quad (\text{IV.29a})$$

gegeben. Dies ist offensichtlich die i -te kartesische Komponente des Gesamtimpulses $\vec{p} = M \dot{\vec{X}}$ des starren Körpers. Dann lautet die damit assoziierte Euler-Lagrange-Gleichung

$$\frac{dp^i}{dt} = M \ddot{X}^i = \frac{\partial \mathcal{L}}{\partial X^i} = -\frac{\partial V}{\partial X^i}. \quad (\text{IV.29b})$$

Auf der rechten Seite ist $-\partial V / \partial X^i$ die i -te Komponente der Gesamtkraft auf den starren Körper. Falls $\vec{X}(t)$ die Position im Inertialsystem \mathcal{B}_I des Schwerpunkts des Körpers ist, entspricht Gl. (IV.29b) dem üblichen Schwerpunktsatz (II.12) für ein Mehrteilchensystem.

Rotationsbewegung des starren Körpers

Benutzt man die auf Gl. (IV.26) offensichtliche Symmetrie $I^{ij} = I^{ji}$ für jedes mögliches Paar (i, j) , so findet man für den zu φ^i kanonisch konjugierten verallgemeinerten Impuls

$$L^i \equiv \frac{\partial \mathcal{L}}{\partial \dot{\varphi}^i} = \sum_{j=1}^3 I^{ij} \dot{\varphi}^j. \tag{IV.30a}$$

Dies ist die i -te kartesische Komponente des *Eigendrehimpulses* des starren Körpers um den Nullpunkt von \mathcal{B} .

Die Zeitentwicklung dieses generalisierten Impulses wird durch die Euler-Lagrange-Gleichung

$$\frac{dL^i}{dt} = \sum_{j=1}^3 I^{ij} \ddot{\varphi}^j = \frac{\partial \mathcal{L}}{\partial \varphi^i} = -\frac{\partial V}{\partial \varphi^i} \tag{IV.30b}$$

gegeben, wobei die „verallgemeinerte Kraft“ auf der rechten Seite der Gleichung das *Drehmoment* ist.

Beispiel: Homogener dünner Stab als physikalisches Pendel

Sei ein homogener unendlich dünner Stab mit Länge ℓ und linearer Massendichte (d.h. Masse pro Längeneinheit) μ . Dementsprechend ist seine Gesamtmasse $M = \mu\ell$.

Ein Endpunkt des Stabs ist am Nullpunkt des Koordinatensystems eines raumfesten Inertialsystems aufgehängt. Der Stab kann sich in der (Y, Z) -Ebene, d.h. um die X -Achse, unter dem Einfluss des Schwerfeldes $-g\vec{e}_Z$ drehen. Somit bildet der Stab ein ebenes Pendel.⁽⁴²⁾

Als körperfestes Koordinatensystem wird ein System mit dem Nullpunkt O im Aufhängepunkt, mit der x -Achse parallel zur X -Achse des raumfesten Systems und der y -Achse entlang des Stabs. Somit streckt sich der Stab von $y = 0$ bis $y = \ell$. Für dieses eindimensionale System kann man dann für eine beliebige Funktion f der Position

$$\int_V \rho(\vec{r}) f(\vec{r}) d^3\vec{r} = \int_0^\ell \mu(y) f(x=0, y, z=0) dy$$

schreiben; mit $f \equiv 1$ ergibt sich z.B. die Gesamtmasse $M = \mu\ell$.

Wer nicht mit einer linearen Massendichte $\mu(y)$, sondern mit der üblichen Massendichte (pro Volumeneinheit) $\rho(\vec{r})$ arbeiten möchte, kann einfach $\rho(\vec{r}) = \mu(y)\delta(x)\delta(z)$ betrachten: mit dieser Massendichte ergibt sich genau die obige Gleichung.

Sei θ der Ablenkwinkel des Stabs aus der Lotrichtung. Da θ einer Drehung um die x -Achse entspricht, gilt in den Notationen der Gl. (IV.28)–(IV.30) (d.h. mit $x = x^1$) $\theta = \varphi^1$. Dann ist $\dot{\varphi}^1 = \dot{\theta}$: um eine Bewegungsgleichung für θ zu erhalten, sollte man Gl. (IV.30b) mit $i = 1$ benutzen. Da wir ausschließlich Rotationsbewegungen um die x -Achse betrachten wollen, lautet die Winkelgeschwindigkeit $\vec{\omega} = \dot{\theta}\vec{e}_X$, d.h. mit nur einer φ^1 -Komponente: $\dot{\varphi}^2 = \dot{\varphi}^3 = 0$. Dementsprechend ist nur die Komponente I^{11} des Trägheitstensors um den Punkt O nötig, um Gl. (IV.30b) für $i = 1$ zu schreiben:

$$\frac{dL^1}{dt} = \sum_{j=1}^3 I^{1j} \ddot{\varphi}^j = I^{11} \ddot{\theta}.$$

Aus Gl. (IV.26b) folgt dann

$$I^{11} = \int_0^\ell \mu(y) y^2 dy = \frac{\mu\ell^3}{3} = \frac{M\ell^2}{3},$$

wobei $\vec{r}^2 \delta^{11} - (x)^2 = y^2$ wenn $x = z = 0$ benutzt wurde.

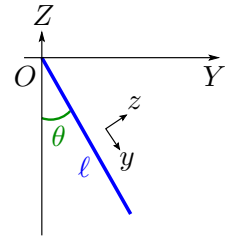


Abbildung IV.3

⁽⁴²⁾Im Gegensatz zum „mathematischen Pendel“ des § III.2.4 mit masselosem Stab und Punktmasse am Ende wird dieses System mit massebehaftetem Stab „physikalisches Pendel“ genannt.

Betrachten wir jetzt das Potential des Stabs im Schwerfeld. Für ein kleines Massenelement μdy im Abstand y vom Aufhängepunkt, d.h. bei der Höhe $Z = -y \cos \theta$ (vgl. Abb. IV.3) lautet es $V = (\mu dy)gZ = -\mu g \cos \theta y dy$. Die potentielle Energie des ganzen Stabs ist dann

$$V = \int_0^\ell (-\mu g \cos \theta) y dy = -\mu g \frac{\ell^2}{2} \cos \theta = -Mg \frac{\ell \cos \theta}{2}.$$

Dabei erkennt man, dass $-\ell \cos \theta/2$ die Höhe des Schwerpunkts des Stabs ist: alles passiert, als ob die ganze Masse in diesem Schwerpunkt konzentriert wäre.

Schließlich ist die Bewegungsgleichung (IV.30b) für θ

$$I^{11}\ddot{\theta} = \frac{M\ell^2}{3} = -\frac{\partial V}{\partial \theta} = -Mg \frac{\ell}{2} \sin \theta \quad \Leftrightarrow \quad \ddot{\theta} = -\frac{3g}{2\ell} \sin \theta.$$

Diese Gleichung hat die gleiche Form wie die Bewegungsgleichung (III.30) des mathematischen Pendels und wird ähnlich gelöst.

IV.2.2c Trägheitstensor

Um zu zeigen, dass die I^{ij} die (kartesischen) Komponenten eines Tensors bilden, soll man ihr Verhalten unter Drehungen untersuchen.⁽⁴³⁾ Sei \mathcal{R} eine Drehmatrix. Unter ihrer Wirkung transformieren sich die Komponenten eines Ortsvektors bezüglich \mathcal{B} gemäß

$$x^i \rightarrow x^{i'} = \mathcal{R}^{i'}_i x^i.$$

Dementsprechend transformiert sich das Produkt $x^i x^j$ in $x^{i'} x^{j'} = \mathcal{R}^{j'}_j \mathcal{R}^{i'}_i x^i x^j$. Andererseits gilt

$$\mathcal{R}^{i'}_i \mathcal{R}^{j'}_j \delta^{ij} = \mathcal{R}^{i'}_i \delta^{ij} (\mathcal{R}^\top)_j^{j'} = \mathcal{R}^{i'}_i (\mathcal{R}^\top)_i^{j'} = \delta^{i'j'},$$

wobei \mathcal{R}^\top die zu \mathcal{R} transponierte Matrix bezeichnet, die gleichzeitig die inverse Drehmatrix \mathcal{R}^{-1} ist, wie in der letzten Gleichung benutzt wurde. Somit transformieren sich beide Terme mit Indizes in der Definition (IV.26) der I^{ij} gleich unter Drehungen, und zwar derart, dass I^{ij} sich insgesamt gemäß

$$I^{ij} \rightarrow I^{i'j'} = \mathcal{R}^{i'}_i \mathcal{R}^{j'}_j I^{ij} = \mathcal{R}^{i'}_i (\mathcal{R}^\top)_i^{j'} I^{ij} \quad (\text{IV.31})$$

transformiert, d.h. wie die Komponenten eines Tensors zweiter Stufe.

In tensorieller Form lautet der Trägheitstensor

$$\mathbf{I} = \sum_a m_a [(\vec{x}_a)^2 \mathbf{g}^{-1} - \vec{x}_a \otimes \vec{x}_a] \quad (\text{IV.32a})$$

falls der starrer Körper aus diskreten Massenpunkten besteht, oder

$$\mathbf{I} = \int \rho(\vec{r}) [\vec{r}^2 \mathbf{g}^{-1} - \vec{r} \otimes \vec{r}] \quad (\text{IV.32b})$$

für einen kontinuierlichen Körper. In diesen Formeln bezeichnet \otimes das Tensorprodukt (vgl. § A.1.3 b–A.1.3 c), während \mathbf{g}^{-1} der inverse metrische Tensor (vgl. § A.1.4) ist, dessen Komponenten in einer kartesischen Basis $g^{ij} = \delta^{ij}$ sind.

Identifiziert man den Tensor \mathbf{I} mit einer 3×3 -Matrix mit Elementen I^{ij} , die ebenfalls mit \mathbf{I} bezeichnet wird, so kann die letzte Gleichung noch in Matrixform als

$$\mathbf{I} \rightarrow \mathbf{I}' = \mathcal{R} \mathbf{I} \mathcal{R}^\top \quad (\text{IV.33})$$

geschrieben werden. Gleichmaßen gilt für die Rotationsenergie in Gl. (IV.27)

$$T_{\text{Rot.}} = \frac{1}{2} \omega^i(t) I^{ij} \omega^j(t) = \frac{1}{2} \vec{\omega}(t)^\top \mathbf{I} \vec{\omega}(t) \quad (\text{IV.34})$$

mit dem zu $\vec{\omega}(t)$ transponierten Zeilenvektor $\vec{\omega}(t)^\top$. Für den Eigendrehimpuls (IV.30a) gilt

$$\vec{L} = \mathbf{I} \vec{\omega}. \quad (\text{IV.35})$$

⁽⁴³⁾Einige Definitionen und Ergebnisse über Tensoren werden im Anhang A zusammengefasst.

Entsprechend der tensoriellen Natur von \mathbf{I} können Rotationsenergie und Drehimpuls noch in „geometrischer“ Form geschrieben werden:

$$T_{\text{Rot.}} = \frac{1}{2} \vec{\omega}(t) \cdot \mathbf{I} \cdot \vec{\omega}(t), \quad \vec{L} = \mathbf{I} \cdot \vec{\omega}$$

wobei \cdot die *Kontraktion* zweier Tensoren bezeichnet.

Trägheitsmoment

Sei \vec{e} der Einheitsvektor entlang einer beliebigen Richtung in \mathbb{R}^3 . Das *Trägheitsmoment* des starren Körpers bezüglich der (Dreh)Achse parallel zu dieser Richtung, die durch den Ursprungspunkt des körperfesten Bezugssystem geht, wird durch

$$I = e^i I^{ij} e^j = \vec{e}^T \mathbf{I} \vec{e} = \vec{e} \cdot \mathbf{I} \cdot \vec{e} \quad (\text{IV.36})$$

definiert, wobei $\{e^i\}_{i=1,2,3}$ die Koordinaten von \vec{e} bezeichnen. Beispielsweise ist das relevante Trägheitsmoment für eine Rotationsbewegung um eine Achse in Richtung des Basisvektors \vec{e}_3 einfach die Komponente $I^{33} = \vec{e}_3^T \mathbf{I} \vec{e}_3$. Dabei gilt

$$I^{33} = \sum_a m_a \left[(\vec{x}_a)^2 \delta^{33} - x_a^3 x_a^3 \right] = \sum_a m_a \left[(x_a^1)^2 + (x_a^2)^2 \right]$$

d.h. unter Einführung der Projektion $\vec{x}_{a,\perp}$ von \vec{x}_a auf der (x^1, x^2) -Ebene orthogonal zur Drehachse

$$I^{33} = \sum_a m_a (\vec{x}_{a,\perp})^2. \quad (\text{IV.37a})$$

Dabei ist $(\vec{x}_{a,\perp})^2$ das Quadrat des Abstands des Massenpunkts a von der Drehachse. Im Kontinuumslimites wird dieses Trägheitsmoment zu

$$I^{33} = \int \rho(\vec{r}) \vec{r}_{\perp}^2 d^3 \vec{r}. \quad (\text{IV.37b})$$

Für eine Rotationsbewegung mit Winkelgeschwindigkeit $\vec{\omega} = \omega \vec{e}_3$ um diese Richtung ist die Rotationsenergie $T_{\text{Rot.}} = \frac{1}{2} I^{33} \omega^2$. Da I^{33} automatisch positiv ist, gilt das auch für diese kinetische Energie.

Beispiel: homogener Zylinder

Sei ein homogener Vollzylinder mit Massendichte ρ , Radius R und Höhe h . Dementsprechend ist seine Gesamtmasse $M = \pi R^2 h \rho$. Der Zylinder dreht sich um seine eigene Achse, welche die Richtung x^3 definiert.

In Zylinderkoordinaten $(r, \theta, z \equiv x^3)$ ist der Abstand eines Punkts von der Drehachse genau gleich der Radialkoordinate r . Somit lautet das durch Gl. (IV.37b) gegebene Trägheitsmoment des Zylinders um diese Achse

$$I^{33} = \int_0^h \left[\int_0^{2\pi} \left(\int_0^R \rho r^2 r dr \right) d\theta \right] dz = \rho \frac{\pi R^4 h}{2} = \frac{MR^2}{2}. \quad (\text{IV.38})$$

Das letztere Ergebnis ist unabhängig von der Höhe h (genauer ist diese versteckt in der Gesamtmasse), so dass es auch im Grenzfall einer unendlich dünnen Scheibe $h \rightarrow 0$ gilt.

Eigenschaften des Trägheitstensors

Wie schon erwähnt wurde ist der Trägheitstensor ein symmetrischer Tensor, d.h. $I^{ij} = I^{ji}$ für alle $i, j = 1, 2, 3$. Dies bedeutet zuerst, dass nur 6 seiner Komponenten unabhängig voneinander sind.

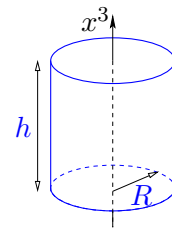


Abbildung IV.4

Ähnlich einer symmetrischen reellen Matrix besitzt ein symmetrischer reeller Tensor zweiter Stufe nur reelle Eigenwerte und ist orthogonal diagonalisierbar. Das heißt, man kann eine Orthonormalbasis $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ finden, in welcher der Tensor, oder genauer seine Matrixdarstellung, die Diagonalform

$$\mathbf{I} = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} \quad (\text{IV.39})$$

annimmt, mit den reellen Eigenwerten I_1, I_2, I_3 der Matrix, die wie oben schon bemerkt alle positiv sind. Diese Eigenwerte werden *Hauptträgheitsmomente* des starren Körpers genannt, während die zugehörigen Eigenvektoren $\{\vec{e}_i\}_{i=1,2,3}$ Einheitsvektoren entlang der *Hauptträgheitsachsen* sind.

Im *Hauptachsensystem* — d.h. im Koordinatensystem, dessen Achsen die Hauptträgheitsachsen des starren Körpers sind — nimmt die Rotationsenergie des Körpers die einfache Form

$$T_{\text{Rot.}} = \frac{1}{2} \left[I_1 (\omega^1)^2 + I_2 (\omega^2)^2 + I_3 (\omega^3)^2 \right]$$

an. Dann folgt aus Gl. (IV.35), dass wenn die Winkelgeschwindigkeit $\vec{\omega}$ entlang einer Hauptträgheitsachse ist, so ist der Drehimpuls auch parallel dazu.

Bemerkung: Genau wie der Trägheitstensor selber hängen seine Hauptträgheitselemente von der Wahl des Bezugspunkts O ab. Die folgende Eigenschaft bezieht sich auf den Fall, wo der Bezugspunkt im Schwerpunkt des starren Körpers ist.

Besitzt ein starrer Körper Symmetrien, wie z.B. Invarianz unter der Spiegelung bezüglich einer Ebene oder Rotationssymmetrie, so sind die entsprechenden Symmetrieelemente, insbesondere Symmetrieachsen, einfach mit den Hauptträgheitsachsen relativ zum Körperschwerpunkt verknüpft. Um dieses Ergebnis zu illustrieren, betrachten wir die nichtdiagonalen Elemente des Trägheitstensors, z.B. I^{12} . Die Definition (IV.26b) gibt

$$\begin{aligned} I^{12} &= \int \rho(\vec{r}) (\vec{r}^2 \delta^{12} - x^1 x^2) d^3 \vec{r} = \int \rho(x^1, x^2, x^3) (-x^1 x^2) dx^1 dx^2 dx^3 \\ &= \frac{1}{2} \int \rho(x^1, x^2, x^3) (-x^1 x^2 - x^1 x^2) dx^1 dx^2 dx^3. \end{aligned}$$

Wie unter Gl. (IV.26b) schon diskutiert wurde, kann der Integrationsbereich mit einer geeigneten Definition von ρ auf \mathbb{R}^3 erweitert werden, was wir hier annehmen. Führt man die Substitution $x^1 \rightarrow u \equiv -x^1$ für den zweiten Summanden in den Klammern, so ergibt sich

$$\begin{aligned} I^{12} &= \frac{1}{2} \left[- \int [\rho(x^1, x^2, x^3) x^1 x^2 dx^1 dx^2 dx^3 - \int \rho(-u, x^2, x^3) (-u) x^2 du dx^2 dx^3] \right] \\ &= \frac{1}{2} \int [-\rho(x^1, x^2, x^3) + \rho(-x^1, x^2, x^3)] x^1 x^2 dx^1 dx^2 dx^3, \end{aligned}$$

wobei in der zweiten Zeile die Integrationsvariable u des zweiten Terms in x^1 umbenannt wurde. Falls der starre Körper symmetrisch unter Spiegelungen bezüglich der Ebene $x^1 = 0$ ist, so dass $\rho(-x^1, x^2, x^3) = \rho(x^1, x^2, x^3)$ für alle Werte von x^2, x^3 , dann ist $I^{12} = 0$.

Allgemeiner findet man, dass die Symmetrieachsen des starren Körpers, falls einige vorhanden sind, auch seine Hauptträgheitsachsen sind.

Klassifikation von Trägheitstensoren

Je nach den Werten der Hauptträgheitsmomente lassen sich drei Fälle unterscheiden:

- Für $I_1 \neq I_2 \neq I_3 \neq I_1$ sind die Eigenvektoren $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ und somit die Hauptträgheitsachsen eindeutig⁽⁴⁴⁾ definiert. Dann spricht man von einem „unsymmetrischen Kreisel“.

⁽⁴⁴⁾ ... bis auf ein Minus Zeichen für die Einheitseigenvektoren.

Ein einfaches Beispiel davon ist ein homogener Quader, dessen Kanten unterschiedliche Längen $a \neq b \neq c \neq a$ haben: die Symmetrieachsen des Quaders sind seine Hauptträgheitsachsen (relativ zum Schwerpunkt), die zugehörigen Trägheitsmomente sind unterschiedlich.

- Für $I_1 = I_2 \neq I_3$ ist der Eigenvektor \vec{e}_3 bzw. die entsprechende Hauptachse eindeutig definiert. Dagegen gibt es Entartung in der (x^1, x^2) -Ebene, in welcher jedes Paar von orthogonalen Achsen als Hauptträgheitsachsen betrachtet werden kann. Dies entspricht dem Modell des *symmetrischen Kreisels*.

Beispiele sind ein eindimensionaler Stab parallel zur x^3 -Achse: $x^1 = x^2 = 0$ geben sofort $I^{33} = 0$ und $I^{11} = I^{22} \neq 0$; oder ein homogener Zylinder (im allgemeinen Fall), insbesondere eine zweidimensionale Scheibe in der Ebene $x^3 = 0$, für die man einfach $I^{11} = I^{22} = \frac{1}{2}I^{33}$ nachprüft.

- Für $I_1 = I_2 = I_3$ sind die Achsen jedes Koordinatensystems Hauptträgheitsachsen, entsprechend einem *Kugelkreisel*.

Dies ist z.B. der triviale Fall des Trägheitstensors einer homogenen Kugel um drei Achsen, die durch ihren Schwerpunkt verläuft. Falls der Radius der Kugel verschwindet, d.h. die Kugel wird zu einem Punkt, ist $\mathbf{I} = \mathbf{0}$: somit verschwinden der Eigendrehimpuls und die Rotationsenergie eines Massenpunkts, wie bisher stillschweigend angenommen wurde!

Offensichtlich sind die drei (Haupt)Trägheitsmomente eines homogenen Würfels um seine drei Symmetrieachsen alle gleich, d.h. der Matrixdarstellung des Trägheitstensors \mathbf{I} ist in der zugehörigen Basis proportional zur Einheitsmatrix. Dann gilt dies noch in jeder beliebigen Basis, so dass die Momente um jedes Triplet von Achsen, die durch seinen Schwerpunkt gehen, sind alle gleich.

Während die Unabhängigkeit des Trägheitsmoments von der Richtung einer durch den Schwerpunkt durchlaufenden Drehachse intuitiv im Fall der Kugel ist, wirkt sie bei dem Würfel überraschend. Dabei ist unsere⁽⁴⁵⁾ Intuition falsch, denn wir stellen uns die ganzen Massenverteilungen $\rho(\vec{r})$ vor — die offensichtlich unterschiedlich für eine Kugel und einen Würfel sind —, während das Trägheitsmoment bzw. der Trägheitstensor nur einem sehr geringen Anteil der in dieser Massenverteilungen enthaltenen Information entspricht. Für die Rotationsbewegung der Körper reicht aber diese Information aus.

Steinerscher Satz

Trägheitsmomente um die Symmetrieachsen eines starren Körpers — d.h. entsprechend dem Fall, wo der Ursprungspunkt des Koordinatensystems in der Definition (IV.26) im Schwerpunkt des Körpers liegt — sind unter Betrachtung der betreffenden Symmetrie meist einfacher zu berechnen. Oft dreht sich der starre Körper aber um eine Achse, die nicht durch seinen Schwerpunkt durchläuft. Anstatt das Trägheitsmoment um die verschobene Drehachse explizit aus Definition (IV.26) zu berechnen, ist es dann schneller, das folgende Resultat zu benutzen.

Theorem (Satz von Steiner^(t)):

Das Trägheitsmoment eines starren Körpers mit Masse M um eine beliebige Achse im Abstand L von seinem Schwerpunkt ist gegeben durch die Summe aus ML^2 und dem Trägheitsmoment des Körpers um die parallele Drehachse, die durch den Schwerpunkt geht.

(IV.40a)

Bezeichnet man mit I bzw. I' das Drehmoment um eine durch den Schwerpunkt durchlaufende Drehachse bzw. um eine dazu parallele Achse im Abstand L vom Schwerpunkt, so lautet der

⁽⁴⁵⁾ ... oder zumindest die des Autors!

^(t) J. STEINER, 1796–1863

Steinersche Satz

$$I' = I + ML^2. \quad (\text{IV.40b})$$

Zum Beweis dieses Satzes betrachte man neben dem körperfesten Bezugssystem \mathcal{B} mit Ursprung im Schwerpunkt des starren Körpers ein zweites körperfestes Bezugssystem \mathcal{B}' , dessen Nullpunkt um \vec{b} relativ zu \mathcal{B} verschoben ist, während die Koordinatenachsen von \mathcal{B}' parallel zu denen von \mathcal{B} bleiben. Dann gilt für die Position jedes (Massen)Punkts⁽⁴⁶⁾ des starren Körpers relativ zu beiden Bezugssystemen $\vec{x}'_a = \vec{x}_a - \vec{b}$. Dies gibt für den Trägheitstensor relativ zu \mathcal{B}'

$$\begin{aligned} I'^{ij} &= \sum_a m_a \left[(\vec{x}'_a)^2 \delta^{ij} - x'^i_a x'^j_a \right] = \sum_a m_a \left[(\vec{x}_a - \vec{b})^2 \delta^{ij} - (x_a^i - b^i)(x_a^j - b^j) \right] \\ &= I^{ij} - 2\delta^{ij} \vec{b} \cdot \sum_a m_a \vec{x}_a + b^i \sum_a m_a x_a^j + b^j \sum_a m_a x_a^i + \sum_a m_a (\vec{b}^2 \delta^{ij} - b^i b^j). \end{aligned}$$

Dabei verschwinden der zweite, der dritte und der vierte Beiträge der zweiten Zeile, und die Summe der Massen im fünften Term kann durch die Gesamtmasse ersetzt werden. Somit ergibt sich die tensorielle Form des Steinerschen Satzes:

$$I'^{ij} = I^{ij} + M(\vec{b}^2 \delta^{ij} - b^i b^j). \quad (\text{IV.40c})$$

Um das Drehmoment um eine gegebene Achse zu erhalten, multipliziert man diese Gleichung mit den Komponenten des Einheitsvektors in Richtung der Achse, vgl. Gl. (IV.36), woraus sich Gl. (IV.40b) ergibt.

Beispiel: homogener Zylinder (2)

Wir betrachten nochmals den homogenen Vollzylinder (Masse M , Radius R) der Abb. (IV.4), der sich jetzt um eine Achse drehen soll, die parallel zu seiner Symmetrieachse ist, um den Abstand L aber verschoben ist.

Dann ist gemäß dem Steinerschen Satz (IV.40) das Trägheitsmoment des Zylinders um diese Achse

$$I'^{33} = I^{33} + ML^2 = \frac{MR^2}{2} + ML^2,$$

wobei der Ausdruck (IV.38) des Trägheitsmoments um die Symmetrieachse benutzt wurde.

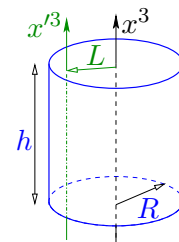


Abbildung IV.5

Literatur zum Kapitel IV

- Feynman, *Vorlesungen über Physik. Band 1* [24] = *Lectures on Physics. Vol. I* [25], Kap. 18–20 & *Band 2* [22] = *Vol. II* [23], Kap. 31.4;
- Fließbach, *Mechanik* [2] Teil V, Kap. 19–23 & Teil VI, Kap. 24–26.
- Goldstein, *Klassische Mechanik* [4] = *Classical Mechanics* [5], Kap. 4.1–4.4, 5, & 6.1–6.3.
- Greiner, *Klassische Mechanik II* [7] Kap. III & IV.
- Nolting, *Klassische Mechanik* [15] Kap. 4.
- Scheck, *Mechanik* [18] Kap. 3.

⁽⁴⁶⁾Hier wird der Beweis für ein System aus diskreten Massenpunkten dargelegt. Der Beweis lässt sich ohne Schwierigkeit auf den Fall eines kontinuierlichen Systems übersetzen.

KAPITEL V

Hamilton-Formalismus

V.1	Hamiltonsche Bewegungsgleichungen	106
V.1.1	Kanonisch konjugierter Impuls	107
V.1.2	Hamilton-Funktion	107
V.1.3	Kanonische Bewegungsgleichungen	108
V.1.4	Beispiele	110
V.2	Phasenraum	111
V.2.1	Phasenraumtrajektorien	111
V.2.2	Satz von Liouville	115
V.3	Poisson-Mechanik	118
V.3.1	Phasenraum-Funktionen	118
V.3.2	Poisson-Klammer	118
V.3.3	Poisson-Klammer und Zeitentwicklung	120
V.3.4	Kanonische Transformationen	122
V.3.5	Poisson-Klammer und Symmetrien	125

Dieses Kapitel befasst sich mit einer weiteren Formulierung der klassischen Mechanik. Diese beruht auf alternativen Variablen — den (verallgemeinerten) Koordinaten und den dazu konjugierten Impulsen —, woraus natürlich neue Bewegungsgleichungen folgen, die sog. *Hamilton-Gleichungen* (Abschn. V.1). Die möglichen Lösungen dieser Gleichungen für ein bestimmtes physikalisches System lassen sich auf natürliche Weise interpretieren als die Trajektorien eines fiktiven Punkts in einem (meistens) hochdimensionalen Raum, dem *Phasenraum* des Systems (Abschn. V.2). Schließlich wird in Abschn. V.3 eine alternative Form der Hamiltonschen Bewegungsgleichungen eingeführt. Anhand dieser Form kann der Zusammenhang zwischen Erhaltungsgrößen und Symmetrien eines Systems leicht ausgedrückt werden.

Wie wir hiernach sehen werden, besitzen die Grundgleichungen der Hamilton-Formulierung einen größeren Grad an Symmetrie als im Lagrange-Formalismus, was natürlich „ästhetisch“ wirkt — was bei manchen theoretischen Physikern wichtig ist. Daneben ist der Hamilton-Formalismus aus zwei anderen Gründen interessant. Zum einen nähert er sich am meistens dem kanonischen Formalismus der Quantenmechanik. Zum anderen stellt er auch einen natürlichen Rahmen für die Formulierung der (klassischen) statistischen Mechanik von Vielteilchensystemen dar.

V.1 Hamiltonsche Bewegungsgleichungen

In diesem Abschnitt werden die Grundlagen der Hamilton-Formulierung der klassischen Mechanik eingeführt, und zwar die Grundvariablen (§ V.1.1) und die davon abhängige Grundgröße (§ V.1.2), sowie die Bewegungsgleichungen (§ V.1.3). Die ganze Konstruktion wird dann in § V.1.4 auf einige Beispiele angewandt.

V.1.1 Kanonisch konjugierter Impuls

Sei ein mechanisches System mit s Freiheitsgraden. Im Rahmen des im Kap. III eingeführten Lagrange-Formalismus benutzt man zur Beschreibung dieses Systems und dessen Zeitentwicklung einerseits s verallgemeinerte Koordinaten

$$\mathbf{q} \equiv \{q^\alpha\}_{\alpha=1,\dots,s} \quad (\text{V.1a})$$

und andererseits die zugehörigen verallgemeinerten Geschwindigkeiten

$$\dot{\mathbf{q}} \equiv \{\dot{q}^\alpha\}_{\alpha=1,\dots,s}. \quad (\text{V.1b})$$

Zusammen mit der Zeit bilden die generalisierten Koordinaten und Geschwindigkeiten die $2s + 1$ Argumente der Lagrange-Funktion $\mathcal{L}(t, \mathbf{q}, \dot{\mathbf{q}})$, die die ganze Information über das System enthält.

Mithilfe der Lagrange-Funktion werden noch die verallgemeinerten Impulse definiert; somit ist

$$p_\alpha \equiv \frac{\partial \mathcal{L}(t, \mathbf{q}, \dot{\mathbf{q}})}{\partial \dot{q}^\alpha} \quad (\text{V.1c})$$

der zur Koordinate q^α konjugierte Impuls.

Im Folgenden werden diese Impulse kollektiv mit

$$\mathbf{p} \equiv \{p_\alpha\}_{\alpha=1,\dots,s} \quad (\text{V.1d})$$

bezeichnet.

Die Grundidee des Hamilton-Formalismus besteht darin, den Bewegungszustand eines mechanischen Systems zur Zeit t durch die verallgemeinerten Koordinaten $\{q^\alpha\}$ und die dazu konjugierten Impulse $\{p_\alpha\}$ zu charakterisieren, statt durch die Variablen $\{q^\alpha\}, \{\dot{q}^\alpha\}$.

Definition: Der von den $2s$ Variablen $\{q^\alpha\}$ und $\{p_\alpha\}$ aufgespannte Raum Γ heißt *Phasenraum* des Systems. Dementsprechend werden die $\{q^\alpha\}$ und $\{p_\alpha\}$ für $\alpha \in \{1, \dots, s\}$ gemeinsam *Phasenraumkoordinaten* genannt.

Jedem möglichen Bewegungszustand eines gegebenen Systems wird ein Punkt in dessen Phasenraum Γ zugeordnet, und die Bewegung entspricht einer Phasenraumtrajektorie. Dieses Thema wird im Abschn. V.2 ausführlicher diskutiert.

Bemerkung: Der Phasenraum ist im Allgemeinen kein Vektorraum, weil einige Variablen — wie z.B. Winkel — ihre Werte nur in einem endlichen Intervall von \mathbb{R} annehmen können.

Mathematisch ist der Phasenraum eine *Mannigfaltigkeit* der Dimension $2s$.

V.1.2 Hamilton-Funktion

Um die Variablen $\{q^\alpha\}, \{\dot{q}^\alpha\}$ des Lagrange-Formalismus durch die Koordinaten $\{q^\alpha\}$ und die konjugierten Impulse $\{p_\alpha\}$ zu ersetzen, muss eine Funktion der neuen Variablen definiert werden, welche die gesamte Information über das System enthält, um die Rolle der Lagrange-Funktion zu übernehmen.

Definition: Gegeben die Lagrange-Funktion eines System wird seine *Hamilton-Funktion* durch

$$\mathcal{H}(t, \mathbf{q}, \mathbf{p}) \equiv \sum_{\alpha=1}^s p_\alpha \dot{q}^\alpha - \mathcal{L}(t, \mathbf{q}, \dot{\mathbf{q}}) \quad (\text{V.2})$$

definiert.

Auf der rechten Seite dieser Definition sollen die Geschwindigkeiten \dot{q}^α als Funktionen der Zeit t und der Phasenraumkoordinaten $\{q^\beta\}$ und $\{p_\beta\}$ betrachtet werden: $\dot{q}^\alpha = \dot{q}^\alpha(t, \mathbf{q}, \mathbf{p})$. Das heißt, dass die definierende Beziehung des konjugierten Impulses (V.1c) invertiert werden soll. Wenn dies möglich ist, gilt bei festen $\{p_\beta\}_{\beta=1, \dots, s}$

$$\frac{\partial \mathcal{H}}{\partial \dot{q}^\alpha} = \frac{\partial}{\partial \dot{q}^\alpha} \left(\sum_{\beta=1}^s p_\beta \dot{q}^\beta - \mathcal{L} \right) = p_\alpha - \frac{\partial \mathcal{L}}{\partial \dot{q}^\alpha} = 0,$$

d.h. \mathcal{H} hängt nicht explizit von \dot{q}^α ab.

Laut dem Satz von der Umkehrabbildung soll die Determinante der $s \times s$ -Matrix mit Elementen $\partial^2 \mathcal{L} / \partial \dot{q}^\alpha \partial \dot{q}^\beta$ ungleich Null sein, damit Gl. (V.1c) lokal invertierbar sei; vgl. auch Anhang D.

Wie oben erwähnt muss die Hamilton-Funktion \mathcal{H} die gleiche Information wie die Lagrange-Funktion \mathcal{L} enthalten. Dass dies der Fall ist, lässt sich beweisen, indem man \mathcal{L} aus \mathcal{H} rekonstruieren kann.

Sei somit angenommen, dass die Hamilton-Funktion $\mathcal{H}(t, \mathbf{q}, \mathbf{p})$ bekannt ist. Definiert man zunächst

$$\dot{Q}^\alpha \equiv \frac{\partial \mathcal{H}(t, \mathbf{q}, \mathbf{p})}{\partial p_\alpha},$$

so folgt aus Gl. (V.2) und der Kettenregel

$$\dot{Q}^\alpha = \frac{\partial}{\partial p_\alpha} \left(\sum_{\beta=1}^s p_\beta \dot{q}^\beta - \mathcal{L} \right) = \dot{q}^\alpha + \sum_{\beta=1}^s p_\beta \frac{\partial \dot{q}^\beta}{\partial p_\alpha} - \sum_{\beta=1}^s \left(\frac{\partial \mathcal{L}}{\partial q^\beta} \frac{\partial q^\beta}{\partial p_\alpha} + \frac{\partial \mathcal{L}}{\partial \dot{q}^\beta} \frac{\partial \dot{q}^\beta}{\partial p_\alpha} \right).$$

Unter Verwendung von $\partial q^\beta / \partial p_\alpha = 0$, entsprechend der Unabhängigkeit der Variablen q_β und p_α , und der Beziehung (V.1c) im letzten Term kommt dann $\dot{Q}^\alpha = \dot{q}^\alpha$. Somit lassen sich die verallgemeinerten Geschwindigkeiten aus der Hamilton-Funktion über

$$\dot{q}^\alpha(t, \mathbf{q}, \mathbf{p}) = \frac{\partial \mathcal{H}(t, \mathbf{q}, \mathbf{p})}{\partial p_\alpha} \quad (\text{V.3a})$$

wiederfinden. Dann liefert eine einfache Berechnung

$$\sum_{\alpha=1}^s p_\alpha \frac{\partial \mathcal{H}}{\partial p_\alpha} - \mathcal{H} = \sum_{\alpha=1}^s p_\alpha \dot{q}^\alpha - \left(\sum_{\alpha=1}^s p_\alpha \dot{q}^\alpha - \mathcal{L} \right) = \mathcal{L}.$$

Das heißt, die Lagrange-Funktion kann aus der Hamilton-Funktion gemäß

$$\mathcal{L}(t, \mathbf{q}, \dot{\mathbf{q}}) \equiv \sum_{\alpha=1}^s p_\alpha \dot{q}^\alpha(t, \mathbf{q}, \mathbf{p}) - \mathcal{H}(t, \mathbf{q}, \mathbf{p}) \quad (\text{V.3b})$$

rekonstruiert werden, wobei \dot{q}^α durch Gl. (V.3a) gegeben ist.

Mathematisch ist der Übergang von \mathcal{L} zu \mathcal{H} ein Beispiel von *Legendre^(u)-Transformation*,⁽⁴⁷⁾ und der Übergang von \mathcal{H} zu \mathcal{L} ist die Rücktransformation (*inverse Legendre-Transformation*).

Bemerkung: Vergleicht man Gl. (V.2) mit Gl. (III.35c), so stimmt die Definition der Hamilton-Funktion mit jener der Noether-Ladung überein, die mit Invarianz unter Zeittranslationen assoziiert und als gesamte Energie des Systems interpretiert wurde.

⁽⁴⁷⁾S. Anhang D.

^(u)A.-M. LEGENDRE, 1752–1833

V.1.3 Kanonische Bewegungsgleichungen

Die Position eines Systems im Phasenraum zu einer gegebenen Zeit t gibt seine verallgemeinerten Koordinaten und konjugierten Impulse zu diesem Zeitpunkt, die hiernach mit $\{q^\alpha(t)\}$ und $\{p_\alpha(t)\}$ bezeichnet werden. Zur Charakterisierung der Bewegung des Systems sind noch Bewegungsgleichungen erforderlich, welche die Zeitentwicklung von den $\{q^\alpha(t)\}$ und $\{p_\alpha(t)\}$ bestimmen.

Im vorigen Paragraphen wurde schon Gl. (V.3a) gefunden, welche die Zeitableitung der verallgemeinerten Koordinate $q^\alpha(t)$ durch Größen des Hamilton-Formalismus ausdrückt. Betrachtet man jetzt die Ableitung der Hamilton-Funktion (V.2) nach der verallgemeinerten Koordinate q^α , so ergibt sich

$$\frac{\partial \mathcal{H}}{\partial q^\alpha} = \frac{d\mathcal{H}}{dq^\alpha} = \frac{d}{dq^\alpha} \left(\sum_{\beta=1}^s p_\beta \dot{q}^\beta - \mathcal{L} \right) = \sum_{\beta=1}^s p_\beta \frac{\partial \dot{q}^\beta}{\partial q^\alpha} - \frac{\partial \mathcal{L}}{\partial q^\alpha} - \sum_{\beta=1}^s \frac{\partial \mathcal{L}}{\partial \dot{q}^\beta} \frac{\partial \dot{q}^\beta}{\partial q^\alpha}.$$

Im letzten Summanden kann $\partial \mathcal{L} / \partial \dot{q}^\beta$ durch p_β ersetzt werden, so dass dieser Term sich mit dem ersten herauskürzt. Wird der zweite Term $\partial \mathcal{L} / \partial q^\alpha$ dann im Punkt $(t, \mathbf{q}(t), \dot{\mathbf{q}}(t))$ der Trajektorie des Systems ausgewertet, so ist er laut der Euler-Lagrange-Gleichung (III.11) gleich der Zeitableitung von $\partial \mathcal{L} / \partial \dot{q}^\alpha$. Nach Gl. (V.1c) ist dies auch die Zeitableitung des konjugierten Impulses p_α . Somit ergibt sich

$$\frac{\partial \mathcal{H}(t, \mathbf{q}(t), \mathbf{p}(t))}{\partial q^\alpha} = - \frac{dp_\alpha(t)}{dt}.$$

Insgesamt gelten die (*kanonischen*) *Hamiltonschen Gleichungen*

$$\begin{aligned} \frac{dq^\alpha(t)}{dt} &= \frac{\partial \mathcal{H}(t, \mathbf{q}(t), \mathbf{p}(t))}{\partial p_\alpha} \\ \frac{dp_\alpha(t)}{dt} &= - \frac{\partial \mathcal{H}(t, \mathbf{q}(t), \mathbf{p}(t))}{\partial q^\alpha} \end{aligned} \quad \text{für } \alpha = 1, \dots, s. \quad (\text{V.4})$$

Im Gegensatz zum zweiten newtonschen Gesetz (I.14) oder zu den Euler-Lagrange-Gleichungen (III.11), die zweiter Ordnung sind, sind die $2s$ Hamiltonschen Bewegungsgleichungen Differentialgleichungen erster Ordnung. Daher ist für jede Phasenraumkoordinate nur eine einzige Anfangsbedingung nötig, um die Lösung der Bewegungsgleichungen, d.h. die Phasenraumtrajektorie, zu bestimmen.

Bemerkungen:

* Die totale Ableitung der Hamilton-Funktion nach der Zeit gibt unter Verwendung der Kettenregel und der Hamiltonschen Gleichungen (V.4)

$$\frac{d\mathcal{H}}{dt} = \frac{\partial \mathcal{H}}{\partial t} + \sum_{\alpha=1}^s \frac{\partial \mathcal{H}}{\partial q^\alpha} \dot{q}^\alpha + \sum_{\alpha=1}^s \frac{\partial \mathcal{H}}{\partial p_\alpha} \dot{p}_\alpha = \frac{\partial \mathcal{H}}{\partial t} + \sum_{\alpha=1}^s (-\dot{p}_\alpha \dot{q}^\alpha + \dot{q}^\alpha \dot{p}_\alpha) = \frac{\partial \mathcal{H}}{\partial t}.$$

Falls die Hamilton-Funktion nicht explizit von der Zeit abhängt ($\partial \mathcal{H} / \partial t = 0$), ist sie somit eine Konstante der Bewegung, entsprechend Energieerhaltung im System.

* Die Hamiltonschen Bewegungsgleichungen (V.4) können aus einem Extremalprinzip, dem schon angetroffenen Hamilton(!)-Prinzip (III.10), hergeleitet werden. Demgemäß ist die Wirkung

$$S[\mathbf{q}, \mathbf{p}] \equiv \int_{t_1}^{t_2} \left[\sum_{\alpha=1}^s p_\alpha(t) \dot{q}^\alpha(t, \mathbf{q}(t), \mathbf{p}(t)) - \mathcal{H}(t, \mathbf{q}(t), \mathbf{p}(t)) \right] dt \quad (\text{V.5})$$

unter allen Phasenraumtrajektorien mit festen Endpunkten extremal für die physikalisch realisierte Bewegung.

V.1.4 Beispiele

V.1.4 a Allgemeines eindimensionales System

Das einfachste Beispiel ist das eines zeittranslationsinvarianten Systems mit einem einzigen Freiheitsgrad, der durch eine verallgemeinerte Koordinate q parametrisiert wird. Sei

$$\mathcal{L}(q, \dot{q}) = \frac{m}{2} \dot{q}^2 - V(q) \quad (\text{V.6a})$$

die zugehörige Lagrange-Funktion, mit dem Potential V für die generalisierte Koordinate. Der zu q konjugierte Impuls ist

$$p \equiv \frac{\partial \mathcal{L}(q, \dot{q})}{\partial \dot{q}} = m\dot{q}, \quad (\text{V.6b})$$

woraus $\dot{q}(t, q, p) = p/m$ folgt.

Die Hamilton-Funktion (V.2) für dieses System lautet

$$\mathcal{H}(q, p) \equiv p\dot{q} - \mathcal{L}(q, \dot{q}) = m\dot{q}^2 - \frac{m}{2}\dot{q}^2 + V(q) = \frac{m}{2}\dot{q}^2 + V(q).$$

Ersetzt man \dot{q} durch p/m , so ergibt sich schließlich

$$\mathcal{H}(q, p) = \frac{p^2}{2m} + V(q). \quad (\text{V.6c})$$

Ausgehend von dieser Hamilton-Funktion lauten die Hamiltonschen Gleichungen (V.4)

$$\frac{dq(t)}{dt} = \frac{\partial \mathcal{H}(q(t), p(t))}{\partial p} = \frac{p(t)}{m} \quad (\text{V.6d})$$

und

$$\frac{dp(t)}{dt} = -\frac{\partial \mathcal{H}(q(t), p(t))}{\partial q} = -\frac{\partial V(q(t))}{\partial q}. \quad (\text{V.6e})$$

Wird die zweite in der Ableitung der ersten nach der Zeit eingesetzt, so findet man die „übliche“ Bewegungsgleichung des Lagrange-Formalismus wieder.

Schließlich kann man schnell nachprüfen, dass die Lagrange-Funktion aus der Hamilton-Funktion rekonstruiert werden kann (s. Gl. (V.3b)):

$$p \frac{\partial \mathcal{H}(q, p)}{\partial p} - \mathcal{H}(q, p) = p \frac{p}{m} - \frac{p^2}{2m} - V(q) = \frac{p^2}{2m} - V(q).$$

V.1.4 b Harmonischer Oszillator

Ein wichtiges Beispiel für die im letzten Paragraphen gefundenen Ergebnisse ist das des eindimensionalen harmonischen Oszillators mit Potential $V(q) = \frac{1}{2}m\omega^2 q^2$. Mit der Hamilton-Funktion

$$\mathcal{H}(q, p) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2, \quad (\text{V.7})$$

lauten die Hamiltonschen Bewegungsgleichungen

$$\frac{dq(t)}{dt} = \frac{p(t)}{m} \quad (\text{V.8a})$$

$$\frac{dp(t)}{dt} = -m\omega^2 q(t). \quad (\text{V.8b})$$

Diese zwei gekoppelten Differentialgleichungen können natürlich kombiniert werden, um die übliche Differentialgleichung zweiter Ordnung $m\ddot{q}(t) + m\omega^2 q(t) = 0$ zu geben. Stattdessen wird hiernach eine geeignete Linearkombination von $q(t)$ und $p(t)$ eingeführt, um eine einfach lösbare Gleichung erster Ordnung zu erhalten.

Sei

$$\alpha(t) \equiv \sqrt{\frac{m\omega}{2}} q(t) + i \frac{p(t)}{\sqrt{2m\omega}}. \quad (\text{V.9})$$

Die Phasenraumtrajektorie lässt sich durch $\alpha(t)$ und die komplex konjugierte Funktion $\alpha(t)^*$ ausdrücken:

$$q(t) = \frac{1}{\sqrt{2m\omega}} [\alpha(t) + \alpha(t)^*] = \sqrt{\frac{2}{m\omega}} \operatorname{Re} \alpha(t) \quad (\text{V.10a})$$

$$p(t) = \sqrt{\frac{m\omega}{2}} \frac{\alpha(t) - \alpha(t)^*}{i} = \sqrt{2m\omega} \operatorname{Im} \alpha(t) \quad (\text{V.10b})$$

wobei $\operatorname{Re} \alpha(t)$ bzw. $\operatorname{Im} \alpha(t)$ den Realteil bzw. Imaginärteil von $\alpha(t)$ bezeichnet. Die Summe aus Gl. (V.8a), multipliziert mit $\sqrt{m\omega/2}$, und Gl. (V.8b) multipliziert mit $i/\sqrt{2m\omega}$ gibt

$$\frac{d\alpha(t)}{dt} = -i\omega \alpha(t). \quad (\text{V.11})$$

Diese gewöhnliche Differentialgleichung erster Ordnung lässt sich sofort lösen: wenn α_0 die Anfangsbedingung bei $t = 0$ bezeichnet, gilt

$$\alpha(t) = \alpha_0 e^{-i\omega t}, \quad (\text{V.12})$$

woraus $q(t)$ und $p(t)$ über Gl. (V.10) folgen. Insbesondere ergibt sich das bekannte oszillatorische Verhalten von $q(t)$.

Bemerkung: Man kann noch die Hamilton-Funktion durch die Variablen α, α^* ausdrücken:

$$\mathcal{H} = \omega \alpha^* \alpha.$$

Die Transformation $(q, p) \rightarrow (\alpha, \alpha^*)$ ist aber keine kanonische Transformation (vgl. § V.3.4).

V.2 Phasenraum

Jeder Bewegungszustand eines mechanischen Systems lässt sich durch einen Punkt $(\{q^\alpha\}, \{p_\alpha\})$ im Phasenraum Γ darstellen, während die Bewegung im physikalischen Raum durch eine Phasenraumtrajektorie dargestellt wird. Insbesondere können die Gleichgewichtspositionen eines Systems sowie ihre Stabilität anhand des Verhaltens der Trajektorie im Phasenraum des Systems einfach erkannt werden.

In § V.2.1 werden die Phasenraumtrajektorien einiger eindimensionaler Probleme vorgestellt und ihre Eigenschaften diskutiert. Dann wird in § V.2.2 ein wichtiges Ergebnis dargelegt, das Liouville-Theorem, das insbesondere in der klassischen statistischen Physik eine große Rolle spielt.

V.2.1 Phasenraumtrajektorien

V.2.1 a Erste Eigenschaften

Sei $(\mathbf{q}(t_0), \mathbf{p}(t_0))$ der Bewegungszustand eines Systems zu irgendeinem Zeitpunkt t_0 . Laut den kanonischen Gleichungen (V.4) ist der Zustand des Systems zu einer späteren Zeit völlig bestimmt: z.B. gilt nach einer infinitesimal kleinen Zeitspanne δt

$$q^\alpha(t_0 + \delta t) \simeq q^\alpha(t_0) + \frac{dq^\alpha(t_0)}{dt} \delta t = q^\alpha(t_0) + \frac{\partial \mathcal{H}(t_0, \mathbf{q}(t_0), \mathbf{p}(t_0))}{\partial p_\alpha} \delta t$$

$$p_\alpha(t_0 + \delta t) \simeq p_\alpha(t_0) + \frac{dp_\alpha(t_0)}{dt} \delta t = p_\alpha(t_0) - \frac{\partial \mathcal{H}(t_0, \mathbf{q}(t_0), \mathbf{p}(t_0))}{\partial q^\alpha} \delta t$$

für alle $\alpha \in \{1, \dots, s\}$, wobei \mathcal{H} die Hamilton-Funktion des Systems ist.

Wenn \mathcal{H} nicht explizit von der Zeit abhängt, wie wir in diesem Abschnitt annehmen werden, sind ihre partiellen Ableitungen in einem Punkt von Γ völlig durch die Koordinaten des Punkts bestimmt. Somit kann in diesem Fall durch jeden (physikalisch realisierbaren) Punkt des Phasenraums nur eine einzige Trajektorie gehen, d.h. Phasenraumtrajektorien kreuzen sich nicht. ⁽⁴⁸⁾

⁽⁴⁸⁾Die scheinbaren Ausnahmen zu dieser Regel in Abb. V.3 und V.4 werden hiernach erklärt.

Bei periodischen Bewegungen befindet sich das System nach einer Periodendauer \mathcal{T} erneut im gleichen Bewegungszustand. Dementsprechend kommt die Phasenraumtrajektorie des Systems zurück zum gleichen Punkt: die Trajektorie in Γ ist dann eine geschlossene Kurve.

Ein Sonderfall davon besteht aus den in Abschn. IV.1 eingeführten Gleichgewichtslösungen des Problems: da die Position zeitunabhängig ist, bleiben die generalisierten Koordinaten q konstant. dazu verschwinden die verallgemeinerten Kräfte, d.h. die Zeitableitung der Impulse, so dass auch die letzteren konstant bleiben. Demzufolge wird eine solche stationäre Lösung im Phasenraum durch einen einzigen Punkt, einen *Fixpunkt*, dargestellt.

Dabei lassen sich stabile und instabile Gleichgewichtslösungen noch leicht voneinander unterscheiden. In der Nachbarschaft eines stabilen Gleichgewichts finden nämlich kleine Schwingungen statt, die wie oben gesagt durch geschlossene Kurven im Phasenraum dargestellt sind. Im Gegensatz dazu befinden sich in der Nähe einer instabilen Gleichgewichtslösung sowohl Trajektorien, die sich dem Fixpunkt nähern, als andere, die sich davon entfernen.

V.2.1 b Beispiele

Der Phasenraum Γ eines Systems mit s Freiheitsgraden ist $2s$ -dimensional: somit lassen sich Phasenräume günstig nur für Probleme mit $s = 1$ Freiheitsgrad darstellen. Im Folgenden werden die Phasenräume von drei solchen Problemen mit zeitunabhängiger Hamilton-Funktion diskutiert.

Harmonischer Oszillator

Betrachten wir zuerst den eindimensionalen harmonischen Oszillator mit Hamilton-Funktion

$$\mathcal{H}(q, p) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2. \quad (\text{V.13})$$

Das System wurde schon in § V.1.4 b untersucht; die Lösungen der Bewegungsgleichungen sind der Form

$$q(t) = A \cos \omega(t - t_0) \quad , \quad p(t) = -m\omega A \sin \omega(t - t_0), \quad (\text{V.14})$$

wobei A der Wert von q zur Zeit t_0 ist, während die letztere so gewählt wird, dass $p(t_0) = 0$.

In der (q, p) -Ebene, entsprechend dem Phasenraum des Oszillators, ist die durch t parametrisierte Kurve (V.14) für einen bestimmten Wert von A eine Ellipse mit Halbachsen A und $m\omega A$, deren Zentrum im Nullpunkt $q = 0, p = 0$ liegt. In Abb. V.1 werden einige solche Phasenraumtrajektorien

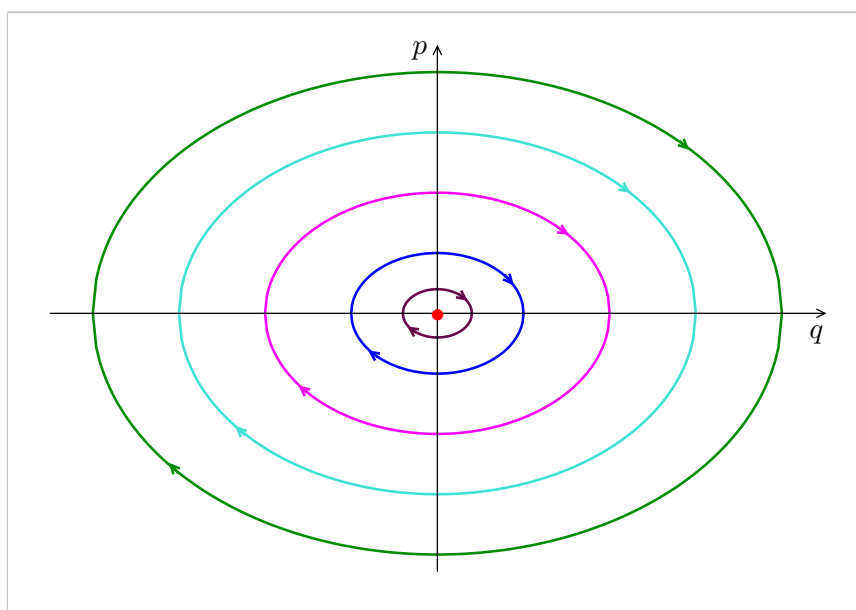


Abbildung V.1 – Phasenraumtrajektorien eines eindimensionalen harmonischen Oszillators.

dargestellt, insbesondere der Sonderfall mit $A = 0$, welcher der (stabilen) Gleichgewichtslösung des Systems entspricht.

Doppelmuldenpotential

Als nächstes Beispiel betrachten wir die eindimensionale Bewegung eines Massenpunkts in dem „Doppelmuldenpotential“

$$V(q) = V_0 - \frac{\mu}{2}(q - q_0)^2 + \frac{\lambda}{4}(q - q_0)^4 \quad \text{mit } \lambda, \mu > 0, \quad (\text{V.15})$$

wobei $q_0, V_0 \in \mathbb{R}$. Dieses Potential hat ein lokales Maximum bei $q = q_0$ und zwei globale Minima bei $q_1 = q_0 - \sqrt{\mu/\lambda}$ und $q_2 = q_0 + \sqrt{\mu/\lambda}$, wie in Abb. V.2 gezeigt wird.

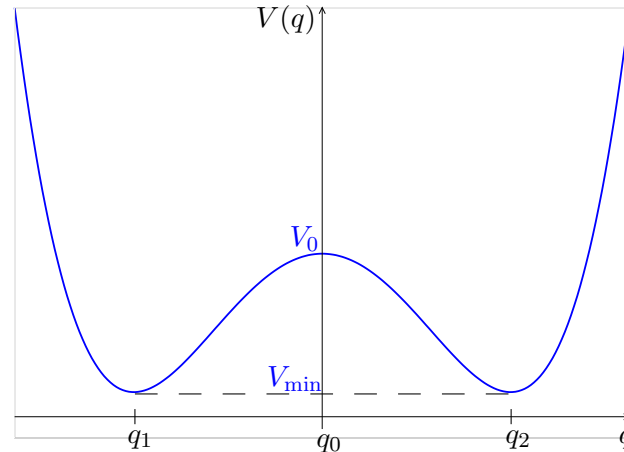


Abbildung V.2 – Verlauf des Doppelmuldenpotentials (V.15)

Sei $V_{\min} \equiv V(q_1) = V(q_2)$. Je nach dem Wert der Energie $E \geq V_{\min}$ des Massenpunkts treten unterschiedliche Verhalten auf, denen wiederum verschiedenartige Phasenraumtrajektorien zugeordnet sind, s. Abb. V.3.

- Wenn $E = V_{\min}$ kann sich der Massenpunkt entweder in q_1 oder in q_2 , d.h. in einer der stabilen Gleichgewichtspositionen, befinden. Im Phasenraum werden diese Möglichkeiten durch (rot gemalte) Fixpunkte dargestellt.
- Wenn $V_{\min} < E < V_0 = V(q_0)$ kann sich der Massenpunkt in einer der Mulden periodisch bewegen. Somit sind die entsprechenden Phasenraumtrajektorien geschlossene Kurven um den jeweiligen Fixpunkt.
- Der Energie $E = V_0$ entsprechen zwei Möglichkeiten:
 - Falls der Massenpunkt sich genau in q_0 befindet, dann hat er keine kinetische Energie. Dementsprechend wird das System in dieser Gleichgewichtsposition, die in Abb. V.3 mit einem schwarzen Punkt dargestellt ist, stationär bleiben.
 - Ist der Massenpunkt nicht in q_0 , sondern in einer der beiden Mulden, wird er sich bewegen. Sei z.B. angenommen, dass er sich in der rechten Mulde mit $q > q_0$ befindet — die Diskussion für den Fall $q < q_0$ ist analog. Wenn der Impuls p des Massenpunkts negativ ist, bewegt er sich nach links, bis er (asymptotisch) den Fixpunkt bei $q = q_0$ erreicht. Wenn dagegen p positiv ist, wird der Massenpunkt sich erstens weiter in die positive q -Richtung bewegen, bis er den Punkt rechts von q_2 erreicht, wo $V(q) = V_0$: in diesem Punkt⁽⁴⁹⁾ verschwindet sein Impuls, der danach negativ wird, so dass die Bewegung nun nach links bis zum Fixpunkt in $q = q_0$ stattfindet.

⁽⁴⁹⁾ Im physikalischen Raum handelt es sich um einen Umkehrpunkt.

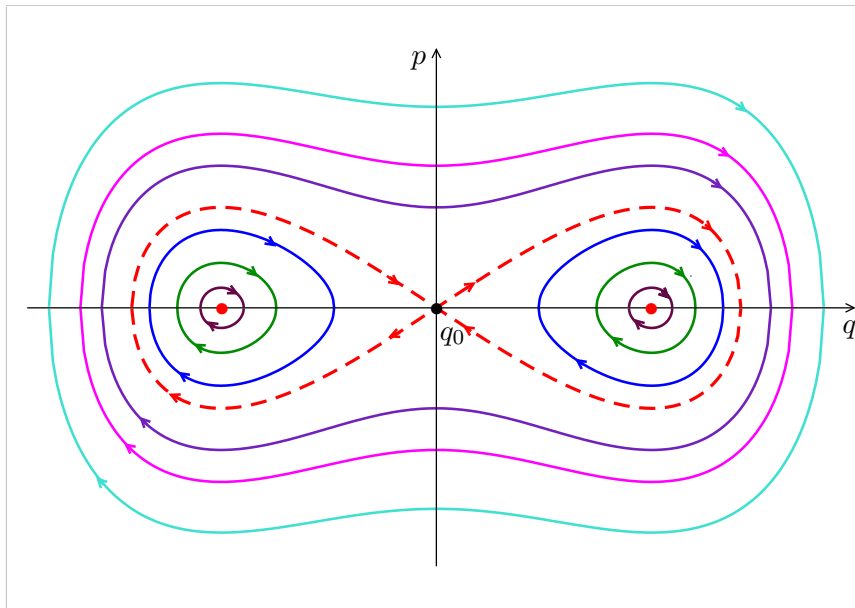


Abbildung V.3 – Phasenraum für die Bewegung im Doppelmuldenpotential (V.15).

Diesen Bewegungen wird im Phasenraum der Abb. V.3 die gestrichelte (rote) Kurve zugeordnet. Dabei bleibt der Massenpunkt je nach der Anfangsbedingung entweder rechts oder links von q_0 , und er nähert sich dem instabilen Fixpunkt nur für $t \rightarrow +\infty$. Somit treffen sich die zwei Teile der gestrichelten Kurve mit $q > q_0$ bzw. $q < q_0$ in der Tat nicht.

- Wenn $E > V_0$ kann sich der Massenpunkt in der ganzen Doppelmulde periodisch bewegen. Die zugehörigen Phasenraumtrajektorien sind geschlossene Kurven, welche die beiden stabilen Fixpunkte umkreisen und bei $q = q_0$ einen nicht-verschwindenden Impuls p — entweder positiv oder negativ — haben.

Die gestrichelte Kurve in Abb. V.3 trennt Phasenraumgebiete mit unterschiedlichen Verhalten — Bewegung in einer der kleinen Mulden oder in der ganzen Doppelmulde — voneinander. Eine solche Kurve wird *Separatrix* genannt.

Ebenes Pendel

Schließlich betrachten wir ein ebenes Pendel mit einer Masse m am Ende eines masselosen Stabs der Länge l in einem Schwerfeld (vgl. § III.2.4). Wenn φ den Ablenkwinkel von der Richtung des Schwerfeldes bezeichnet, führt die Standard-Lagrange-Funktion $\mathcal{L}(\varphi, \dot{\varphi}) = \frac{1}{2}ml^2\dot{\varphi}^2 + mgl \cos \varphi$ [Gl. (III.29)] zum konjugierten Impuls $p_\varphi = ml^2\dot{\varphi}$ und somit zur Hamilton-Funktion

$$\mathcal{H}(\varphi, p_\varphi) = \frac{p_\varphi^2}{2ml^2} - mgl \cos \varphi. \quad (\text{V.16})$$

Das System hat eine stabile Gleichgewichtsposition bei $\varphi = 0$ und eine instabile bei $\varphi = \pi$.

Auf erster Sicht könnte man die Werte des Winkels φ auf das Intervall $[-\pi, \pi]$ einschränken. Dies reicht aus, um die periodischen Schwingungen mit Amplitude $\varphi_0 < \pi$ um die stabile Gleichgewichtsposition zu beschreiben. Wie üblich werden solche Oszillationen im Phasenraum durch geschlossene Kurven um den Fixpunkt ($\varphi = 0, p_\varphi = 0$) dargestellt (Abb. V.4).

Wenn die Energie der Masse groß genug ist, kann das Pendel periodisch um den Aufhängepunkt drehen, anstatt Schwingungen durchzuführen. Um solche Bewegungen günstiger — d.h. ohne Sprung des Winkels — zu beschreiben, erweitert man die Wertemenge von φ auf die ganze reelle Gerade \mathbb{R} , wie in Abb. V.4 gemacht wird. Dann werden diese periodischen Rotationsbewegungen durch φ -periodische Kurven dargestellt — und nicht mehr durch geschlossene Kurven.

Die Drehbewegungen werden von den Schwingungsbewegungen durch Separatrizen getrennt, die durch die periodisch wiederholte Darstellung des instabilen Fixpunkts gehen. Wie in der Diskussion

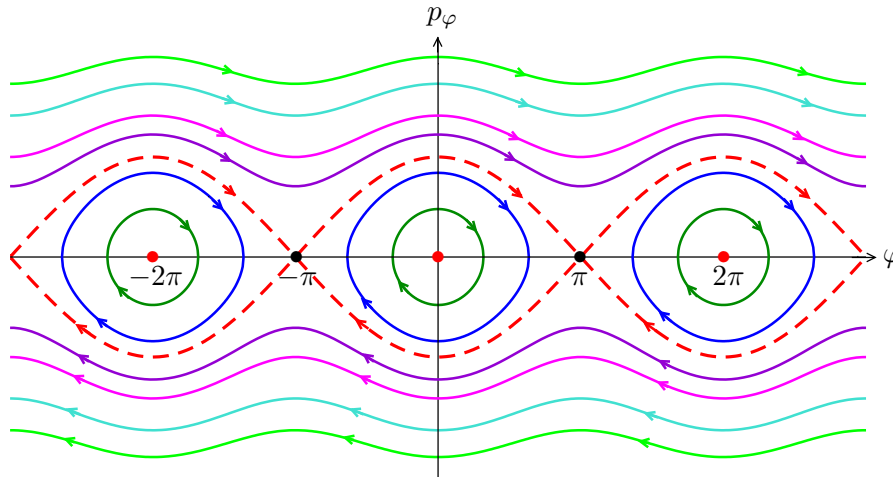


Abbildung V.4 – Phasenraum eines ebenen Pendels

des Phasenraums für die Bewegung im Doppelmuldenpotential entspricht die Portion einer Separatrix zwischen zwei sukzessiven Fixpunkten einer möglichen Bewegung, die einen Fixpunkt für $t \rightarrow +\infty$ erreicht: solche Phasenraumtrajektorien kreuzen eigentlich nicht die φ -Achse.

V.2.2 Satz von Liouville

Betrachten wir zu einer bestimmten Zeit t_0 ein ($2s$ -dimensionaler) Bereich $d\Gamma(t_0)$ im Phasenraum Γ eines mechanischen Systems mit s Freiheitsgraden und zeitunabhängiger Hamilton-Funktion \mathcal{H} . Die Punkte dieses Bereichs stellen unterschiedliche mögliche Bewegungszustände des Systems dar — die man sich auch als die jeweiligen Bewegungszustände von identischen Kopien des Systems vorstellen kann.

Zu einem späteren Zeitpunkt t_1 hat jede Phasenraumtrajektorie, dessen Punkt zu t_0 in $d\Gamma(t_0)$ lag, einen neuen Punkt des Phasenraums erreicht: sei $d\Gamma(t_1)$ der neue Bereich, den diese Punkte aufspannen. Dann ist das Volumen⁽⁵⁰⁾ von $d\Gamma(t_1)$ gleich dem von $d\Gamma(t_0)$, in Übereinstimmung mit dem

Theorem (Satz von Liouville^(v)):⁽⁵¹⁾

Für ein konservatives Hamiltonsches System bleibt das von benachbarten Phasenraumtrajektorien eingeschlossene Volumen konstant in der Zeitentwicklung. (V.17)

Eine erste Anwendung dieses Satzes wird in Abb. V.5 am Beispiel des eindimensionalen harmonischen Oszillators illustriert. In diesem Fall sind *alle* Phasenraumtrajektorien Ellipsen mit dem gleichen Zentrum, der gleichen Orientierung und der gleichen Exzentrizität, und die „Winkelgeschwindigkeit“ in der (q, p) -Ebene entlang dieser Trajektorien ist ebenfalls immer die gleiche: die Punkte des Phasenraums drehen sich um den Fixpunkt $(q = 0, p = 0)$ wie die Punkte eines rotierenden starren Körpers. Infolgedessen bewegen sich Phasenraumvolumina in dieser Zeitentwicklung ohne sich zu ändern, was genau dem Satz von Liouville entspricht.

Der harmonische Oszillator stellt einen Ausnahmefall dar, indem sich die Phasenraumgebiete in der Zeitentwicklung des Systems nicht verformen. Ein klassischeres Verhalten ist das des in Abb. V.6 gezeigten ebenen Pendels: in der Zeitentwicklung ändert ein typisches Phasenraumgebiet seine Form, auch wenn sein Volumen gemäß dem Satz von Liouville unverändert bleibt.

⁽⁵⁰⁾Um diese Volumina zu quantifizieren, müsste man ein Maß für die Gebiete von Γ spezifizieren.

⁽⁵¹⁾Dieses Theorem wird hier ohne Beweis angegeben.

^(v)J. LIOUVILLE, 1809–1882

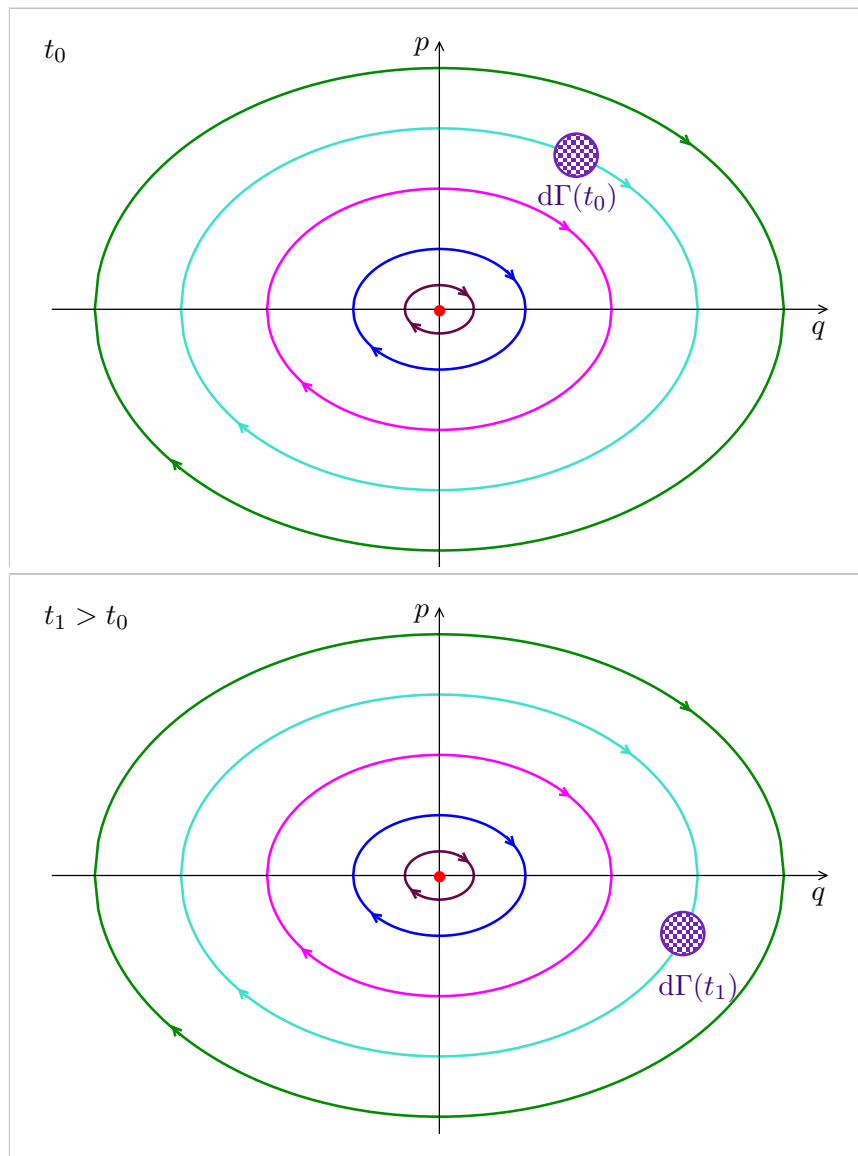


Abbildung V.5—Zeitentwicklung eines Volumenelements im Phasenraum des harmonischen Oszillators.

Bemerkungen:

* Betrachtet man die Punkte des Phasenraums eines Systems als Bestandteile eines fiktiven „kontinuierlichen Mediums“, so entspricht die Zeitentwicklung des Systems der Bewegung dieses Mediums. Wie oben schon gesagt, ist im Fall des eindimensionalen harmonischen Oszillators das entsprechende Phasenraum-Medium ein starrer Körper, und dessen Bewegung ist eine Drehbewegung.

Im allgemeineren Fall, z.B. für das ebene Pendel, ist das Phasenraum-Medium aber deformierbar, so dass es sich eher um ein *Fluid*⁽⁵²⁾ handelt. Da das Volumen jedes Teilbereichs dieses Fluids sich laut dem Liouville-Theorem in der Bewegung nicht ändert, kann dieses Fluid nicht komprimiert (oder ausgedehnt) werden: es verhält sich wie ein sog. *inkompressibles Fluid*.

* Am Beispiel des ebenen Pendels in Abb. V.6 wird auch ein anderes mögliches Verhalten illustriert. Zur Zeit t_0 ist das Phasenraumgebiet $d\Gamma(t_0)$ noch relativ eng um einen Zentralpunkt konzentriert, d.h. die verschiedenen Punkte im Bereich haben ziemlich ähnliche Positionen und Impulse. Als die Zeit vergeht verformt sich der Bereich in dessen Bewegung durch den Phasenraum immer mehr, so dass Trajektorien, die am Anfang nah aneinander lagen, sich nach einer Weile wie in $d\Gamma(t_2)$ weit

⁽⁵²⁾... wie z.B., in der echten physikalischen Welt, eine Flüssigkeit oder ein Gas.

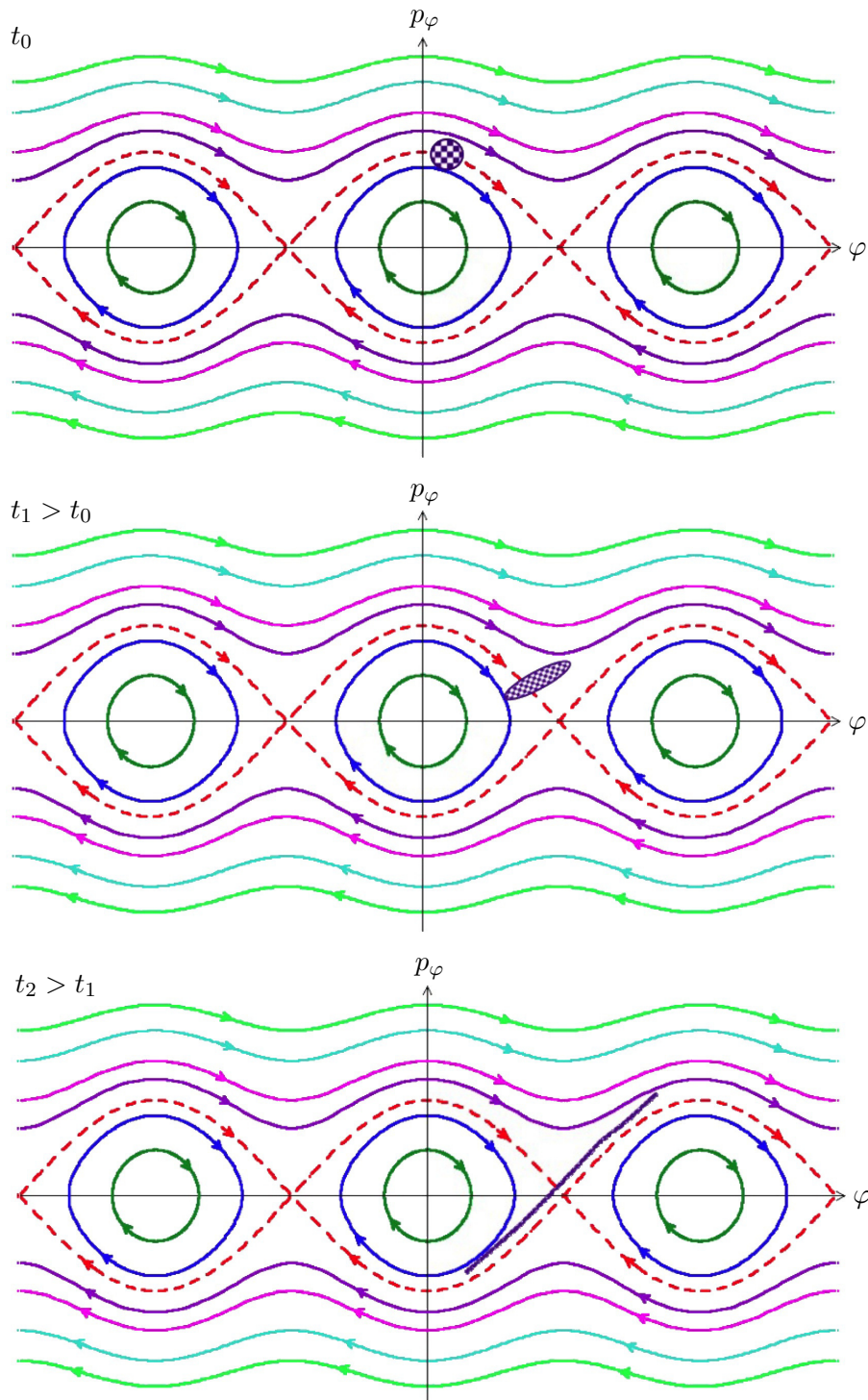


Abbildung V.6 – Entwicklung eines Volumenelements im Phasenraum des ebenen Pendels zu drei sukzessiven Zeiten $t_2 > t_1 > t_0$ (von oben nach unten).

voneinander entfernen. Somit kann eine kleine Unbestimmtheit über den „Anfangs-“Zustand eines Systems nach einiger Zeit groß werden, entsprechend *Chaos* im System. Dabei bleibt die Dynamik aber völlig deterministisch, weil die Bewegungsgleichungen — die Hamiltonschen Gleichungen — selbst völlig deterministisch sind.

Zum Quantifizieren des Verhaltens sollte man einen Abstand d_Γ im Phasenraum einführen. Chaos wird dadurch charakterisiert, dass der Abstand zwischen zwei typischen Trajektorien mit der Zeit

gemäß

$$d_{\Gamma}(t) \sim d_{\Gamma}(t_0) e^{\lambda(t-t_0)} \quad (\text{V.18})$$

wächst, wobei $\lambda > 0$ der *Ljapunow*^(w)-*Exponent* mit physikalischer Dimension Γ^{-1} ist. Dann stellt die *Ljapunow-Zeit* λ^{-1} die typische Zeitskala dar, über die sich sinnvolle Vorhersagen über das System machen lassen.

V.3 Poisson-Mechanik

Die Hamilton-Mechanik kann noch in einer alternativen Form formuliert werden, in welcher der Phasenraum Γ eines Systems und die darauf definierten Funktionen (§ V.3.1) eine zentrale Rolle spielen. Dabei wird eine bilineare Abbildung auf Γ eingeführt, die Poisson^(x)-Klammer, die mit zwei Phasenraumfunktionen eine neue Phasenraumfunktion assoziiert (§ V.3.2). Anhand dieser Abbildung lässt sich die Zeitableitung irgendeiner Funktion auf Γ einfach ausdrücken, und insbesondere die Hamiltonschen Bewegungsgleichungen symmetrisch umschreiben (§ V.3.3). Abschnitt V.3.4 befasst sich mit den Koordinatentransformationen, welche die Form der Bewegungsgleichungen invariant lassen. Schließlich wird der Zusammenhang zwischen den Symmetrien eines physikalischen Systems und den Erhaltungsgrößen im Rahmen des Hamilton-Formalismus wiedergefunden (§ V.3.5).

Es sei hier schon erwähnt, dass das große Interesse der hier eingeführten Formulierung daran liegt, dass sie den „klassischen Limes“ des Hamilton-Formalismus der Quantenmechanik darstellt.

V.3.1 Phasenraum-Funktionen

Die verallgemeinerten Koordinaten $\{q^{\alpha}\}$ und die dazu konjugierten Impulse $\{p_{\alpha}\}$ bestimmen den Bewegungszustand eines mechanischen Systems vollständig. Somit lässt sich jede mögliche Größe, die diesen Zustand charakterisiert — wie z.B. die Position, der Drehimpuls oder die gesamte Energie —, durch die Phasenraumkoordinaten des Systems ausdrücken.

Dementsprechend ist es sinnvoll, Funktionen von der Zeit t und den $2s$ Phasenraumkoordinaten $\{q^{\alpha}\}, \{p_{\alpha}\}$ mit $\alpha = 1, \dots, s$ zu betrachten. Im Rest dieses Kapitels werden solche Funktionen der Kürze halber „Phasenraum-Funktionen“ oder „Funktionen auf dem Phasenraum“ genannt, auch wenn die Zeit auch Argument der Funktion ist. Dazu wird angenommen, dass sie beliebig differenzierbar sind, ohne dass das jedes Mal erwähnt wird.

Bemerkung: Wenn die Funktion einer messbaren physikalischen Größe entspricht, anstatt nur ein mathematisches Konstrukt zu sein, wird sie auch *Observable* genannt.

V.3.2 Poisson-Klammer

V.3.2 a Definition

Definition: Seien f und g zwei Funktionen von der Zeit t und den $2s$ Phasenraumkoordinaten $\{q^{\alpha}\}, \{p_{\alpha}\}$ mit $\alpha = 1, \dots, s$. Ihre *Poisson-Klammer* ist ebenfalls eine Phasenraum-Funktion derselben Variablen, definiert durch

$$\{f, g\} \equiv \sum_{\alpha=1}^s \left(\frac{\partial f}{\partial q^{\alpha}} \frac{\partial g}{\partial p_{\alpha}} - \frac{\partial f}{\partial p_{\alpha}} \frac{\partial g}{\partial q^{\alpha}} \right), \quad (\text{V.19})$$

wobei die $(t, \mathbf{q}, \mathbf{p})$ -Abhängigkeit aller Funktionen nicht geschrieben wurde.

Bemerkungen:

* In der Literatur sind die Poisson-Klammern manchmal mit einem globalen Minus-Vorzeichen vor der rechten Seite definiert. Um nur internationale Standardreferenzen zu nennen ist die hier

^(w)A. M. LJAPUNOW (auch LYAPUNOV), 1857–1918 ^(x)S. POISSON, 1781–1740

verwendete Konvention die gleiche wie bei Arnold [1] oder Goldstein [4, 5], während Landau & Lifschitz [13, 26] die alternative Konvention benutzen.

Auf ähnliche Weise ist die Notation nicht universell: somit benutzen viele Autoren rechteckige Klammern $[\cdot, \cdot]$ — z.B. Arnold, Goldstein oder Landau & Lifschitz — um die formelle Analogie mit dem Kommutator der Quantenmechanik zu betonen.

* Hiernach wird die Poisson-Klammer (V.19) manchmal auch mit $\{f, g\}_{\mathbf{q}, \mathbf{p}}$ bezeichnet, d.h. mit expliziter Angabe der relevanten Phasenraumkoordinaten.

Für die spätere Diskussion über kanonische Transformationen in § V.3.4 ist es günstig, die Poisson-Klammer in einer Matrixform zu schreiben. Dafür führt man die s -dimensionalen Spaltenvektoren

$$\nabla_{\mathbf{q}} f \equiv \begin{pmatrix} \partial f / \partial q_1 \\ \vdots \\ \partial f / \partial q_s \end{pmatrix} \quad \text{und} \quad \nabla_{\mathbf{p}} f \equiv \begin{pmatrix} \partial f / \partial p_1 \\ \vdots \\ \partial f / \partial p_s \end{pmatrix} \quad (\text{V.20})$$

ein. Diese können wiederum in einen Spaltenvektor mit insgesamt $2s$ Komponenten kombiniert werden. Dann ist die Poisson-Klammer von f und g gegeben durch

$$\{f, g\} = ((\nabla_{\mathbf{q}} f)^{\top} \quad (\nabla_{\mathbf{p}} f)^{\top}) \begin{pmatrix} 0 & \mathbb{1}_s \\ -\mathbb{1}_s & 0 \end{pmatrix} \begin{pmatrix} \nabla_{\mathbf{q}} g \\ \nabla_{\mathbf{p}} g \end{pmatrix}, \quad (\text{V.21})$$

wobei $(\nabla_{\mathbf{q}} f)^{\top}$ und $(\nabla_{\mathbf{p}} f)^{\top}$ die zu den Spaltenvektoren (V.20) transponierten Zeilenvektoren sind, während $\mathbb{1}_s$ die $s \times s$ -Einheitsmatrix bezeichnet.

V.3.2b Eigenschaften

Die Poisson-Klammer besitzt einige mathematischen Eigenschaften, die sich generell problemlos beweisen lassen und deshalb hiernach nur aufgelistet werden. Der Kürze halber wird die Abhängigkeit der verschiedenen Funktionen von ihren Variablen $(t, \mathbf{q}, \mathbf{p})$ nicht geschrieben.

Bilinearität Seien f, g_1, g_2 bzw. f_1, f_2, g drei Funktionen auf dem Phasenraum und $\lambda_1, \lambda_2 \in \mathbb{C}$. Dann gelten

$$\{\lambda_1 f_1 + \lambda_2 f_2, g\} = \lambda_1 \{f_1, g\} + \lambda_2 \{f_2, g\}, \quad (\text{V.22a})$$

$$\{f, \lambda_1 g_1 + \lambda_2 g_2\} = \lambda_1 \{f, g_1\} + \lambda_2 \{f, g_2\}. \quad (\text{V.22b})$$

Antisymmetrie / Antikommutativität Für jedes Paar (f, g) von Funktionen auf dem Phasenraum gilt

$$\{f, g\} = -\{g, f\}. \quad (\text{V.23})$$

Daraus folgt trivial $\{f, f\} = 0$.

Nullelemente Sei K eine Zahl; die Funktion auf dem Phasenraum, die identisch konstant gleich K ist, ist ein Nullelement, d.h. ihre Poisson-Klammer mit jeder Funktion f auf dem Phasenraum verschwindet

$$\{f, K\} = 0. \quad (\text{V.24})$$

Produktregel Für jedes Triplett (f, g, h) von Funktionen auf dem Phasenraum gelten

$$\{f, gh\} = \{f, g\}h + g\{f, h\} \quad (\text{V.25})$$

sowie die

Jacobi-Identität

$$\{\{f, g\}, h\} + \{\{g, h\}, f\} + \{\{h, f\}, g\} = 0. \quad (\text{V.26})$$

Im Gegensatz zu den anderen Eigenschaften, die sich in einer Zeile nachprüfen lassen, ist der Beweis dieser Identität mühsam. Jeder der drei Terme ist eine Summe über zwei Freiheitsgrade-Indizes von 8 Beiträgen, die selbst Produkte von zwei Ableitungen und einer doppelten Ableitung sind. Das Spiel besteht darin, Indizes zu umbenennen und die Vertauschung der Ordnung der Ableitungen zu benutzen, um das gesuchte Ergebnis zu finden.

Bemerkungen:

* Für die dritte Eigenschaft ist eigentlich nur die Unabhängigkeit der Funktion K von den Phasenraumkoordinaten nötig: die „Konstante“ kann noch von der Zeit abhängen — entsprechend einer auf dem Phasenraum gleichförmigen Funktion —, ohne den Nullwert deren Poisson-Klammer mit jeder anderen Funktion zu ändern.

* Versetzen mit der Addition und der Poisson-Klammer bildet die Menge der Funktionen auf dem Phasenraum eines Systems eine Algebra.

V.3.2c Fundamentale Poisson-Klammern

Seien $\{q^\alpha\}_{\alpha=1,\dots,s}$ verallgemeinerte Koordinaten und $\{p_\alpha\}_{\alpha=1,\dots,s}$ die zugehörigen konjugierten Impulse. Aus der Definition der Poisson-Klammer zweier Funktionen (V.19) und der Tatsache, dass die Phasenraumkoordinaten unabhängig voneinander sind, folgen die *fundamentalen Poisson-Klammern*

$$\begin{aligned} \{q^\alpha, q^\beta\} &= \{p_\alpha, p_\beta\} = 0 \\ \{q^\alpha, p_\beta\} &= \delta_\beta^\alpha \end{aligned} \quad (\text{V.27})$$

Bemerkung: Da die Poisson-Klammer zweier Funktionen selbst eine Funktion von der Zeit und der Phasenraumkoordinaten ist, bedeutet δ_β^α hier eine Funktion, die für $\alpha \neq \beta$ identisch Null, für $\alpha = \beta$ identisch gleich 1 ist.

Der Beweis der Beziehungen (V.27) ist trivial. Beispielsweise gilt

$$\{q^\alpha, p_\beta\} = \sum_{\gamma=1}^s \left(\frac{\partial q^\alpha}{\partial q^\gamma} \frac{\partial p_\beta}{\partial p_\gamma} - \frac{\partial q^\alpha}{\partial p_\gamma} \frac{\partial p_\beta}{\partial q^\gamma} \right) = \sum_{\gamma=1}^s \delta_\gamma^\alpha \delta_\beta^\gamma,$$

wobei die zweite Gleichung die Unabhängigkeit der Koordinaten ausdrückt und zu $\{q^\alpha, p_\beta\} = \delta_\beta^\alpha$ führt.

Definition: Wenn Phasenraumkoordinaten $\{q^\alpha\}$, $\{p_\alpha\}$ die Gleichungen (V.27) erfüllen, so heißen sie *kanonische Variablen*.

V.3.3 Poisson-Klammer und Zeitentwicklung

Mit Hilfe der Poisson-Klammer lässt sich die Zeitentwicklung einer Phasenraumfunktion elegant umschreiben.

V.3.3a Zeitentwicklung einer Phasenraumfunktion

Sei jetzt f eine beliebige Funktion auf dem Phasenraum eines physikalischen Systems, dessen Hamilton-Funktion \mathcal{H} als bekannt angenommen wird. Die Anwendung der Kettenregel gibt für die totale Ableitung von f nach der Zeit

$$\frac{df(t, \mathbf{q}(t), \mathbf{p}(t))}{dt} = \frac{\partial f(t, \mathbf{q}(t), \mathbf{p}(t))}{\partial t} + \sum_{\alpha=1}^s \frac{\partial f(t, \mathbf{q}(t), \mathbf{p}(t))}{\partial q^\alpha} \dot{q}^\alpha(t) + \sum_{\alpha=1}^s \frac{\partial f(t, \mathbf{q}(t), \mathbf{p}(t))}{\partial p_\alpha} \dot{p}_\alpha(t).$$

Dabei können die Zeitableitungen $\dot{q}^\alpha(t)$, $\dot{p}_\alpha(t)$ der Koordinatenfunktionen mithilfe der Hamiltonschen Bewegungsgleichungen (V.4) umgeschrieben werden:

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \sum_{\alpha=1}^s \left(\frac{\partial f}{\partial q^\alpha} \frac{\partial \mathcal{H}}{\partial p_\alpha} - \frac{\partial f}{\partial p_\alpha} \frac{\partial \mathcal{H}}{\partial q^\alpha} \right),$$

wobei alle Funktionen im gleichen Punkt $(t, \mathbf{q}(t), \mathbf{p}(t))$ auszuwerten sind. Unter Verwendung der Definition (V.19) der Poisson-Klammer lautet diese Zeitableitung

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \{f, \mathcal{H}\}. \quad (\text{V.28})$$

Insbesondere werden die Hamiltonschen Bewegungsgleichungen (V.4) zu

$$\begin{aligned} \frac{dq^\alpha}{dt} &= \{q^\alpha, \mathcal{H}\} \\ \frac{dp_\alpha}{dt} &= \{p_\alpha, \mathcal{H}\} \end{aligned} \quad \text{für } \alpha = 1, \dots, s, \quad (\text{V.29})$$

weil die Projektionen $q^\alpha(t, \mathbf{q}(t), \mathbf{p}(t)) = q^\alpha$, $p_\alpha(t, \mathbf{q}(t), \mathbf{p}(t)) = p_\alpha$ auf die Koordinatenachsen der Position im Phasenraum trotz ihrer Notation keine explizite Funktion der Zeit sind. In dieser Form besitzen die Bewegungsgleichungen für alle Koordinatenfunktionen die gleiche Form.

V.3.3b Integrale der Bewegung

In Übereinstimmung mit Definition (II.9) ist eine Funktion \mathcal{K} der Zeit t und der Phasenraumkoordinaten eine Konstante der Bewegung, auch *Integral der Bewegung* genannt, wenn sie konstant entlang der Trajektorie eines Systems im Phasenraum bleibt, d.h. wenn $d\mathcal{K}/dt = 0$. Laut Gl. (V.29) ist diese Anforderung äquivalent zu

$$\frac{\partial \mathcal{K}}{\partial t} + \{\mathcal{K}, \mathcal{H}\} = 0,$$

d.h. unter Verwendung der Antikommutativität (V.23) der Poisson-Klammer

$$\mathcal{K}(t, \mathbf{q}, \mathbf{p}) \text{ Integral der Bewegung} \Leftrightarrow \frac{\partial \mathcal{K}}{\partial t} = \{\mathcal{H}, \mathcal{K}\}. \quad (\text{V.30})$$

Mithilfe der Poisson-Klammer können auch — zumindest prinzipiell — neue Integrale der Bewegung gefunden werden, und zwar dank dem

Theorem (Satz von Poisson): Seien \mathcal{K}_1 , \mathcal{K}_2 zwei Konstanten der Bewegung. Dann ist ihre Poisson-Klammer $\{\mathcal{K}_1, \mathcal{K}_2\}$ auch eine Erhaltungsgröße.

Beweis: die Jacobi-Identität (V.26) mit $f = \mathcal{K}_1$, $g = \mathcal{K}_2$, $h = \mathcal{H}$ gibt

$$\{\mathcal{H}, \{\mathcal{K}_1, \mathcal{K}_2\}\} = \{\{\mathcal{K}_2, \mathcal{H}\}, \mathcal{K}_1\} + \{\{\mathcal{H}, \mathcal{K}_1\}, \mathcal{K}_2\}$$

d.h. nach Verwendung der Beziehung (V.30) für beide Konstanten der Bewegung

$$\{\mathcal{H}, \{\mathcal{K}_1, \mathcal{K}_2\}\} = \left\{ \mathcal{K}_1, \frac{\partial \mathcal{K}_2}{\partial t} \right\} + \left\{ \frac{\partial \mathcal{K}_1}{\partial t}, \mathcal{K}_2 \right\} = \frac{\partial}{\partial t} \{\mathcal{K}_1, \mathcal{K}_2\},$$

wobei die letzte Gleichung aus dem Austauschen von partiellen Ableitung nach der Zeit und nach Phasenraumkoordinaten folgt. Somit ist $\{\mathcal{K}_1, \mathcal{K}_2\}$ laut Gl. (V.30) erhalten. \square

Bemerkung: In der Praxis ist dieser Satz nicht immer nützlich, denn seine Anwendung führt schnell entweder zur Nullfunktion, oder zu einem Integral der Bewegung, das abhängig von den schon bekannten Integralen ist, und somit nicht „neu“ ist.

Eigentlich kann ein System mit s Freiheitsgraden maximal $2s - 1$ unabhängige Konstanten der Bewegung haben.

V.3.4 Kanonische Transformationen

Es kann günstig sein, Koordinatentransformationen im Phasenraum durchzuführen.

V.3.4 a Koordinatentransformation im Phasenraum

Wir betrachten eine allgemeine Koordinatentransformation im Phasenraum

$$(\{q^\alpha\}, \{p_\alpha\}) \rightarrow (\{Q^\alpha\}, \{P_\alpha\}) \quad \text{mit} \quad \begin{cases} Q^\alpha = Q^\alpha(t, \mathbf{q}, \mathbf{p}) \\ P_\alpha = P_\alpha(t, \mathbf{q}, \mathbf{p}) \end{cases} \quad \text{für } \alpha = 1, \dots, s. \quad (\text{V.31})$$

Unter dieser Transformation der Koordinaten ändert sich im Allgemeinen auch die funktionale Form der Phasenraum-Funktion, welche eine gewisse physikalische Größe ausdrückt. Das heißt, eine Funktion f der alten Koordinaten $\mathbf{q} \equiv \{q^\alpha\}$, $\mathbf{p} \equiv \{p_\alpha\}$ soll durch eine Funktion F der neuen Koordinaten $\mathbf{Q} \equiv \{Q^\alpha\}$, $\mathbf{P} \equiv \{P_\alpha\}$ ersetzt werden. Die beiden Funktionen stehen in Zusammenhang zu einander, weil sie die Beziehung

$$F(t, \mathbf{Q}(t, \mathbf{q}, \mathbf{p}), \mathbf{P}(t, \mathbf{q}, \mathbf{p})) = f(t, \mathbf{q}, \mathbf{p})$$

erfüllen müssen.

Bemerkung: Hier dürfen die Koordinatentransformationen nicht nur von den Koordinaten $\{q^\beta\}$ (und der Zeit) abhängen, wie es bei den Punkttransformationen (III.13) im Lagrange-Formalismus der Fall war, sondern auch von den Impulsen $\{p_\beta\}$. Somit sind weitere Transformationen möglich.

Beispiel: Sei ein System mit einem einzigen Freiheitsgrad, beschrieben durch kanonisch konjugierte Variablen q, p , mit der Hamilton-Funktion

$$h(q, p) = \frac{a}{2} \frac{1}{q^2} + \frac{b}{2} p^2 q^4, \quad (\text{V.32})$$

wobei a und b positiv sind.

Die Hamiltonschen Bewegungsgleichungen lauten

$$\dot{q} = \frac{\partial h}{\partial p} = bpq^4, \quad \dot{p} = -\frac{\partial h}{\partial q} = \frac{a}{q^3} - 2bp^2q^3.$$

Wird die erste Gleichung nach der Zeit abgeleitet bzw. nach p umgestellt, so ergibt sich

$$\ddot{q} = b\dot{p}q^4 + 4bpq^3\dot{q}, \quad \text{bzw.} \quad p = \frac{\dot{q}}{bq^4}.$$

Daraus folgt die Bewegungsgleichung

$$\ddot{q} = abq + 2\frac{\dot{q}^2}{q},$$

deren Lösung auf erster Sicht nicht trivial aussieht.

Führt man aber die Koordinatentransformation

$$Q \equiv pq^2, \quad P \equiv \frac{1}{q} \quad (\text{V.33})$$

durch, so nimmt die Hamilton-Funktion die bekannte Form

$$H(Q, P) = \frac{a}{2} P^2 + \frac{b}{2} Q^2$$

an, entsprechend einem harmonischen Oszillator mit Masse $m = 1/a$ und natürlicher Kreisfrequenz $\omega^2 = ab$.

Was aber noch nicht klar ist, ist ob die Bewegungsgleichungen in den neuen Koordinaten (Q, P) die gleiche Form wie in den alten erhalten, mit Hamilton-Funktion $H(Q, P)$ statt $h(q, p)$.

V.3.4 b Kanonische Transformationen

Sei nun g eine neue Funktion der alten Koordinaten \mathbf{q}, \mathbf{p} und G die damit assoziierte Funktion der neuen Koordinaten \mathbf{Q}, \mathbf{P} :

$$G(t, \mathbf{Q}(t, \mathbf{q}, \mathbf{p}), \mathbf{P}(t, \mathbf{q}, \mathbf{p})) = g(t, \mathbf{q}, \mathbf{p}).$$

Definition: Eine Koordinatentransformation (V.31) im Phasenraum heißt *kanonische Transformation*, wenn sie die Poisson-Klammer zweier beliebigen Funktionen unverändert lässt, d.h.

$$\{f, g\}_{\mathbf{q}, \mathbf{p}} = \{F, G\}_{\mathbf{Q}, \mathbf{P}}. \quad (\text{V.34})$$

Insbesondere sollen die fundamentalen Poisson-Klammern in den neuen Koordinaten

$$\{Q^\alpha, Q^\beta\}_{\mathbf{Q}, \mathbf{P}} = \{P_\alpha, P_\beta\}_{\mathbf{Q}, \mathbf{P}} = 0 \quad , \quad \{Q^\alpha, P_\beta\}_{\mathbf{Q}, \mathbf{P}} = \delta_\beta^\alpha \quad (\text{V.35})$$

sein.

Reziprok kann man zeigen, dass wenn die fundamentalen Poisson-Klammern in den neuen Phasenraumvariablen die kanonische Form (V.35) annehmen, dann ist die Koordinatentransformation kanonisch. d.h. Gl. (V.34) gilt für alle Funktionen f, g und die damit assoziierten F und G .

Bemerkung: An Stelle der mathematisch korrekten Notation F, G für die Funktionen der neuen Variablen \mathbf{Q}, \mathbf{P} benutzen Physiker oft die gleiche Notation f und g , auch wenn die mathematische Form nicht die gleiche wie in den alten Variablen \mathbf{q}, \mathbf{p} ist. Dementsprechend wird

$$\{f, g\}_{\mathbf{q}, \mathbf{p}} = \{f, g\}_{\mathbf{Q}, \mathbf{P}} \quad (\text{V.36})$$

statt Gl. (V.34) geschrieben.

Transformation der Phasenraumkoordinaten für $s = 1$

Seien f, g zwei Funktionen der Phasenraumkoordinaten q, p eines Systems mit einem einzigen Freiheitsgrad. Unter einer Koordinatentransformation $(q, p) \rightarrow (Q, P)$ werden sie durch neue Funktionen F, G ersetzt mit

$$f(t, q, p) = F(t, Q(t, q, p), P(t, q, p)), \quad g(t, q, p) = G(t, Q(t, q, p), P(t, q, p)).$$

Die Ableitung der ersten dieser Gleichungen nach einer der Phasenraumvariablen, zB q , gibt

$$\frac{\partial f}{\partial q} = \frac{dF}{dq}$$

d.h. unter Verwendung der Kettenregel

$$\frac{\partial f}{\partial q} = \frac{\partial F}{\partial Q} \frac{\partial Q}{\partial q} + \frac{\partial F}{\partial P} \frac{\partial P}{\partial q}.$$

Ähnliche Gleichungen gelten für die Ableitung nach p oder für die Ableitungen von g , so dass die Poisson-Klammer von f und g sich als

$$\begin{aligned} \{f, g\}_{q, p} &= \frac{\partial f}{\partial q} \frac{\partial g}{\partial p} - \frac{\partial f}{\partial p} \frac{\partial g}{\partial q} \\ &= \left(\frac{\partial F}{\partial Q} \frac{\partial Q}{\partial q} + \frac{\partial F}{\partial P} \frac{\partial P}{\partial q} \right) \left(\frac{\partial G}{\partial Q} \frac{\partial Q}{\partial p} + \frac{\partial G}{\partial P} \frac{\partial P}{\partial p} \right) - \left(\frac{\partial F}{\partial Q} \frac{\partial Q}{\partial p} + \frac{\partial F}{\partial P} \frac{\partial P}{\partial p} \right) \left(\frac{\partial G}{\partial Q} \frac{\partial Q}{\partial q} + \frac{\partial G}{\partial P} \frac{\partial P}{\partial q} \right) \end{aligned}$$

geschrieben werden kann. Beim Ausmultiplizieren des Terms auf der rechten Seite kommen acht

Beiträge, wovon sich vier aufheben, während die vier übrigen Terme faktorisiert werden können. Insgesamt ergibt sich

$$\{f, g\}_{q,p} = \left(\frac{\partial F}{\partial Q} \frac{\partial G}{\partial P} - \frac{\partial F}{\partial P} \frac{\partial G}{\partial Q} \right) \left(\frac{\partial Q}{\partial q} \frac{\partial P}{\partial p} - \frac{\partial Q}{\partial p} \frac{\partial P}{\partial q} \right),$$

d.h.

$$\{f, g\}_{q,p} = \{F, G\}_{Q,P} \frac{\partial(Q, P)}{\partial(q, p)}, \quad (\text{V.37a})$$

mit

$$\frac{\partial(Q, P)}{\partial(q, p)} \equiv \frac{\partial Q}{\partial q} \frac{\partial P}{\partial p} - \frac{\partial Q}{\partial p} \frac{\partial P}{\partial q} = \det \begin{pmatrix} \frac{\partial Q}{\partial q} & \frac{\partial Q}{\partial p} \\ \frac{\partial P}{\partial q} & \frac{\partial P}{\partial p} \end{pmatrix} \quad (\text{V.37b})$$

Gleichung (V.37a) zeigt, dass die Koordinatentransformation $(q, p) \rightarrow (Q, P)$ kanonisch ist, wenn deren Jacobi-Determinante (V.37b) gleich 1 ist.

Beispiel: Kommt man zurück zum Beispiel des § V.3.4 a, man prüft einfach nach, dass die Transformation (V.33) eigentlich kanonisch ist.

Transformation der Phasenraumkoordinaten für $s > 1$

Wir betrachten nur eine Transformation $(\mathbf{q}, \mathbf{p}) \rightarrow (\mathbf{Q}, \mathbf{P})$ der Phasenraumkoordinaten für ein System mit s Freiheitsgraden, wobei $s > 1$ ist.⁽⁵³⁾ Laut Gl. (V.21) lässt sich die Poisson-Klammer zweier Funktionen der Variablen (\mathbf{q}, \mathbf{p}) als

$$\{f, g\}_{\mathbf{q}, \mathbf{p}} = ((\nabla_{\mathbf{q}} f)^T \quad (\nabla_{\mathbf{p}} f)^T) \begin{pmatrix} 0 & \mathbf{1}_s \\ -\mathbf{1}_s & 0 \end{pmatrix} \begin{pmatrix} \nabla_{\mathbf{q}} g \\ \nabla_{\mathbf{p}} g \end{pmatrix} \equiv ((\nabla_{\mathbf{q}} f)^T \quad (\nabla_{\mathbf{p}} f)^T) \mathbf{J}_{2s} \begin{pmatrix} \nabla_{\mathbf{q}} g \\ \nabla_{\mathbf{p}} g \end{pmatrix} \quad (\text{V.38})$$

schreiben, wobei die zweite Gleichung die (antisymmetrische) $2s \times 2s$ -Matrix \mathbf{J}_{2s} definiert. Wie im Fall des Problems mit $s = 1$ Freiheitsgrad lassen sich die partiellen Ableitungen von f, g nach den Variablen q^α, p_α durch die Ableitungen von F, G nach den zugehörigen Variablen Q^α, P_α ausdrücken; beispielsweise gelten

$$\frac{\partial g}{\partial q^\alpha} = \frac{dG}{dq^\alpha} = \sum_{\beta=1}^s \left(\frac{\partial G}{\partial Q^\beta} \frac{\partial Q^\beta}{\partial q^\alpha} + \frac{\partial G}{\partial P_\beta} \frac{\partial P_\beta}{\partial q^\alpha} \right)$$

und

$$\frac{\partial g}{\partial p_\alpha} = \frac{dG}{dp_\alpha} = \sum_{\beta=1}^s \left(\frac{\partial G}{\partial Q^\beta} \frac{\partial Q^\beta}{\partial p_\alpha} + \frac{\partial G}{\partial P_\beta} \frac{\partial P_\beta}{\partial p_\alpha} \right)$$

sowie ähnliche Gleichungen mit g bzw. G ersetzt durch f bzw. F . Führt man die vier folgenden $s \times s$ -Matrizen definiert durch ihre Elemente⁽⁵⁴⁾

$$(\nabla_{\mathbf{q}} \mathbf{Q})_{\alpha\beta} = \frac{\partial Q^\beta}{\partial q^\alpha}, \quad (\nabla_{\mathbf{q}} \mathbf{P})_{\alpha\beta} = \frac{\partial P_\beta}{\partial q^\alpha}, \quad (\nabla_{\mathbf{p}} \mathbf{Q})_{\alpha\beta} = \frac{\partial Q^\beta}{\partial p_\alpha}, \quad (\nabla_{\mathbf{p}} \mathbf{P})_{\alpha\beta} = \frac{\partial P_\beta}{\partial p_\alpha}, \quad (\text{V.39a})$$

ein, so lassen sich die Beziehungen zwischen den partiellen Ableitungen von g und denen von G in der kürzeren Matrix-Schreibweise

$$\begin{pmatrix} \nabla_{\mathbf{q}} g \\ \nabla_{\mathbf{p}} g \end{pmatrix} = \begin{pmatrix} \nabla_{\mathbf{q}} \mathbf{Q} & \nabla_{\mathbf{q}} \mathbf{P} \\ \nabla_{\mathbf{p}} \mathbf{Q} & \nabla_{\mathbf{p}} \mathbf{P} \end{pmatrix} \begin{pmatrix} \nabla_{\mathbf{Q}} G \\ \nabla_{\mathbf{P}} G \end{pmatrix} \equiv \mathbf{\Lambda} \begin{pmatrix} \nabla_{\mathbf{Q}} G \\ \nabla_{\mathbf{P}} G \end{pmatrix} \quad (\text{V.39b})$$

⁽⁵³⁾Eigentlich gilt die folgende Herleitung auch im Fall $s = 1$.

⁽⁵⁴⁾Hier ist die tief- oder hochgestellte Position der Indizes irrelevant. Wichtiger ist der Unterschied zwischen Zeilen- und Spaltenindizes.

schreiben. Dabei stellt die $2s \times 2s$ -Matrix $\mathbf{\Lambda}$ die Transformationsmatrix für die Koordinatentransformation unter Betrachtung dar.

Die Transposition dieser Beziehung, mit g bzw. G ersetzt durch f bzw. F , lautet

$$((\nabla_{\mathbf{q}}f)^{\top} \quad (\nabla_{\mathbf{p}}f)^{\top}) = ((\nabla_{\mathbf{Q}}F)^{\top} \quad (\nabla_{\mathbf{P}}F)^{\top}) \mathbf{\Lambda}^{\top}.$$

Dies kann in den Ausdruck (V.38) der Poisson-Klammer von f und g eingesetzt werden. Somit ergibt sich

$$\{f, g\}_{\mathbf{q}, \mathbf{p}} = ((\nabla_{\mathbf{Q}}F)^{\top} \quad (\nabla_{\mathbf{P}}F)^{\top}) \mathbf{\Lambda}^{\top} \mathbf{J}_{2s} \mathbf{\Lambda} \begin{pmatrix} \nabla_{\mathbf{Q}}G \\ \nabla_{\mathbf{P}}G \end{pmatrix}.$$

Andererseits gilt in den neuen Phasenraumkoordinaten

$$\{F, G\}_{\mathbf{Q}, \mathbf{P}} = ((\nabla_{\mathbf{Q}}F)^{\top} \quad (\nabla_{\mathbf{P}}F)^{\top}) \mathbf{J}_{2s} \begin{pmatrix} \nabla_{\mathbf{Q}}G \\ \nabla_{\mathbf{P}}G \end{pmatrix}.$$

Vergleicht man beide Ausdrücke, so sieht man, dass die Transformation $(\mathbf{q}, \mathbf{p}) \rightarrow (\mathbf{Q}, \mathbf{P})$ kanonisch ist, wenn die Anforderung

$$\mathbf{\Lambda}^{\top} \mathbf{J}_{2s} \mathbf{\Lambda} = \mathbf{J}_{2s} \quad (\text{V.40})$$

an die Transformationsmatrix $\mathbf{\Lambda}$ erfüllt ist.

Bemerkung: Eine Matrix, welche die Beziehung (V.40) erfüllt, wird *symplektische Matrix* genannt. Man prüft einfach nach, dass wenn $\mathbf{\Lambda}_1, \mathbf{\Lambda}_2$ zwei solche Matrizen sind, dann ist das Produkt $\mathbf{\Lambda}_1 \mathbf{\Lambda}_2$ auch eine symplektische Matrix: anders gesagt ist das Hintereinanderausführen zweier kanonischer Transformationen wieder eine kanonische Transformation.

Aus der Beziehung (V.40) und $\det \mathbf{J}_{2s} = 1 \neq 0$ folgt $(\det \mathbf{\Lambda})^2 = 1$,⁽⁵⁵⁾ so dass jede symplektische Matrix invertierbar ist; eigentlich gilt

$$\mathbf{\Lambda}^{-1} = \mathbf{J}_{2s}^{-1} \mathbf{\Lambda}^{\top} \mathbf{J}_{2s}. \quad (\text{V.41})$$

Da die Einheitsmatrix $\mathbb{1}_{2s}$ offensichtlich symplektisch ist, bilden symplektische $2s \times 2s$ -Matrizen eine Gruppe, die sog. *symplektische Gruppe* $\text{Sp}(2s)$.

V.3.5 Poisson-Klammer und Symmetrien

Sei $G(\mathbf{q}, \mathbf{p})$ eine (mindestens zwei mal kontinuierlich differenzierbare) Funktion auf dem Phasenraum eines Systems, die nicht explizit von der Zeit abhängt. Anhand dieser Funktion lassen sich infinitesimale Koordinatentransformationen der Form

$$\begin{cases} q^{\alpha} \rightarrow Q^{\alpha} = q^{\alpha} + \delta q^{\alpha} \\ p_{\alpha} \rightarrow P_{\alpha} = p_{\alpha} + \delta p_{\alpha} \end{cases} \quad (\text{V.42a})$$

mit

$$\begin{cases} \delta q^{\alpha} \equiv \epsilon \frac{\partial G(\mathbf{q}, \mathbf{p})}{\partial p_{\alpha}} = \epsilon \{q^{\alpha}, G(\mathbf{q}, \mathbf{p})\} \\ \delta p_{\alpha} \equiv -\epsilon \frac{\partial G(\mathbf{q}, \mathbf{p})}{\partial q^{\alpha}} = \epsilon \{p_{\alpha}, G(\mathbf{q}, \mathbf{p})\} \end{cases} \quad (\text{V.42b})$$

generieren, wobei ϵ ein kleiner reeller Parameter ist. Die Funktion G wird oft *Erzeugende* oder *Generator* der Transformation genannt.

Zuerst kann man zeigen, dass eine solche Koordinatentransformation immer kanonisch ist, unabhängig von der Wahl der Funktion G . Die zugehörige Transformationsmatrix (V.39) lautet

⁽⁵⁵⁾Man kann sogar zeigen, dass die Determinante einer symplektischen Matrix $\det \mathbf{\Lambda} = 1$ ist. Ein elementarer Beweis wird in Ref. [27] gegeben.

$$\mathbf{\Lambda} = \begin{pmatrix} \left(\delta_{\alpha}^{\beta} + \epsilon \frac{\partial^2 G}{\partial q^{\alpha} \partial p_{\beta}} \right) & \left(-\epsilon \frac{\partial^2 G}{\partial q^{\alpha} \partial q^{\beta}} \right) \\ \left(\epsilon \frac{\partial^2 G}{\partial p_{\alpha} \partial p_{\beta}} \right) & \left(\delta_{\beta}^{\alpha} - \epsilon \frac{\partial^2 G}{\partial p_{\alpha} \partial q^{\beta}} \right) \end{pmatrix},$$

wobei die Matrixelemente der $s \times s$ -Subblocks angegeben wurden. Dann gilt

$$\mathbf{J}_{2s} \mathbf{\Lambda} = \begin{pmatrix} \left(\epsilon \frac{\partial^2 G}{\partial p_{\alpha} \partial p_{\beta}} \right) & \left(\delta_{\beta}^{\alpha} - \epsilon \frac{\partial^2 G}{\partial p_{\alpha} \partial q^{\beta}} \right) \\ \left(-\delta_{\alpha}^{\beta} - \epsilon \frac{\partial^2 G}{\partial q^{\alpha} \partial p_{\beta}} \right) & \left(\epsilon \frac{\partial^2 G}{\partial q^{\alpha} \partial q^{\beta}} \right) \end{pmatrix}.$$

Nach einer zweiten Matrixmultiplikation findet man

$$\mathbf{\Lambda}^T \mathbf{J}_{2s} \mathbf{\Lambda} = \mathbf{J}_{2s} + \epsilon^2 \mathbf{M},$$

wobei die Matrixelemente von \mathbf{M} von den zweifachen Ableitungen von G abhängen. Zur Ordnung ϵ ist dies genau die charakteristische Beziehung (V.40) einer symplektischen Matrix, d.h. die Koordinatentransformation (V.42) ist kanonisch.

Sei nun $f(t, \mathbf{q}, \mathbf{p})$ eine beliebige Funktion der Phasenraumkoordinaten (und möglicherweise der Zeit). Anhand einer Taylor-Entwicklung kann man die Variation dieser Funktion zwischen dem "transformierten" Phasenraumpunkt $(\mathbf{q} + \delta\mathbf{q}, \mathbf{p} + \delta\mathbf{p})$ und (\mathbf{q}, \mathbf{p}) schreiben:

$$\delta f(t, \mathbf{q}, \mathbf{p}) \equiv f(t, \mathbf{q} + \delta\mathbf{q}, \mathbf{p} + \delta\mathbf{p}) - f(t, \mathbf{q}, \mathbf{p}) \simeq \sum_{\alpha=1}^s \left(\frac{\partial f}{\partial q^{\alpha}} \delta q^{\alpha} + \frac{\partial f}{\partial p_{\alpha}} \delta p_{\alpha} \right).$$

Unter Verwendung der expliziten Form (V.42b) von δq^{α} , δp_{α} kommt

$$\delta f(t, \mathbf{q}, \mathbf{p}) = \epsilon \{ f(t, \mathbf{q}, \mathbf{p}), G(\mathbf{q}, \mathbf{p}) \}_{\mathbf{q}, \mathbf{p}}. \quad (\text{V.43})$$

Das heißt, die Änderung der Funktion f unter der infinitesimalen Transformation (V.42) ist einfach durch die Poisson-Klammer von f und der Erzeugenden G der Transformation gegeben.

Beziehung (V.43) gilt insbesondere, wenn f die Hamilton-Funktion \mathcal{H} des Systems ist. Sei angenommen, dass G eine Konstante der Bewegung ist. Laut Gl. (V.30) ist die Poisson-Klammer $\{ \mathcal{H}, G \}$ gleich der partiellen Ableitung von G nach der Zeit, d.h. hier Null, denn es wurde angenommen, dass G nicht explizit von der Zeit abhängt. Somit gilt für ein Integral der Bewegung

$$\delta \mathcal{H} = \epsilon \{ \mathcal{H}, G \} = 0. \quad (\text{V.44})$$

Anders gesagt ändert sich die Hamilton-Funktion nicht unter der infinitesimalen kanonischen Transformation, die durch eine Konstante der Bewegung generiert wird: eine solche Koordinatentransformation ist eine Symmetrietransformation des Systems, und wir finden den im Abschn. III.3 diskutierten Zusammenhang zwischen Symmetrien und Erhaltungsgrößen wieder.

Beispiel 1: Invarianz unter Raumtranslationen und Impulserhaltung

Sei zuerst angenommen, dass die Erzeugende einer der Impulse ist, $G(\mathbf{q}, \mathbf{p}) = p_{\beta}$. Dann wird die dadurch generierte infinitesimale Koordinatentransformation (V.42) zu

$$q^{\beta} \rightarrow Q^{\beta} = q^{\beta} + \epsilon, \quad q^{\alpha} \rightarrow Q^{\alpha} = q^{\alpha} \text{ für } \alpha \neq \beta, \quad p_{\alpha} \rightarrow P_{\alpha} = p_{\alpha} \text{ für alle } \alpha.$$

Das heißt, die kanonische Transformation ist eine Translation entlang q^{β} . Nun, wenn die Physik des Systems, charakterisiert durch dessen Hamilton-Funktion, nicht von q^{β} abhängt, sollte sich \mathcal{H} unter der Translation nicht ändern: $\delta \mathcal{H} = 0$. Aus Gl. (V.43) mit $f = \mathcal{H}$ und $G = p_{\beta}$ folgt dann $\{ \mathcal{H}, p_{\beta} \} = 0$, was laut Gl. (V.30) bedeutet, dass p_{β} erhalten ist — wie in § III.3.1 b schon gefunden wurde.

Beispiel 2: Invarianz unter Zeittranslationen und Energieerhaltung

Wenn die Erzeugende die Hamilton-Funktion selber ist, $G(\mathbf{q}, \mathbf{p}) = \mathcal{H}(\mathbf{q}, \mathbf{p})$, lautet die zugehörige Koordinatentransformation (V.42)

$$q^\alpha \rightarrow Q^\alpha = q^\alpha + \epsilon \frac{\partial \mathcal{H}}{\partial p_\alpha} = q^\alpha + \epsilon \dot{q}^\alpha \quad , \quad p_\alpha \rightarrow P_\alpha = p_\alpha - \epsilon \frac{\partial \mathcal{H}}{\partial q^\alpha} = p_\alpha + \epsilon \dot{p}_\alpha \quad \text{für alle } \alpha.$$

Wenn die Phasenraumkoordinaten entlang einer Trajektorie betrachtet sind, erkennt man dabei eine infinitesimale Zeittranslation: $Q^\alpha(t) = q^\alpha(t + \epsilon)$, $P_\alpha(t) = p_\alpha(t + \epsilon)$.

Da die Poisson-Klammer $\{\mathcal{H}, \mathcal{H}\}$ Null ist, ändert sich \mathcal{H} laut Gl. (V.43) nicht unter einer Zeittranslation: das heißt, dass die Hamilton-Funktion die Erzeugende der Zeittranslationen ist, und dass sie darunter erhalten bleibt, entsprechend der Erhaltung der Gesamtenergie (§ III.3.1 a).

Literatur zum Kapitel V

- Fließbach, *Mechanik* [2] Teil V, Kap. 19–23 & Teil VI, Kap. 24–26.
- Goldstein, *Klassische Mechanik* [4] = *Classical Mechanics* [5], Kap. 4.1–4.4, 5, & 6.1–6.3.
- Greiner, *Klassische Mechanik II* [7] Kap. III & IV.
- Nolting, *Klassische Mechanik* [15] Kap. 4.
- Scheck, *Mechanik* [18] Kap. 3.

Zweiter Teil

Klassische Elektrodynamik

KAPITEL VI

Einleitung

VI.1 Grundbegriffe und Definitionen 132

VI.2 Einheiten 134

Die Elektrodynamik beruht auf ein paar gekoppelten Gleichungen, die das Wechselspiel von elektrischen und magnetischen Feldern \vec{E} , \vec{B} mit elektrischen Ladungen und Strömen $\rho_{\text{el.}}$, $\vec{j}_{\text{el.}}$ beschreiben. Somit können die *Maxwell-Gleichungen*

$$\vec{\nabla} \cdot \vec{E}(t, \vec{r}) = \frac{1}{\epsilon_0} \rho_{\text{el.}}(t, \vec{r}) \quad (\text{VI.1a})$$

$$\vec{\nabla} \cdot \vec{B}(t, \vec{r}) = 0 \quad (\text{VI.1b})$$

$$\vec{\nabla} \times \vec{E}(t, \vec{r}) + \frac{\partial \vec{B}(t, \vec{r})}{\partial t} = \vec{0} \quad (\text{VI.1c})$$

$$\vec{\nabla} \times \vec{B}(t, \vec{r}) - \epsilon_0 \mu_0 \frac{\partial \vec{E}(t, \vec{r})}{\partial t} = \mu_0 \vec{j}_{\text{el.}}(t, \vec{r}). \quad (\text{VI.1d})$$

als die Gleichungen angesehen werden, welche die „Antwort“ vom elektromagnetischen Feld auf die Wirkung dessen Quellen modellieren.

Jede der Maxwell-Gleichungen trägt einen eigenen Namen, der sich meistens auf die davon abgeleitete — jedoch historisch früher gefundene — integrale Formulierung der Gleichung. Somit ist Gl. (VI.1a) die *Maxwell-Gauß^(y)-Gleichung* — entsprechend dem in § VII.1.2 diskutierten Gaußschen Gesetz. Dann wird Gl. (VI.1b) (relativ selten) *Maxwell-Thomson^(z)-Gleichung* genannt. Gleichung (VI.1c) heißt *Maxwell-Faraday-Gleichung*, denn sie stellt die differentiale Formulierung des Faraday^(aa)-Induktionsgesetzes (§ IX.1.1 a) dar. Schließlich ist Gl. (VI.1d) die *Maxwell-Ampère^(ab)-Gleichung*, die das Ampère-Gesetz (§ VIII.1.3 c) verallgemeinert.

Wiederum erfahren elektrische Ladungen und Ströme in einem elektromagnetischen Feld eine Kraft. Die zugehörige Kraftdichte, d.h. Kraft pro Volumeneinheit, ist die *Lorentz-Kraftdichte*

$$\vec{f}_{\text{L}}(t, \vec{r}) = \rho_{\text{el.}} \vec{E}(t, \vec{r}) + \vec{j}_{\text{el.}}(t, \vec{r}) \times \vec{B}(t, \vec{r}). \quad (\text{VI.2})$$

In diesem Einführungskapitel werden die verschiedenen Größen und Faktoren, die in den Gleichungen (VI.1)–(VI.2) auftreten, erläutert (Abschn. VI.1). Dann befasst sich Kapitel VI.2 mit dem Problem der „Einheiten“ in der Elektrodynamik.

^(y)C. F. GAUSS, 1777–1855

^(z)W. THOMSON, Lord KELVIN, 1824–1907

^(aa)M. FARADAY, 1791–1867

^(ab)A.-M. AMPÈRE, 1775–1836

VI.1 Grundbegriffe und Definitionen

Die Größen $\vec{E}(t, \vec{r})$ und $\vec{B}(t, \vec{r})$ sind mathematische Vektorfelder, die jeweils das elektrische und das magnetische Feld beschreiben. Beide Felder — genau wie jedes Feld, das in der Elektrodynamik eingeführt wird — sind im Prinzip in jedem Punkt \vec{r} des Raums und zu jeder Zeit t definiert, sind also Felder auf der Raumzeit oder auf einem Gebiet davon.

Genauer heißt das Vektorfeld \vec{E} *elektrische Feldstärke*, mit physikalischer Dimension $\text{ML}^{-1}\text{T}^{-3}$ und SI-Einheit $\text{V} \cdot \text{m}^{-1}$, wobei V für das Volt steht. Wiederum wird \vec{B} als *magnetische Flussdichte* oder *magnetische Induktion* bezeichnet; die zugehörige Dimension bzw. SI-Einheit ist $\text{ML}^{-1}\text{T}^{-2}$ bzw. das Tesla T.

Zur Beschreibung der elektrischen und magnetischen Felder in einem Medium werden weitere Vektorfelder eingeführt, und zwar die *elektrische Flussdichte* (oder *dielektrische Verschiebung*) \vec{D} und die *magnetische Feldstärke* \vec{H} , weshalb die genaueren Bezeichnungen manchmal nötig sind. Im Vakuum, wie es im Rahmen dieses Skripts der Fall sein wird, sind die Ausdrücke „elektrisches Feld“ und „magnetisches Feld“ eindeutig.

Die materiellen „Quellen“ des elektromagnetischen Feldes auf der rechten Seite der Maxwell-Gleichungen (VI.1) sind die elektrischen Ladungsdichte und -stromdichte $\rho_{\text{el.}}$ und $\vec{j}_{\text{el.}}$. Die erstere ist natürlich die elektrische Ladung pro Volumeneinheit, so dass deren Integral über ein Gebiet \mathcal{V}

$$Q_{\mathcal{V}}(t) \equiv \int_{\mathcal{V}} \rho_{\text{el.}}(t, \vec{r}) \, d^3\vec{r} \quad (\text{VI.3})$$

die gesamte *elektrische Ladung* im Volumen \mathcal{V} zur Zeit t darstellt.

Wiederum ist die elektrische Stromdichte derart definiert, dass deren Flächenintegral

$$I(t) \equiv \int_{\mathcal{S}} \vec{j}_{\text{el.}}(t, \vec{r}) \cdot d^2\vec{\mathcal{S}} \quad (\text{VI.4})$$

die *elektrische Stromstärke* durch die Fläche \mathcal{S} zur Zeit t ist, wobei die Stromstärke die Gesamtladung ist, die pro Zeiteinheit durch \mathcal{S} strömt. Dementsprechend stellt die elektrische Stromdichte $j_{\text{el.}}$ die Ladung dar, die pro Zeiteinheit durch eine Einheitsfläche fließt.

Eine elektrische Ladung Q bzw. Stromstärke I hat im SI-System die physikalische Dimension $[Q] = \text{IT}$ bzw. $[I] = \text{I}$ — die elektrische Stromstärke ist eine Basisgröße des Internationalen Größensystems —, mit der SI-Einheit Coulomb^(ac) (C) bzw. Ampere (A), wobei $1 \text{ C} = 1 \text{ A} \cdot \text{s}$. Daraus folgt, dass die elektrische Ladungsdichte bzw. Stromdichte die Dimension $[\rho_{\text{el.}}] = \text{ITL}^{-3}$, gemessen in $\text{C} \cdot \text{m}^{-3}$ bzw. $[\vec{j}_{\text{el.}}] = \text{ITL}^{-2}$ (Einheit $\text{A} \cdot \text{m}^{-2}$) hat.

Man sieht also, dass die elektrische Stromdichte die Dimension des Produkts aus Ladungsdichte und Geschwindigkeit hat. Dies lässt sich an einem einfachen Beispiel gut nachvollziehen: betrachte man Punktladungen, die sich alle mit der gleichen Geschwindigkeit \vec{v} bewegen. Diejenigen, die im Zeitintervall $[t, t + \delta t]$ durch eine ebene Fläche \mathcal{S} strömen, sind solche, die sich zur Zeit t innerhalb eines Zylinders mit Basis \mathcal{S} und den Seiten parallel zu \vec{v} befinden (Abb. VI.1). Dabei ist die Höhe des Zylinders $\vec{v} \cdot \vec{e}_n \delta t$, wobei \vec{e}_n den Normaleinheitsvektor zur Fläche \mathcal{S} bezeichnet. Da das Zylindervolumen $\mathcal{S} \vec{v} \cdot \vec{e}_n \delta t$ beträgt, stellen die Punktladungen innerhalb des Zylinders eine Gesamtladung

$$\rho_{\text{el.}}(t, \vec{r}) \mathcal{S} \vec{v} \cdot \vec{e}_n \delta t$$

dar. Laut der Gl. (VI.4) soll diese elektrische Ladung auch gleich

$$I(t) \delta t = \int_{\mathcal{S}} \vec{j}_{\text{el.}}(t, \vec{r}) \cdot \mathcal{S} \vec{e}_n \delta t$$

sein, weil $\mathcal{S} \vec{e}_n$ der mit der Fläche assoziierte Vektor ist. Somit gilt in diesem Fall

$$\vec{j}_{\text{el.}}(t, \vec{r}) = \rho_{\text{el.}}(t, \vec{r}) \vec{v}. \quad (\text{VI.5})$$

^(ac)C. A. COULOMB, 1736–1806

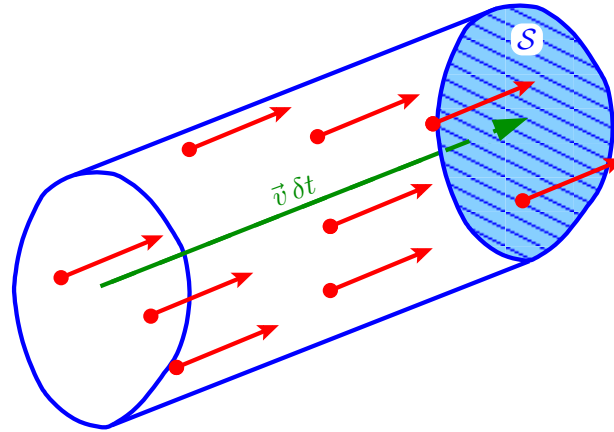


Abbildung VI.1

Das Ergebnis lässt sich direkt verallgemeinern: für eine Menge aus N Punktladungen $\{q_a\}$ mit jeweiligen Ortsvektoren $\{\vec{x}_a(t)\}$ und Geschwindigkeiten $\{\vec{v}_a(t)\}$ zur Zeit t lautet die elektrische Ladungsdichte

$$\rho_{\text{el.}}(t, \vec{r}) = \sum_{a=1}^N q_a \delta^{(3)}(\vec{r} - \vec{x}_a(t)). \quad (\text{VI.6a})$$

Dabei ist $\delta^{(3)}$ die dreidimensionale Dirac-Distribution, die dafür sorgt, dass nur die Raumpunkte beitragen, wo sich eine Punktladung gerade befindet. Integriert über irgendein Volumen ergibt diese Ladungsdichte die Summe der Ladungen, die innerhalb des Volumens sind, wie in Gl. (VI.3). Andererseits ist die elektrische Stromdichte der Punktladungen

$$\vec{j}_{\text{el.}}(t, \vec{r}) = \sum_{a=1}^N q_a \vec{v}_a(t) \delta^{(3)}(\vec{r} - \vec{x}_a(t)). \quad (\text{VI.6b})$$

Falls alle Geschwindigkeiten den gleichen Wert \vec{v} haben, findet man die Beziehung (VI.5) wieder.

Betrachtet man eine einzige Punktladung mit Position $\vec{x}(t)$ und Geschwindigkeit $\vec{v}(t)$ zur Zeit t , so führen die Gl. (VI.6), eingesetzt in die Lorentz-Kraftdichte (VI.2), zu

$$\vec{f}_{\text{L}}(t, \vec{r}) = q \delta^{(3)}(\vec{r} - \vec{x}(t)) [\vec{E}(t, \vec{r}) + \vec{v}(t) \times \vec{B}(t, \vec{r})]. \quad (\text{VI.7})$$

Das Integral dieser Kraftdichte über irgendein Volumen, das die Punktladung enthält, gibt

$$\vec{F}_{\text{L}}(t) = q [\vec{E}(t, \vec{x}(t)) + \vec{v}(t) \times \vec{B}(t, \vec{x}(t))], \quad (\text{VI.8})$$

d.h. die *Lorentz-Kraft* auf die Punktladung.

In den Maxwell-Gleichungen treten (im SI-System, vgl. Abschn. VI.2) zwei Konstanten auf:

- die *elektrische Feldkonstante* ϵ_0 , die früher *Permittivität des Vakuums* genannt wurde;
- die *magnetische Feldkonstante* μ_0 , früher *Permeabilität des Vakuums*.

Im SI-System beträgt die erstere $\epsilon_0 \simeq 8,854 \cdot 10^{-12} \text{ F} \cdot \text{m}^{-1}$ mit F dem Farad ($1 \text{ F} = 1 \text{ C/V}$), und die letztere ist definitionsgemäß $\mu_0 \equiv 4\pi \cdot 10^{-7} \text{ N} \cdot \text{A}^{-2}$. Die jeweiligen physikalischen Dimensionen im internationalen Größensystem sind $[\epsilon_0] = \text{M}^{-1} \text{ L}^{-3} \text{ T}^4 \text{ I}^2$ und $[\mu_0] = \text{L M T}^{-2} \text{ I}^{-2}$.

Beispiele von Ladungs- und Stromverteilungen

Ein erstes Beispiel von elektrischer Ladungs- und Stromdichte ist das von Punktladungen, das oben diskutiert wurde.

Betrachte man eine Metallkugel mit Radius R , auf deren Oberfläche eine Gesamtladung Q homogen verteilt ist. In einem sphärischen Koordinatensystem mit dem Nullpunkt im Zentrum der Kugel soll die zugehörige Ladungsdichte der Form $\rho_{\text{el.}}(\vec{r}) \propto \delta(r - R)$ sein, wobei $r \equiv |\vec{r}|$; sei K der Proportionalitätsfaktor. Laut Beziehung (VI.3) soll

$$\int \rho_{\text{el.}}(\vec{r}) d^3\vec{r} = Q$$

gelten. Das Integral lässt sich als

$$\int \rho_{\text{el.}}(\vec{r}) 4\pi r^2 dr = \int K\delta(r - R) 4\pi r^2 dr = 4\pi R^2 K$$

umschreiben, woraus $K = Q/4\pi R^2$ folgt.

Bemerkung: Zur Beschreibung einer solchen Ladungsverteilung, die auf einer Ober- oder Grenzfläche eingeschränkt ist, wird manchmal eine *Flächenladungsdichte* eingeführt, die die elektrische Ladung pro Einheitsfläche darstellt. Dabei handelt es sich um eine Funktion von nur zwei Ortskoordinaten, welche die Fläche parametrisieren, statt drei: im obigen Beispiel könnte man zwei Winkel (Polar- und Azimutwinkel θ, φ) wählen. Auf ähnlicher Weise definiert man auch *elektrische Flächenstromdichten*.

Als nächstes Beispiel sei ein elektrischer Strom mit Stärke I durch einen unendlich dünnen geradlinigen Draht lokalisiert auf der z -Achse eines Systems kartesischer Koordinaten. Die zugehörige Ladungsstromdichte ist der Form $\vec{j}_{\text{el.}}(\vec{r}) = K\delta(x)\delta(y)\vec{e}_z$ mit einer festzustellenden Konstanten K . Sei \mathcal{S} eine durch den Draht durchlaufene Fläche parallel zur (x, y) -Ebene. Beziehung (VI.4) lautet

$$I = \int_{\mathcal{S}} \vec{j}_{\text{el.}}(\vec{r}) \cdot d^2\vec{S} = \int K\delta(x)\delta(y)\vec{e}_z \cdot d^2\mathcal{S}\vec{e}_z = K,$$

und legt somit $\vec{j}_{\text{el.}}$ fest.

Bemerkung: Hier handelt es sich um eine *Linienstromdichte*, die nur von z abhängt. Ähnlich gibt es *Linienladungsdichten*.

VI.2 Einheiten

In diesem Skript werden die SI-Einheiten und dementsprechend die physikalischen Größen des Internationalen Größensystems verwendet.

Manche Referenzen benutzen aber andere Systeme, insbesondere um die Feldkonstanten ϵ_0 und μ_0 loszuwerden. Dabei handelt es sich⁽⁵⁶⁾ nicht um eine Wahl von Einheiten, wie der Einfachheit halber oft gesagt wird, sondern wirklich um eine Neudefinition der elektromagnetischen Größen (Ladung, elektrisches und magnetisches Feld, usw.). Dementsprechend nehmen die Gleichungen eine andere Form an.

Um die Gleichungen aus dem Internationalen Größensystem in ein anderes System zu übersetzen, braucht man nur einige Konversionsfaktoren zwischen den Größen in beiden Systemen. Hiernach beziehen sich die physikalischen Größen mit einem Sternchen auf ein alternatives System.

Somit schreibt man für die elektrische Ladung

$$q \equiv \sqrt{\epsilon_0\psi} q^*, \quad (\text{VI.9a})$$

was für die Ladungsdichte und -stromdichte zu den Skalierungen

$$\rho_{\text{el.}} = \sqrt{\epsilon_0\psi} \rho_{\text{el.}}^*, \quad \vec{j}_{\text{el.}} = \sqrt{\epsilon_0\psi} \vec{j}_{\text{el.}}^* \quad (\text{VI.9b})$$

führt. Das elektrische Feld wird gewöhnlich derart umdefiniert, dass dessen Produkt mit der Ladung

⁽⁵⁶⁾... trotz des Titels dieses Abschnitts!

die gleiche Form erhält: mit

$$\vec{E} \equiv \frac{1}{\sqrt{\epsilon_0 \psi}} \vec{E}^* \quad (\text{VI.9c})$$

gilt $q\vec{E} = q^*\vec{E}^*$. Wiederum transformiert sich die magnetische Induktion gemäß

$$\vec{B} \equiv \sqrt{\frac{\mu_0}{\psi'}} \vec{B}^*. \quad (\text{VI.9d})$$

In der „Übersetzung“ der Gleichungen sollen schließlich alle auftretenden Produkte $\sqrt{\epsilon_0 \mu_0}$ durch $1/c$ ersetzt werden.

Mit diesen Rezepten wird die Lorentz-Kraftdichte (VI.2) zu

$$\vec{f}_L^* = q^* \left(\vec{E}^* + \sqrt{\frac{\psi}{\psi'}} \frac{1}{c} \vec{v} \times \vec{B}^* \right), \quad (\text{VI.10a})$$

mit dem charakteristischen Faktor $1/c$ vor dem magnetischen Feld. Im neuen Größensystem lauten die Maxwell-Gleichungen nach einiger trivialen Umschreibung

$$\vec{\nabla} \cdot \vec{E}^* = \psi \rho_{\text{el}}^* \quad (\text{VI.10b})$$

$$\vec{\nabla} \cdot \vec{B}^* = 0 \quad (\text{VI.10c})$$

$$\vec{\nabla} \times \vec{E}^* + \sqrt{\frac{\psi}{\psi'}} \frac{1}{c} \frac{\partial \vec{B}^*}{\partial t} = \vec{0} \quad (\text{VI.10d})$$

$$\vec{\nabla} \times \vec{B}^* - \sqrt{\frac{\psi'}{\psi}} \frac{1}{c} \frac{\partial \vec{E}^*}{\partial t} = \frac{\sqrt{\psi \psi'}}{c} \vec{j}_{\text{el}}^*. \quad (\text{VI.10e})$$

Dazu werden natürlich auch weitere Gleichungen geändert, wie z.B. der Ausdruck des Coulomb-Potentials (VII.14b) zwischen zwei statischen Punktladungen q_1, q_2 im Abstand r_{12} voneinander, das zu

$$V = \frac{\psi q_1^* q_2^*}{4\pi r_{12}} \quad (\text{VI.11})$$

wird.

Die Konversionsfaktoren und die Konventionen für ϵ_0, μ_0 und c für die am meisten benutzten Systeme werden in der Tabelle (VI.1) zusammengefasst.

Tabelle VI.1 – Konversionsfaktoren zwischen Größensystemen.

	ϵ_0	μ_0	c	ψ	ψ'
„SI-System“	$\frac{1}{\mu_0 c^2}$	$\mu_0 \equiv 4\pi \cdot 10^{-7} \text{ N A}^{-2}$	$c = 299792458 \text{ m s}^{-1}$	$\mu_0 c^2$	$\frac{\mu_0}{c^2}$
Gauß-System	1	1	c	4π	4π
Heaviside ^(ad) -Lorentz-System	1	1	c	1	1

Bemerkung: In den Gauß- und Heaviside-Lorentz-Systemen haben die Felder \vec{E}^* und \vec{B}^* die gleiche physikalische Dimension, weshalb sie hoch beliebt sind. Dementsprechend nehmen die linken Seiten der Gl. (VI.10d) und (VI.10e) eine ähnliche Form an — bis auf das Minus-Zeichen —, was auch sehr angenehm ist.

Dagegen gilt die Beziehung $\epsilon_0 \mu_0 c^2 = 1$ [vgl. Gl. (IX.29c)] nicht mehr — was der Autor dieser Zeilen schade findet —, es sei denn, man wählt ein System von (sog. „natürlichen“) Einheiten mit $c = 1$.

^(ad)O. HEAVISIDE, 1850–1925

Literatur zum Kapitel VI

- Fließbach, *Elektrodynamik* [2] Teil II.
- Greiner, *Klassische Elektrodynamik II* [8] Kap. I.
- Griffiths, *Elektrodynamik* [9] = *Introduction to Electrodynamics* [10] Kap 2.1–2.4 & 3.
- Jackson, *Klassische Elektrodynamik* [11] = *Classical Electrodynamics* [12] Kap. 1.1–1.5, 1.7–1.11, 2.1, 3.5–3.6 & 4.1–4.2.
- Nolting, *Elektrodynamik* [17] Kap. 2.1–2.3.
- Scheck, *Klassische Feldtheorie* [19] Kap. 1.5 & 1.7.

KAPITEL VII

Elektrostatik

VII.1	Elektrostatistisches Potential	138
VII.1.1	Skalarpotential	138
VII.1.2	Poisson-Gleichung	138
VII.1.3	Elektrisches Feld und Potential von Ladungen	139
VII.1.4	Elektrostatistische potentielle Energie	141
VII.2	Bestimmung des Skalarpotentials aus der Poisson-Gleichung	144
VII.2.1	Greensche Funktionen	144
VII.2.2	Lösung der Poisson-Gleichung auf \mathbb{R}^3	144
VII.2.3	Lösung der Poisson-Gleichung auf einem endlichen Gebiet von \mathbb{R}^3	146
VII.3	Multipolentwicklung	150
VII.3.1	Kartesische Multipolmomente	150
VII.3.2	Beispiele von elektrischen Multipolmomenten	152
VII.3.3	Wechselwirkung zwischen zwei Ladungsverteilungen	154
VII.3.4	Multipolentwicklung in Kugelkoordinaten	157

Die Elektrostatik beschäftigt sich mit den elektrischen Feldern, die von statischen Ladungen bzw. einer stationären Ladungsverteilung $\rho_{\text{el.}}(\vec{r})$ herrühren.⁽⁵⁷⁾ Die angenommene Zeitunabhängigkeit der Konfiguration führt auch dazu, dass das elektrische und das magnetische Feld ebenfalls zeitlich konstant sind, d.h. nur vom Ort abhängen: $\vec{E}(\vec{r})$, $\vec{B}(\vec{r})$. Unter diesen Bedingungen vereinfachen sich die Maxwell–Gauß- und die Maxwell–Faraday-Gleichung (VI.1a), (VI.1c) zu

$$\vec{\nabla} \cdot \vec{E}(\vec{r}) = \frac{\rho_{\text{el.}}(\vec{r})}{\epsilon_0} \quad (\text{VII.1a})$$

$$\vec{\nabla} \times \vec{E}(\vec{r}) = \vec{0}. \quad (\text{VII.1b})$$

Somit hängt die Rotation der elektrischen Feldstärke $\vec{E}(\vec{r})$ nicht mehr vom Magnetfeld ab.

Die zwei Gleichungen (VII.1) sind die Grundgleichungen der Elektrostatik, aus denen sich das elektrische Feld für eine gegebene Ladungsverteilung berechnen lässt. Aus diesen Gleichungen folgt die Existenz eines elektrischen Potentials, aus dem das elektrische Feld abgeleitet werden kann, und das einer relativ einfachen partiellen Differentialgleichung genügt (Abschn. VII.1). Methoden zur Lösung dieser Differentialgleichung, die auch in anderen Gebieten der Physik — insbesondere in der in Kap. VIII behandelten Magnetostatik — auftaucht, werden in Abschn. VII.2 dargelegt. Schließlich befasst sich Abschn. VII.3 mit einer systematischen Entwicklung des elektrostatischen Potentials bzw. Feldes in sukzessiven, immer kleiner werdenden Termen, die in großen Abständen von der Quelle des Feldes gilt.

Wegen der Zeitunabhängigkeit entkoppeln die zwei Gleichungen (VII.1) von den zwei anderen Maxwell-Gleichungen; sie besagen aber nichts über das Magnetfeld.

Dank der angenommenen Abwesenheit von bewegten Ladungen gibt es keinen elektrischen

⁽⁵⁷⁾Natürlich kann die Forderung der Abwesenheit von Bewegung nur in einem Bezugssystem gelten.

Ladungsstrom, $\vec{j}_{\text{el.}} = \vec{0}$. Somit wird die Maxwell–Ampère-Gleichung (VI.1d) zu $\vec{\nabla} \times \vec{B} = \vec{0}$: diese Gleichung und die Maxwell–Thomson-Gleichung (VI.1b), $\vec{\nabla} \cdot \vec{B} = 0$, führen dazu, dass das Magnetfeld \vec{B} nicht nur zeitunabhängig, sondern auch gleichförmig ist. Da dieses Feld keine Kraft auf statische Ladungen übt, wie aus dem Ausdruck der Lorentz-Kraftdichte (VI.2) folgt, kann es einfach vergessen werden.

VII.1 Elektrostatistisches Potential

In diesem Abschnitt wird erstens das elektrostatistische Potential (§ VII.1.1) mit der dadurch erfüllten Differentialgleichung (§ VII.1.2) eingeführt. Dann wird das elektrische Feld für verschiedene einfache Ladungsverteilungen berechnet und eine Formel für das von einer beliebigen Ladungsverteilung herrührende elektrostatistische Potential hergeleitet (§ VII.1.3). Schließlich wird die Coulomb-Kraft zwischen zwei statischen Punktladungen aus den vorigen Ergebnissen wiedergefunden, und die potentielle Energie einer elektrischen Ladungsverteilung diskutiert (§ VII.1.4).

VII.1.1 Skalarpotential

Betrachte ein einfach zusammenhängendes Gebiet⁽⁷⁾ des Raums. Aus der stationären Maxwell–Faraday-Gleichung (VII.1b) folgt die Existenz einer differenzierbaren skalaren Funktion $\Phi(\vec{r})$ derart, dass die Beziehung

$$\vec{E}(\vec{r}) = -\vec{\nabla}\Phi(\vec{r}) \quad (\text{VII.2})$$

in jedem Punkt \vec{r} des Gebiets erfüllt wird. Nach Angabe eines beliebigen Punktes \vec{r}_0 kann man nämlich Φ durch

$$\Phi(\vec{r}) = \Phi(\vec{r}_0) - \int_{\vec{r}_0}^{\vec{r}} \vec{E}(\vec{r}') \cdot d\vec{r}' \quad (\text{VII.3})$$

definieren, wobei das Wegintegral dank dem Satz von Stokes unabhängig vom Weg ist. Somit ist $\Phi(\vec{r})$ eindeutig festgelegt, abgesehen von der beliebigen Wahl des Zahlenwerts für $\Phi(\vec{r}_0)$.

Φ wird als *elektrostatistisches Potential* oder *Skalarpotential* bezeichnet. Die zugehörige SI-Einheit ist das *Volt* (V) mit physikalischer Dimension $[\Phi] = \text{M L}^2 \text{T}^{-3} \text{I}^{-1}$.

Bemerkungen:

* Offensichtlich ist das elektrostatistische Potential Φ nur bis auf eine additive Konstante bestimmt, entsprechend der Wahl von $\Phi(\vec{r}_0)$ in Gl. (VII.3).

* Wie wir im Kap. IX sehen werden, gilt Gl. (VII.2) nicht mehr im nicht-stationären Fall.

VII.1.2 Poisson-Gleichung

Das Ausdrücken des elektrostatistischen Feldes in der Maxwell–Gauß-Gleichung (VII.1a) durch das Skalarpotential gemäß Gl. (VII.2) führt zur *Poisson-Gleichung*

$$\Delta\Phi(\vec{r}) = -\frac{\rho_{\text{el.}}(\vec{r})}{\epsilon_0} \quad (\text{VII.4})$$

mit dem Laplace-Operator Δ .

Insbesondere ergibt sich im Vakuum, d.h. in Abwesenheit von Ladungen, die *Laplace-Gleichung*

$$\Delta\Phi(\vec{r}) = 0, \quad (\text{VII.5})$$

deren Lösungen die sog. *harmonischen Funktionen* sind.

Die Poisson-Gleichung (VII.4) mit $\rho_{\text{el.}} \neq 0$ ist eine inhomogene lineare partielle Differentialgleichung zweiter Ordnung; die Laplace-Gleichung (VII.5) ist die assoziierte homogene Differentialgleichung. Die Bestimmung einer speziellen Lösung der Poisson-Gleichung wird in Abschn. VII.2 diskutiert.

Gaußsches Gesetz

Sei $\partial\mathcal{V}$ eine geschlossene Fläche, die ein zusammenhängendes Raumvolumen \mathcal{V} einschließt. Der elektrische Fluss durch die Oberfläche $\partial\mathcal{V}$ wird als

$$\Phi_E \equiv \oint_{\partial\mathcal{V}} \vec{E}(\vec{r}) \cdot d^2\vec{S} \quad (\text{VII.6})$$

definiert. Dabei ist $d^2\vec{S} = d^2\mathcal{S} \vec{e}_n(\vec{r})$ das vektorielle Flächenelement, wobei $\vec{e}_n(\vec{r})$ den nach außen gerichteten Normaleinheitsvektor zu $d^2\mathcal{S}$ im Punkt \vec{r} bezeichnet.

Unter Verwendung des Integralsatzes von Gauß ist das Oberflächenintegral auf der rechten Seite gleich dem Volumenintegral der Divergenz des Integranden, d.h.

$$\oint_{\partial\mathcal{V}} \vec{E}(\vec{r}) \cdot d^2\vec{S} = \int_{\mathcal{V}} \vec{\nabla} \cdot \vec{E}(\vec{r}) d^3\vec{r}.$$

Unter Betrachtung der Maxwell–Gauß-Gleichung (VII.1a) lässt sich $\vec{\nabla} \cdot \vec{E}(\vec{r})$ ersetzen

$$\Phi_E = \int_{\mathcal{V}} \frac{\rho_{\text{el.}}(\vec{r})}{\epsilon_0} d^3\vec{r}.$$

Daher ergibt sich das *Gaußsche Gesetz*

$$\Phi_E = \frac{Q_{\mathcal{V}}}{\epsilon_0}, \quad (\text{VII.7})$$

wobei $Q_{\mathcal{V}}$ die Gesamtladung im Volumen \mathcal{V} bezeichnet. Dieses Resultat stellt die globale Formulierung der lokalen Maxwell–Gauß-Gleichung (VII.1a) dar.

VII.1.3 Elektrisches Feld und Potential von Ladungen

VII.1.3a Elektrisches Feld und Potential einer Punktladung

Eine Punktladung q befinde sich im Punkt \vec{r}_0 , während im Rest des Raums Vakuum herrscht. Dann ist die Situation kugelsymmetrisch um \vec{r}_0 . Insbesondere sollte das elektrische Feld $\vec{E}(\vec{r})$ in konstantem Abstand $|\vec{r} - \vec{r}_0|$ von der Punktladung einen konstanten Betrag $E(|\vec{r} - \vec{r}_0|)$ haben. Für die Richtung des elektrischen Feldes, die am meistens symmetrische Möglichkeit ist, das es in jedem Punkt entlang der Radialrichtung mit Einheitsvektor $(\vec{r} - \vec{r}_0)/|\vec{r} - \vec{r}_0|$ zeigt. Somit ist ein plausibler Ansatz

$$\vec{E}(\vec{r}) = E(|\vec{r} - \vec{r}_0|) \frac{\vec{r} - \vec{r}_0}{|\vec{r} - \vec{r}_0|}. \quad (\text{VII.8})$$

Betrachte man jetzt eine in \vec{r}_0 zentrierte Kugel \mathcal{K} mit Radius R . In jedem Punkt der Kugeloberfläche $\partial\mathcal{K}$ ist das elektrische Feld normal, d.h. parallel zur Radialrichtung $\vec{r} - \vec{r}_0$, so dass

$$\oint_{\partial\mathcal{K}} \vec{E}(\vec{r}) \cdot d^2\vec{S} = \oint_{\partial\mathcal{K}} E(|\vec{r} - \vec{r}_0|) d^2\mathcal{S}$$

gilt. Dabei ist $\partial\mathcal{K}$ definitionsgemäß die Menge der Punkte mit $|\vec{r} - \vec{r}_0| = R$, d.h. im Integranden kann $E(|\vec{r} - \vec{r}_0|)$ durch $E(R)$ ersetzt werden. Da dieser Term konstant über $\partial\mathcal{V}$ bleibt, kann er aus dem Integral herausgezogen werden:

$$\oint_{\partial\mathcal{K}} \vec{E}(\vec{r}) \cdot d^2\vec{S} = E(R) \oint_{\partial\mathcal{K}} d^2\mathcal{S} = 4\pi R^2 E(R).$$

Dabei wurde benutzt, dass das Oberflächenintegral im mittleren Glied genau den Flächeninhalt der Kugel mit Radius R gibt.

Dank dem Gaußschen Gesetz (VII.7) ist die linke Seite auch mit der Gesamtladung innerhalb des durch $\partial\mathcal{K}$ abgeschlossenen Volumens verknüpft. Hier ist diese Gesamtladung einfach gleich der Ladung q , die im Zentrum der Kugel liegt:

$$\oint_{\partial\mathcal{K}} \vec{E}(\vec{r}) \cdot d^2\vec{S} = \frac{q}{\epsilon_0}.$$

Somit ergibt sich insgesamt

$$\frac{q}{\epsilon_0} = 4\pi R^2 E(R) \quad \text{bzw.} \quad E(R) = \frac{q}{4\pi\epsilon_0 R^2}.$$

Nach Einsetzen dieses Resultats in den Ansatz (VII.8) ergibt sich schließlich für das elektrische Feld einer Punktladung q im Punkt \vec{r}_0

$$\vec{E}(\vec{r}) = \frac{q}{4\pi\epsilon_0} \frac{\vec{r} - \vec{r}_0}{|\vec{r} - \vec{r}_0|^3} \quad \text{für } \vec{r} \neq \vec{r}_0. \quad (\text{VII.9a})$$

Man prüft problemlos nach, dass das zugehörige Skalarpotential

$$\Phi(\vec{r}) = \frac{q}{4\pi\epsilon_0 |\vec{r} - \vec{r}_0|} \quad \text{für } \vec{r} \neq \vec{r}_0 \quad (\text{VII.9b})$$

lautet, wobei hier $\Phi(\vec{r}) = 0$ im Unendlichen gewählt wurde.

VII.1.3 b Elektrisches Feld und Potential mehrerer Punktladungen

Betrachte man jetzt N Punktladungen q_1, \dots, q_N , die sich jeweils in $\vec{r}_1, \dots, \vec{r}_N$ befinden, in Abwesenheit weiterer Ladungen im ganzen Raum.

Zur Bestimmung des entsprechenden elektrischen Feldes bzw. Skalarpotentials soll man nur darauf beachten, dass die Grundgleichungen der Elektrostatik — sowohl die stationären Maxwell-Gleichungen (VII.1) als äquivalent die Beziehung (VII.2) und die Poisson-Gleichung (VII.4) — linear sind. Demzufolge gilt ein Superpositionsprinzip für die Lösungen der Gleichungen: wenn $(\rho_{\text{el.,1}}(\vec{r}), \Phi_1(\vec{r}))$ und $(\rho_{\text{el.,2}}(\vec{r}), \Phi_2(\vec{r}))$ die Poisson-Gleichung (VII.4) auf \mathbb{R}^3 erfüllen, dann genügen die Funktionen $(\rho_{\text{el.,1}}(\vec{r}) + \rho_{\text{el.,2}}(\vec{r}), \Phi_1(\vec{r}) + \Phi_2(\vec{r}))$ ebenfalls der Poisson-Gleichung auf \mathbb{R}^3 .

Dementsprechend lautet das Skalarpotential für die N Punktladungen

$$\Phi(\vec{r}) = \sum_{a=1}^N \frac{q_a}{4\pi\epsilon_0 |\vec{r} - \vec{r}_a|} \quad \text{für } \vec{r} \notin \{\vec{r}_1, \dots, \vec{r}_N\}, \quad (\text{VII.10a})$$

wieder mit der Wahl $\Phi(\vec{r}) = 0$ für $|\vec{r}| \rightarrow \infty$. Das zugehörige elektrische Feld ist

$$\vec{E}(\vec{r}) = \sum_{a=1}^N \frac{q_a}{4\pi\epsilon_0} \frac{\vec{r} - \vec{r}_a}{|\vec{r} - \vec{r}_a|^3} \quad \text{für } \vec{r} \notin \{\vec{r}_1, \dots, \vec{r}_N\}, \quad (\text{VII.10b})$$

wie sich entweder aus der Gradientenbildung des Skalarpotentials (VII.10a) oder aus der Superposition der elektrischen Felder (VII.9b) der individuellen Punktladungen folgern lässt.

VII.1.3 c Elektrisches Feld und Potential einer Ladungsverteilung

Ähnlich dem Übergang (IV.17) von einer diskreten Massenverteilung $\{m_a\}$, entsprechend einer endlichen Anzahl von Massenpunkten, zu einer durch eine Massendichte $\rho(\vec{r})$ beschriebenen kon-

tinuierlichen Verteilung, kann man von einer endlichen Anzahl von Punktladungen $\{q_a\}$ zu einer kontinuierlichen Ladungsverteilung $\rho_{\text{el.}}$ übergehen, indem man das Rezept

$$\sum_a q_a f(\vec{r}_a) \rightarrow \int_{\mathcal{V}} f(\vec{r}) \rho_{\text{el.}}(\vec{r}) d^3\vec{r} \quad (\text{VII.11})$$

für jede Funktion f der Positionen der Punktladungen anwendet, wobei \mathcal{V} das durch die Ladungsverteilung besetzte Raumvolumen bezeichnet.

Demzufolge lautet das elektrostatische Potential bzw. Feld für eine beliebige elektrische Ladungsverteilung $\rho_{\text{el.}}(\vec{r})$

$$\Phi(\vec{r}) = \int_{\mathcal{V}} \frac{\rho_{\text{el.}}(\vec{r}')}{4\pi\epsilon_0 |\vec{r} - \vec{r}'|} d^3\vec{r}', \quad (\text{VII.12a})$$

bzw.

$$\vec{E}(\vec{r}) = \int_{\mathcal{V}} \frac{\rho_{\text{el.}}(\vec{r}')}{4\pi\epsilon_0} \frac{\vec{r} - \vec{r}'}{|\vec{r} - \vec{r}'|^3} d^3\vec{r}'. \quad (\text{VII.12b})$$

Bemerkungen:

* Der Vorsicht halber sollte der Punkt \vec{r} außerhalb des Volumens \mathcal{V} sein, um unangenehme Divisionen durch 0 zu vermeiden.

* Hiernach wird das Potential (VII.12a) auf eine alternative Weise hergeleitet (§ VII.2.2).

* Setzt man in die Gl. (VII.12a)–(VII.12b) die Ladungsverteilung (VI.6a) für eine Menge von Punktladungen ein, so findet man natürlich Gl. (VII.10a) und (VII.10b) wieder.

VII.1.4 Elektrostatische potentielle Energie

VII.1.4 a Coulomb-Kraft

Betrachte wieder N Punktladungen q_1, \dots, q_N , die sich jeweils in $\vec{r}_1, \dots, \vec{r}_N$ befinden. Da die Punktladung q_a ruht, lautet die darauf wirkende Lorentz-Kraft

$$\vec{F}_a(\vec{r}_a) = q_a \vec{E}_a(\vec{r}_a) \quad (\text{VII.13a})$$

wobei $\vec{E}_a(\vec{r}_a)$ das elektrische Feld bezeichnet, das die anderen Ladungen q_b mit $b \neq a$ im Punkt \vec{r}_a erzeugen. Unter Verwendung der Gl. (VII.10b) für dieses Feld folgt

$$\vec{F}_a(\vec{r}_a) = q_a \sum_{\substack{b=1 \\ b \neq a}}^N \frac{q_b}{4\pi\epsilon_0} \frac{\vec{r}_a - \vec{r}_b}{|\vec{r}_a - \vec{r}_b|^3}. \quad (\text{VII.13b})$$

Falls es nur zwei Punktladungen q_1, q_2 in den Punkten \vec{r}_1, \vec{r}_2 gibt, lautet die Kraft, welche q_2 auf q_1 ausübt

$$\vec{F}_{2 \rightarrow 1} = \frac{q_1 q_2}{4\pi\epsilon_0} \frac{\vec{r}_1 - \vec{r}_2}{|\vec{r}_1 - \vec{r}_2|^3}. \quad (\text{VII.14a})$$

Dies ist die bekannte *Coulomb-Kraft*, die sich aus der *potentiellen Energie*

$$V(\vec{r}_1, \vec{r}_2) = \frac{q_1 q_2}{4\pi\epsilon_0 |\vec{r}_1 - \vec{r}_2|}, \quad (\text{VII.14b})$$

gemäß $\vec{F}_{2 \rightarrow 1} = -\vec{\nabla}_1 V(\vec{r}_1, \vec{r}_2)$ ableiten lässt, wobei $\vec{\nabla}_1$ den Gradienten bezüglich des Ortsvektors \vec{r}_1 bezeichnet.

Bemerkungen:

* Der Ausdruck der Coulomb-Kraft (VII.14a) einer elektrischen Punktladung auf eine andere ist antisymmetrisch unter dem Austausch der zwei Ladungen: $\vec{F}_{2 \rightarrow 1} = -\vec{F}_{1 \rightarrow 2}$, in Übereinstimmung mit dem dritten newtonschen Gesetz (I.19).

* Wenn die Punktladungen sich relativ zu einander bewegen, ist die Kraft zwischen ihnen nicht mehr die Coulomb-Kraft (VII.14a).

VII.1.4 b Potentielle Energie

Sei $\vec{\nabla}_a$ der Gradient bezüglich des Ortsvektors \vec{r}_a der a -ten Punktladung eines Systems von N Ladungen. Wie im vorigen Paragraphen bezeichnen \vec{E}_a bzw. \vec{F}_a das elektrische Feld bzw. die Kraft der anderen Ladungen q_b mit $b \neq a$. Aus der Maxwell-Thomson-Gleichung $\vec{\nabla}_a \times \vec{E}_a = \vec{0}$ und der Gl. (VII.13a) folgt $\vec{\nabla}_a \times \vec{F}_a = \vec{0}$: laut der Behauptung vom § I.1.3 c ist die Kraft \vec{F}_a konservativ. Genauer gilt

$$\vec{F}_a(\vec{r}_a) = -\vec{\nabla}_a V(\{\vec{r}_b\}) \quad (\text{VII.15a})$$

mit einer potentiellen Energie V , die durch⁽⁵⁸⁾

$$V(\{\vec{r}_b\}) = \frac{1}{2} \sum_{\substack{c,d=1 \\ c \neq d}}^N \frac{q_c q_d}{4\pi\epsilon_0 |\vec{r}_c - \vec{r}_d|} \quad (\text{VII.15b})$$

gegeben ist. Alternativ gilt

$$V(\{\vec{r}_b\}) = \frac{1}{2} \sum_{c=1}^N q_c \Phi_c(\vec{r}_c), \quad (\text{VII.15c})$$

wobei Φ_c durch Gl. (VII.10a) mit $\vec{r} = \vec{r}_c$ und Summierung über die Punktladungen q_d mit $d \neq c$ gegeben wird.

Beweis: Mit dem Potential (VII.15b) gilt unter Verwendung der Kettenregel

$$\begin{aligned} -\vec{\nabla}_a V &= -\frac{1}{2} \sum_{\substack{c,d=1 \\ c \neq d}}^N \frac{q_c q_d}{4\pi\epsilon_0} \vec{\nabla}_a \left(\frac{1}{|\vec{r}_c - \vec{r}_d|} \right) = -\frac{1}{2} \sum_{\substack{c,d=1 \\ c \neq d}}^N \frac{q_c q_d}{4\pi\epsilon_0} \left(-\delta_{ca} \frac{\vec{r}_c - \vec{r}_d}{|\vec{r}_c - \vec{r}_d|^3} + \delta_{da} \frac{\vec{r}_c - \vec{r}_d}{|\vec{r}_c - \vec{r}_d|^3} \right) \\ &= \frac{1}{2} \sum_{d \neq a} \frac{q_a q_d}{4\pi\epsilon_0} \frac{\vec{r}_a - \vec{r}_d}{|\vec{r}_a - \vec{r}_d|^3} - \frac{1}{2} \sum_{c \neq a} \frac{q_a q_c}{4\pi\epsilon_0} \frac{\vec{r}_c - \vec{r}_a}{|\vec{r}_c - \vec{r}_a|^3}. \end{aligned}$$

Ersetzt man d bzw. c durch b in der ersten bzw. zweiten Summe der zweiten Zeile, so sind beide Summanden gleich und man findet genau das gesuchte Ergebnis (VII.13b). \square

Im Limes einer kontinuierlichen Ladungsverteilung $\rho_{\text{el.}}$ wird die potentielle Energie (VII.15c) zu

$$V = \frac{1}{2} \int_{\mathcal{V}} \rho_{\text{el.}}(\vec{r}) \Phi(\vec{r}) d^3\vec{r} \quad (\text{VII.16})$$

mit dem Raumvolumen \mathcal{V} , das die Ladungsverteilung besetzt; eigentlich kann \mathcal{V} durch \mathbb{R}^3 ersetzt werden

Bemerkung: Die letztere Gleichung könnte auf erster Sicht verwirrend sein, denn das elektrostatische Potential Φ im Integral wird in einem Punkt betrachtet, wo sich Ladungen befinden. Das heißt, dass

⁽⁵⁸⁾Dieser Ausdruck stimmt natürlich mit der durch Gl. II.18 — hier ohne äußeres Potential — definierten gesamten potentiellen Energie eines Mehrteilchensystems überein.

dies nicht dem Gültigkeitsbereich der Gl. (VII.12a) entspricht, für die der Punkt \vec{r} außerhalb der Ladungsverteilung sein soll. Dies bedeutet nur, dass $\Phi(\vec{r})$ nicht durch Gl. (VII.12a) gegeben ist.

VII.1.4c Energie des elektrostatischen Feldes

Die potentielle Energie (VII.16) einer Ladungsverteilung $\rho_{\text{el.}}$ lässt sich auch durch das elektrostatische Feld $\vec{E}(\vec{r})$, das die Verteilung erzeugt, ausdrücken.

In der Tat gilt

$$V = \int_{\mathbb{R}^3} \frac{\epsilon_0 \vec{E}(\vec{r})^2}{2} d^3\vec{r}, \quad (\text{VII.17a})$$

wobei das Integrand als die *elektrische Energiedichte* $e_{\text{el.}}(\vec{r})$ des Feldes im Punkt \vec{r} interpretiert werden kann:

$$e_{\text{el.}}(\vec{r}) \equiv \frac{\epsilon_0 \vec{E}(\vec{r})^2}{2}. \quad (\text{VII.17b})$$

Zum Beweis dieses Ergebnisses kann man das Integral auf der rechten Seite der Gl. (VII.16) zuerst über das Volumen einer Kugel mit Radius R berechnen, und am Ende den Grenzwert $R \rightarrow \infty$ betrachten. Dank der Maxwell–Gauß–Gleichung (VII.1a) gilt

$$\frac{1}{2} \int_{|\vec{r}| \leq R} \rho_{\text{el.}}(\vec{r}) \Phi(\vec{r}) d^3\vec{r} = \frac{\epsilon_0}{2} \int_{|\vec{r}| \leq R} [\vec{\nabla} \cdot \vec{E}(\vec{r})] \Phi(\vec{r}) d^3\vec{r}.$$

Dabei lässt sich das Integrand mithilfe der Identität

$$[\vec{\nabla} \cdot \vec{E}(\vec{r})] \Phi(\vec{r}) = \vec{\nabla} \cdot [\Phi(\vec{r}) \vec{E}(\vec{r})] - \vec{E}(\vec{r}) \cdot \vec{\nabla} \Phi(\vec{r}) = \vec{\nabla} \cdot [\Phi(\vec{r}) \vec{E}(\vec{r})] + \vec{E}(\vec{r})^2$$

umschreiben:

$$\frac{1}{2} \int_{|\vec{r}| \leq R} \rho_{\text{el.}}(\vec{r}) \Phi(\vec{r}) d^3\vec{r} = \frac{\epsilon_0}{2} \int_{|\vec{r}| \leq R} \vec{\nabla} \cdot [\Phi(\vec{r}) \vec{E}(\vec{r})] d^3\vec{r} + \int_{|\vec{r}| \leq R} \frac{\epsilon_0 \vec{E}(\vec{r})^2}{2} d^3\vec{r}.$$

Der zweite Term auf der rechten Seite gibt das gesuchte Ergebnis im Limes $R \rightarrow \infty$, so dass man nun beweisen soll, dass der erste Term in diesem Limes verschwindet. Tatsächlich kann das erste Volumenintegral auf der rechten Seite mit dem Gaußschen Integralsatz in ein Oberflächenintegral über die Kugeloberfläche transformiert werden:

$$\int_{|\vec{r}| \leq R} \vec{\nabla} \cdot [\Phi(\vec{r}) \vec{E}(\vec{r})] d^3\vec{r} = \oint_{|\vec{r}|=R} \Phi(\vec{r}) \vec{E}(\vec{r}) \cdot d^2\vec{S}.$$

Wenn R groß genug ist, befinden sich alle Ladungen innerhalb der Kugel $|\vec{r}| \leq R$. Dann nehmen das elektrostatische Potential und das Feld mit R ab, gemäß⁽⁵⁹⁾

$$\Phi(R) \propto \frac{1}{R}, \quad |\vec{E}(R)| \propto \frac{1}{R^2}.$$

Dagegen gilt $d^2\mathcal{S} = R^2 d^2\Omega$ mit dem Raumwinkelelement $d^2\Omega$. Insgesamt kommt

$$\oint_{|\vec{r}|=R} \Phi(\vec{r}) \vec{E}(\vec{r}) \cdot d^2\vec{S} \propto 4\pi R^2 \frac{1}{R} \frac{1}{R^2} \propto \frac{1}{R},$$

so dass dieser Term im Grenzwert $R \rightarrow \infty$ Null wird. \square

⁽⁵⁹⁾Dies folgt aus dem Gaußschen Gesetz (VII.7), unter Verwendung des gleichen Arguments wie in § VII.1.3 a.

VII.2 Bestimmung des Skalarpotentials aus der Poisson-Gleichung

Im § VII.1.3 wurde das durch eine Ladungsverteilung erzeugte elektrostatische Potential (VII.12a) mithilfe des Gaußschen Gesetzes hergeleitet. Eine alternative Vorgehensweise zur Bestimmung von $\Phi(\vec{r})$ besteht darin, die Poisson-Gleichung

$$\Delta\Phi(\vec{r}) = -\frac{\rho_{\text{el.}}(\vec{r})}{\epsilon_0} \quad (\text{VII.4})$$

direkt zu lösen. In diesem Abschnitt werden einige allgemeinen⁽⁶⁰⁾ mathematischen Rezepte und Ergebnisse zur Lösung dieser Differentialgleichung vorgestellt.

VII.2.1 Greensche Funktionen

Die Poisson-Gleichung (VII.4) ist ein Beispiel von (inhomogener) linearer partieller Differentialgleichung, hier zweiter Ordnung. Zur Lösung solcher Differentialgleichungen kann man sich die Linearität zu Nutze machen und ein Superpositionsprinzip verwenden. Somit wird die Differentialgleichung erstens mit einer Dirac-Distribution im rechten Glied gelöst. Da jede beliebige Ladungsdichte $\rho_{\text{el.}}$ als Superposition von solchen Distributionen

$$\rho_{\text{el.}}(\vec{r}) = \int \delta^{(3)}(\vec{r} - \vec{r}') \rho_{\text{el.}}(\vec{r}') d^3\vec{r}' \quad (\text{VII.18})$$

geschrieben werden kann, ergibt sich eine Lösung der ursprünglichen linearen Differentialgleichung.

Definition: Eine Funktion $G(\vec{r}, \vec{r}')$ zweier vektoriellen Variablen \vec{r}, \vec{r}' heißt *Greensche^(ae) Funktion* zur Poisson-Gleichung, wenn sie eine Lösung der partiellen Differentialgleichung

$$\Delta_{\vec{r}}G(\vec{r}, \vec{r}') = \delta^{(3)}(\vec{r} - \vec{r}') \quad (\text{VII.19})$$

ist, wobei $\Delta_{\vec{r}}$ den Laplace-Operator bezüglich der Variablen \vec{r} bezeichnet.

Aus einer solchen Lösung der Gl. (VII.19) folgt eine Lösung der Poisson-Gleichung (VII.4) mit fast jedem beliebigen rechten Glied, und zwar

$$\Phi(\vec{r}) = \frac{-1}{\epsilon_0} \int G(\vec{r}, \vec{r}') \rho_{\text{el.}}(\vec{r}') d^3\vec{r}', \quad (\text{VII.20})$$

unter der offensichtlichen Beschränkung, dass das Integral existieren soll.

Wendet man nämlich den Laplace-Operator $\Delta_{\vec{r}}$ auf das so definierte Potential an, so kann er wegen dessen Linearität in das Integral über \vec{r}' eingezogen werden:

$$\Delta_{\vec{r}}\Phi(\vec{r}) = \frac{-1}{\epsilon_0} \int [\Delta_{\vec{r}}G(\vec{r}, \vec{r}')] \rho_{\text{el.}}(\vec{r}') d^3\vec{r}',$$

Unter Verwendung der definierenden Gleichung (VII.19) wird das Integral zur Faltung von der $\delta^{(3)}$ -Distribution mit der Ladungsdichte $\rho_{\text{el.}}$, Gl. (VII.18), so dass sich die rechte Seite zu $-\rho_{\text{el.}}(\vec{r})/\epsilon_0$ vereinfacht. \square

Bemerkung: Im Beweis wurde die spezifische Form des Laplace-Operators nirgendwo benutzt, nur seine Linearität. Somit können allgemeiner Greensche Funktionen auch für andere lineare Differentialoperatoren, d.h. andere lineare Differentialgleichungen — egal, ob sie partiell oder gewöhnlich sind — eingeführt werden. Eine solche Greensche Funktion stellt die „Antwort“ zu einer durch eine δ -Distribution modellierten lokalisierten „Anregung“ dar.

⁽⁶⁰⁾Das heißt, dass sie auch auf andere Differentialgleichungen anwendbar sind.

^(ae)G. GREEN, 1793–1841

VII.2.2 Lösung der Poisson-Gleichung auf \mathbb{R}^3

Die Poisson- (VII.4) und Laplace-Gleichung (VII.5) wurden durch Mathematiker aufwändig untersucht, woraus eine Menge Resultate zur Existenz und Eindeutigkeit ihrer Lösungen resultiert. In diesem und im nächsten Paragraphen werden einige dieser Ergebnisse kurz dargestellt.

Betrachte man zuerst die Poisson-Gleichung, und daher die assoziierte partielle Differentialgleichung (VII.19), auf dem ganzen dreidimensionalen Raum \mathbb{R}^3 . Dann lautet eine Greensche Funktion zur Poisson-Gleichung

$$G(\vec{r}, \vec{r}') = \frac{-1}{4\pi|\vec{r} - \vec{r}'|}. \quad (\text{VII.21})$$

Beweis: Sei $\vec{R} \equiv \vec{r} - \vec{r}'$ und $R \equiv |\vec{R}|$; dank der Kettenregel gilt $\Delta_{\vec{r}} = \Delta_{\vec{R}}$. Da die Funktion (VII.21) einen Pol für $\vec{r} = \vec{r}'$ bzw. $\vec{R} = \vec{0}$ hat, wird zuerst eine „regularisierte“ Version

$$G_\varepsilon(\vec{R}) \equiv \frac{-1}{4\pi\sqrt{R^2 + \varepsilon^2}}$$

ohne Divergenz im Ursprungspunkt eingeführt. Dabei ist $\varepsilon > 0$ ein reeller Parameter: am Ende der Berechnung wird der Limes $\varepsilon \rightarrow 0$ genommen. Die Funktion G_ε hängt nur vom Betrag R ab, so dass sie im Raum der Variablen \vec{R} kugelsymmetrisch um den Nullpunkt ist. Demzufolge ist es interessant, mit Kugelkoordinaten $(R, \theta_{\vec{R}}, \varphi_{\vec{R}})$ zu arbeiten: dann ist das Problem unabhängig von den zwei Winkeln $\theta_{\vec{R}}, \varphi_{\vec{R}}$. Daher vereinfacht sich der Laplace-Operator $\Delta_{\vec{R}}$ zu

$$\frac{\partial^2}{\partial R^2} + \frac{2}{R} \frac{\partial}{\partial R}$$

ohne winkelabhängigen Anteil. Wendet man diesen Operator auf G_ε an, so ergibt sich nach einfachen Ableitungen

$$\Delta_{\vec{R}} G_\varepsilon(\vec{R}) = \frac{3\varepsilon^2}{4\pi} \frac{1}{(R^2 + \varepsilon^2)^{5/2}}.$$

Die Funktion im rechten Glied sei mit $\delta_\varepsilon^{(3)}(\vec{R})$ bezeichnet. Für festes ε besitzt $\delta_\varepsilon^{(3)}$ die folgenden Eigenschaften:

- $|\delta_\varepsilon^{(3)}(\vec{R})| \leq \frac{3\varepsilon^2}{4\pi R^5}$ für $\vec{R} \neq \vec{0}$;
- $|\delta_\varepsilon^{(3)}(\vec{0})| = \frac{3}{4\pi|\varepsilon|}$;
- sei $a > 0$; dann gilt $\int_{|\vec{R}| \leq a} \delta_\varepsilon^{(3)}(\vec{R}) d^3\vec{R} = \int_0^a \delta_\varepsilon^{(3)}(R) 4\pi R^2 dR = \int_0^a \frac{3R^2 \varepsilon^2}{(R^2 + \varepsilon^2)^{5/2}} dR$, d.h.

$$\int_{|\vec{R}| \leq a} \delta_\varepsilon^{(3)}(\vec{R}) d^3\vec{R} = \left[\frac{R^3}{(R^2 + \varepsilon^2)^{3/2}} \right]_0^a = \frac{a^3}{(a^2 + \varepsilon^2)^{3/2}}.$$

Betrachte man jetzt den Grenzwert $\varepsilon \rightarrow 0$. Laut der ersten Eigenschaft gilt in diesem Limes $\lim_{\varepsilon \rightarrow 0} \delta_\varepsilon^{(3)}(\vec{R}) = 0$ für $\vec{R} \neq \vec{0}$. Somit wird der Peak bei $\vec{R} = \vec{0}$ (zweite Eigenschaft) unendlich schmal und hoch. Schließlich geht das Integral von $\delta_\varepsilon^{(3)}$ über eine Kugel mit Radius a gegen 1, unabhängig von der Wahl von a . Diese Eigenschaften zeigen, dass $\delta_\varepsilon^{(3)}$ eine Darstellung der dreidimensionalen δ -Distribution ist, entsprechend dem nachzuprüfenden Ergebnis.⁽⁶¹⁾ \square

Setzt man jetzt die Greensche Funktion (VII.21) in die Beziehung (VII.20) ein, so ergibt sich genau das elektrostatische Skalarpotential (VII.12a)

$$\Phi(\vec{r}) = \int_{\mathbb{R}^3} \frac{\rho_{\text{el}}(\vec{r}')}{4\pi\epsilon_0|\vec{r} - \vec{r}'|} d^3\vec{r}'. \quad (\text{VII.22})$$

Dabei wird die vorsichtige Beschränkung des § VII.1.3c auf Punkte \vec{r} , die nicht zum durch die Ladungsverteilung besetzten Raumvolumen \mathcal{V} gehören, jetzt automatisch aufgehoben.

⁽⁶¹⁾Die Funktion (VII.21) kann auch direkt gefunden werden, indem die definierende Differentialgleichung (VII.19) im Fourier-Raum aufgestellt wird.

Ein Problem bei der Konstruktion des elektrischen Potentials aus der Poisson-Gleichung besteht darin, dass die Funktion (VII.21) nicht die einzige Greensche Funktion zur Poisson-Gleichung auf \mathbb{R}^3 ist.

Betrachtet man nämlich eine Lösung F der Laplace-Gleichung — eine harmonische Funktion — auf \mathbb{R}^3 , so ist $G_2(\vec{r}, \vec{r}') \equiv G(\vec{r}, \vec{r}') + F(\vec{r} - \vec{r}')$ auch eine mögliche Greensche Funktion.

$$\text{In der Tat gilt } \Delta_{\vec{r}} G_2(\vec{r}, \vec{r}') = \Delta_{\vec{r}} G(\vec{r}, \vec{r}') + \Delta_{\vec{r}} F(\vec{r} - \vec{r}') = \delta^{(3)}(\vec{r} - \vec{r}') + 0. \quad \square$$

Da es unendlich viele harmonische Funktionen auf \mathbb{R}^3 gibt, beispielsweise jede lineare Funktion der Form $F(\vec{r}) = ax^1 + bx^2 + cx^3$, gibt es auch unendlich viele Greensche Funktionen.

Unter den Greenschen Funktionen zur Poisson-Gleichung besitzt aber die Funktion (VII.21) eine einzigartige Eigenschaft: sie wird Null im Unendlichen $|\vec{r} - \vec{r}'| \rightarrow \infty$. Dementsprechend verschwindet das Potential (VII.22) weit von der Ladungsverteilung (wenn die letztere nur ein endliches Raumgebiet besetzt).

Genauer kann man zeigen, dass die einzigen harmonischen Funktionen auf \mathbb{R}^3 — und allgemeiner auf \mathbb{R}^n für jede $n \in \mathbb{N}^*$ —, die überall endlich bleiben, die identisch konstanten Funktionen sind.

Bemerkung: Addiert man zur Funktion (VII.21) eine Konstante $K \in \mathbb{R}$, so unterscheidet sich das aus Gl. (VII.20) resultierende Potential von dem Wert (VII.22) um eine additive Konstante — das Produkt aus K und der Gesamtladung der Verteilung geteilt durch ϵ_0 —, die nichts am elektrischen Feld ändert.

VII.2.3 Lösung der Poisson-Gleichung auf einem endlichen Gebiet von \mathbb{R}^3

VII.2.3a Mathematische Fragestellung und Ergebnisse

Anstatt einer Lösung der Poisson-Gleichung auf dem ganzen Raum kann man auch eine Lösung auf einem endlichen Gebiet \mathcal{V} von \mathbb{R}^3 suchen — entsprechend z.B. dem elektrostatischen Potential in einem leeren Hohlraum im Inneren von Materie. Ist die letztere irgendein Metall, in welchem frei bewegliche elektrische Ladungsträger vorhanden sind, so können sich solche Ladungen möglicherweise auf der Oberfläche $\partial\mathcal{V}$ des Bereichs \mathcal{V} befinden: dann dienen diese Ladungen als Quellen für das elektrische Skalarpotential im Hohlraum.

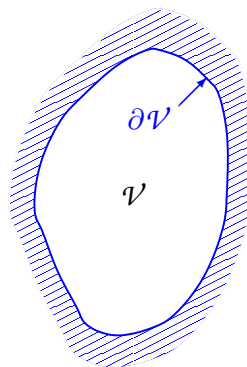


Abbildung VII.1

Meistens wird die entsprechende (Oberflächen)Ladungsdichte nicht explizit präzisiert. Stattdessen werden in der mathematischen Formulierung *Randbedingungen* für das festzustellende Potential $\Phi(\vec{r})$ vorgegeben, entsprechend seinem Verhalten auf der Oberfläche $\partial\mathcal{V}$, das physikalisch „von außen“ gesteuert wird. Dazu wird oft angenommen, dass im Hohlraum Vakuum herrscht, so dass sich die Poisson-Gleichung auf die Laplace-Gleichung $\Delta\Phi(\vec{r}) = 0$ vereinfacht. Bei der Fragestellung handelt es sich dann um ein sog. *Randwertproblem*.

Ein solches mathematisches Randwertproblem ist für physikalische Anwendungen nur dann nützlich, wenn es *wohlgestellt* ist. Dies bedeutet, dass das Problem eine Lösung hat, die eindeutig ist — möglicherweise bis auf eine additive Konstante — und stetig von den vorgegebenen Randbedingungen abhängt. Bei der Poisson-Gleichung ist es insbesondere der Fall für zwei einfache Arten von Randbedingungen:

- bei *Dirichlet*^(af)-Randbedingungen (oder *Randbedingungen erster Art*) werden die Werte des Potentials $\Phi(\vec{r})$ — oder allgemeiner der zu bestimmenden Funktion — an der Oberfläche $\partial\mathcal{V}$ vorgeschrieben;
- bei *Neumann*^(ag)-Randbedingungen (oder *Randbedingungen zweiter Art*) wird die Normalenableitung $\partial_n\Phi(\vec{r})$ vorgegeben, wobei die letztere die Änderungsrate in der Richtung senkrecht zur Oberfläche ist. Diese Ableitung lässt sich auch als $\vec{e}_n(\vec{r}) \cdot \vec{\nabla}\Phi(\vec{r})$ schreiben, wobei $\vec{e}_n(\vec{r})$ den Normaleinheitsvektor zur Oberfläche im Punkt $\vec{r} \in \partial\mathcal{V}$ bezeichnet.

Stimmt die Oberfläche $\partial\mathcal{V}$ lokal mit der (x^1, x^2) -Ebene überein, so ist der lokale Normaleneinheitsvektor \vec{e}_3 , d.h. $\vec{e}_n \cdot \vec{\nabla}\Phi$ ist gleich $\partial\Phi/\partial x^3$, was auch mit $\partial_3\Phi$ bezeichnet wird.

Bemerkung: Im Fall von Neumann-Bedingungen ist die Vorschrift von $\partial_n\Phi(\vec{r})$ auf $\partial\mathcal{V}$ äquivalent zur Angabe der normalen Komponente des elektrischen Feldes $\vec{E}(\vec{r}) = -\vec{\nabla}\Phi(\vec{r})$ in jedem Punkt der Oberfläche. Allgemeiner werden die (physikalischen!) Randbedingungen der Elektrostatik im § VII.2.3 c diskutiert.

Die Existenz einer Lösung der Poisson-Gleichung auf \mathcal{V} und ihre stetige Abhängigkeit von den Randbedingungen sind im allgemeinen Fall nicht trivial zu zeigen. Dagegen ist der Beweis der Eindeutigkeit bis auf eine additive Konstante ziemlich einfach.

Dieser Beweis macht die sog. *erste Greensche Identität*

$$\int_{\mathcal{V}} [\vec{\nabla}f_1(\vec{r}) \cdot \vec{\nabla}f_2(\vec{r}) + f_1(\vec{r})\Delta f_2(\vec{r})] d^3\vec{r} = \oint_{\partial\mathcal{V}} f_1(\vec{r}) \partial_n f_2(\vec{r}) d^2\mathcal{S} \quad (\text{VII.23})$$

zu Nutze, die für zwei (zweimal differenzierbare) reellwertige Funktionen f_1, f_2 gilt.

Beweis: Die Produktregel gibt $\vec{\nabla} \cdot (f_1 \vec{\nabla} f_2) = \vec{\nabla} f_1 \cdot \vec{\nabla} f_2 + f_1 \Delta f_2$, d.h. nach Integration und Anwendung des Gaußschen Integralsatzes

$$\int_{\mathcal{V}} [\vec{\nabla}f_1 \cdot \vec{\nabla}f_2 + f_1 \Delta f_2] d^3\vec{r} = \int_{\mathcal{V}} \vec{\nabla} \cdot (f_1 \vec{\nabla} f_2) d^3\vec{r} = \oint_{\partial\mathcal{V}} f_1 \vec{\nabla} f_2 \cdot d^2\vec{\mathcal{S}}.$$

Dabei ist das vektorielle Flächenelement $d^2\vec{\mathcal{S}}$ gleich $d^2\mathcal{S} \vec{e}_n$, woraus die Identität folgt. \square

Sind Φ_1, Φ_2 nämlich zwei Lösungen der Poisson-Gleichung auf \mathcal{V} mit beliebiger (aber gleicher!) Ladungsdichte ρ_{el} auf der rechten Seite und gegebenen Dirichlet- bzw. Neumann-Randbedingungen, dann ist ihre Differenz $U \equiv \Phi_2 - \Phi_1$ eine Lösung der Laplace-Gleichung $\Delta U(\vec{r}) = 0$ für $\vec{r} \in \mathcal{V}$, mit verschwindenden Dirichlet- bzw. Neumann-Randbedingungen $U(\vec{r}) = 0$ bzw. $\partial_n U(\vec{r}) = 0$ für $\vec{r} \in \partial\mathcal{V}$. Die erste Greensche Identität mit $f_1 = f_2 = U$ lautet dann

$$\int_{\mathcal{V}} [\vec{\nabla}U(\vec{r}) \cdot \vec{\nabla}U(\vec{r}) + U(\vec{r})\Delta U(\vec{r})] d^3\vec{r} = \oint_{\partial\mathcal{V}} U(\vec{r}) \partial_n U(\vec{r}) d^2\mathcal{S}.$$

Das Integrand auf der rechten Seite, und daher das Integral, verschwindet wegen der Randbedingungen. Wiederum ist der zweite Term in den eckigen Klammern auf der linken Seite ebenfalls Null, denn U erfüllt die Laplace-Gleichung. Deshalb bleibt nur

$$\int_{\mathcal{V}} [\vec{\nabla}U(\vec{r})]^2 d^3\vec{r} = 0$$

^(af)P. G. LEJEUNE DIRICHLET, 1805–1859 ^(ag)C. NEUMANN, 1832–1925

übrig, was nur dann möglich ist, wenn $\vec{\nabla}U(\vec{r})$ identisch auf \mathcal{V} verschwindet, d.h. wenn $U(\vec{r})$ konstant auf \mathcal{V} ist: $U(\vec{r}) = K$, mit $K = 0$ falls Dirichlet-Bedingungen vorgegeben sind. Somit unterscheiden sich Φ_1 und Φ_2 nur um eine additive Konstante.

VII.2.3b Physikalische Anwendungen

Faraday Käfig

Als erste Anwendung der obigen mathematischen Ergebnisse kann man das elektrische Feld in einem leeren Hohlraum innerhalb einer metallischen Oberfläche $\partial\mathcal{V}$ bestimmen.

Da $\partial\mathcal{V}$ metallisch ist, können sich darauf freie Ladungsträger befinden. Im statischen Fall dürfen sich diese Ladungen definitionsgemäß nicht bewegen. Demzufolge darf kein elektrisches Feld in der Metallfläche vorhanden sein, $\vec{E} = -\vec{\nabla}\Phi = \vec{0}$ auf $\partial\mathcal{V}$ — sonst würde das elektrische Feld zur Entstehung einer Stromdichte führen. Daher ist das Skalarpotential Φ konstant auf der Oberfläche: $\Phi(\vec{r}) = \Phi_0$ für $\vec{r} \in \partial\mathcal{V}$. Dies stellt eine Dirichlet-Randbedingung für das elektrostatische Potential im Hohlraum dar.

In \mathcal{V} soll Φ wegen der Abwesenheit von Ladungen Lösung der Laplace-Gleichung $\Delta\Phi(\vec{r}) = 0$ sein. Eine Lösung, die auch die Randbedingung erfüllt, ist einfach die konstante Funktion $\Phi(\vec{r}) = \Phi_0$ für $\vec{r} \in \mathcal{V}$. Wegen der Eindeutigkeit ist dies in der Tat die einzige Lösung. Daraus folgt dann $\vec{E} = -\vec{\nabla}\Phi = \vec{0}$ im Hohlraum, der als *Faraday-Käfig* bezeichnet wird.

Spiegelladungsmethode

Basierend auf der Eindeutigkeit der Lösung der Poisson-Gleichung für vorgegebene Dirichlet-Randbedingungen lässt sich eine andere Methode zur Lösung der Poisson-Gleichung in einem Gebiet von \mathbb{R}^3 formulieren.

Sei ein zu lösendes Problem \mathcal{P}_1 , bei dem eine Fläche $\partial\mathcal{V}$ ein Raumbereich \mathcal{V} abgrenzt, in dem Punktladungen gemäß einer Ladungsdichte $\rho_{\text{el}}(\vec{r}) = \sum_a q_a \delta^{(3)}(\vec{r} - \vec{x}_a)$ verteilt sind, wobei \vec{x}_a die Position der a -ten Punktladung bezeichnet. Als (Dirichlet)-Randbedingung wird angenommen, dass $\partial\mathcal{V}$ eine Äquipotentialfläche ist, d.h. $\Phi(\vec{r})$ ist konstant auf $\partial\mathcal{V}$.

Sei ein anderes Problem \mathcal{P}_2 ohne Fläche jedoch mit denselben Punktladungen $\{q_a\}$ an den gleichen Positionen $\{\vec{x}_a\}$ sowie zusätzlichen Ladungen — sog. „Spiegelladungen“ — außerhalb \mathcal{V} . Wenn das (lösbare!) Problem \mathcal{P}_2 zu einem Potential Φ führt, wovon eine Äquipotentialfläche mit der Fläche $\partial\mathcal{V}$ des Problems \mathcal{P}_1 übereinstimmt, dann ist das Potential von \mathcal{P}_2 im Volumen \mathcal{V} genau gleich dem gesuchten Potential von \mathcal{P}_1 .

Als Beispiel dieser Methode kann man das folgende Problem (\mathcal{P}_1) betrachten: im Halbraum $x < 0$ befindet sich ein elektrischer Leiter, dessen Oberfläche bei $x = 0$ eine Äquipotentialfläche ist;⁽⁶²⁾ im Punkt $\vec{x}_q \equiv (a, 0, 0)$ außerhalb des Leiters sitzt eine Punktladung q . Im Bereich $x > 0$ soll das Skalarpotential Lösung der Poisson-Gleichung $\Delta\Phi(\vec{r}) = -q\delta^{(3)}(\vec{r} - \vec{x}_q)/\epsilon_0$ sein, mit der Randbedingung $\Phi(\vec{r}) = \Phi_0$ für $x = 0$.

Ein passendes, einfach lösbares Problem \mathcal{P}_2 besteht aus zwei Punktladungen in einem sonst leeren Raum: q ist immer noch in \vec{x}_q , während eine zweite (Spiegel-)Punktladung q' im Punkt $-\vec{x}_q \equiv (-a, 0, 0)$ sitzt. Zusammen erzeugen diese Ladungen das Skalarpotential

$$\Phi(\vec{r}) = \frac{q}{4\pi\epsilon_0|\vec{r} - \vec{x}_q|} + \frac{q'}{4\pi\epsilon_0|\vec{r} + \vec{x}_q|},$$

das eine Lösung von $\Delta\Phi(\vec{r}) = \frac{-q}{\epsilon_0}\delta^{(3)}(\vec{r} - \vec{x}_q) - \frac{q'}{\epsilon_0}\delta^{(3)}(\vec{r} + \vec{x}_q)$ im ganzen Raum darstellt.

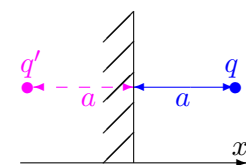


Abbildung VII.2

⁽⁶²⁾Dies ist immer der Fall bei elektrischen Leitern im Gleichgewicht, d.h. ohne Bewegung von Ladungen, wie schon bei der Beschreibung des Faraday-Käfigs argumentiert wurde.

Dieses Potential stellt auch eine spezielle Lösung der Poisson-Gleichung $\Delta\Phi(\vec{r}) = -q\delta^{(3)}(\vec{r} - \vec{x}_q)/\epsilon_0$ für $x \geq 0$ dar, die für $q' = -q$ konstant (und Null) bei $x = 0$ ist. Anders gesagt löst das Potential Φ das ursprüngliche Problem \mathcal{P}_1 für $x \geq 0$.

VII.2.3c Randbedingungen der Elektrostatik

Zur Lösung der Poisson-Gleichung für das elektrostatische Potential in einem abgegrenzten Raumgebiet \mathcal{V} müssen die Randbedingungen an der Oberfläche $\partial\mathcal{V}$ des Gebiets präzisiert werden. Dies ist anhand der stationären Maxwell-Gleichungen (VII.1) möglich.⁽⁶³⁾

Betrachte zuerst eine Kontour Γ um die Oberfläche $\partial\mathcal{V}$, mit zwei Seiten der Länge ℓ parallel zu $\partial\mathcal{V}$, entlang des Einheitsvektors \vec{e}_1 , und den zwei anderen Seiten der Länge δ normal zu $\partial\mathcal{V}$, längs \vec{e}_3 .⁽⁶⁴⁾ Somit ist die durch Γ abgeschlossene Fläche \mathcal{S} in der (x^1, x^2) -Ebene bzw. senkrecht zu \vec{e}_2 . Die Länge δ soll infinitesimal klein sein.

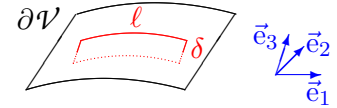


Abbildung VII.3

Laut dem Satz von Stokes ist der Fluss von $\vec{\nabla} \times \vec{E}$ durch \mathcal{S} durch

$$\int_{\mathcal{S}} [\vec{\nabla} \times \vec{E}(\vec{r})] \cdot \vec{e}_2 d^2\mathcal{S} = \oint_{\Gamma} \vec{E}(\vec{r}) \cdot d\vec{r} = [E^1(x^1, x^2, x^3=0^-) - E^1(x^1, x^2, x^3=0^+)]\ell + \mathcal{O}(\delta) = 0,$$

wobei die letzte Gleichheit aus Gl. (VII.1b) folgt, während die Notationen $x^3=0^+$, $x^3=0^-$ den Limes $\delta \rightarrow 0$ stillschweigend einbeziehen. Daher gilt $E^1(x^3=0^-) = E^1(x^3=0^+)$ d.h. E^1 ist stetig an der Oberfläche. Analog prüft man die Stetigkeit von E^2 nach. Insgesamt ist die Tangentialkomponente \vec{E}_{\parallel} des elektrischen Feldes stetig an der Oberfläche.

Im Gegensatz dazu ist die Normalkomponente \vec{E}_{\perp} des Feldes im Allgemeinen nicht stetig an der Oberfläche.

Sei v ein zylindrisches Volumen, abgegrenzt durch Elementarflächen $d\mathcal{S}$ auf den beiden Seiten der Oberfläche $\partial\mathcal{V}$ und durch die Röhre mit Höhe δ , die diese zwei Flächen miteinander verbindet. Die Elementarflächen sind parallel zur (x^1, x^2) -Ebene, die Röhre zwischen ihnen liegt entlang \vec{e}_3 .

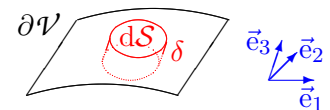


Abbildung VII.4

Bezeichnet man die Oberfläche des Volumens v mit ∂v , so gibt der Gaußsche Integralsatz

$$\int_v \vec{\nabla} \cdot \vec{E}(\vec{r}) d^3\vec{r} = \oint_{\partial v} \vec{E}(\vec{r}) \cdot d^2\vec{s} = \int_{d\mathcal{S}} [E^3(x^1, x^2, x^3=0^+) - E^3(x^1, x^2, x^3=0^-)] dx^1 dx^2 + \mathcal{O}(\delta),$$

wobei der Grenzwert $\delta \rightarrow 0$ implizit betrachtet wurde. Andererseits lässt sich $\vec{\nabla} \cdot \vec{E}(\vec{r})$ auch dank der Maxwell-Gauß-Gleichung (VII.1a) durch die elektrische Ladungsverteilung ρ_{el} ausdrücken. Im Limes $\delta \rightarrow 0$ trägt nur die auf $\partial\mathcal{V}$ verteilte Ladung zu ρ_{el} bei: die Ladungsverteilung kann durch $\rho_{\text{el}}(\vec{r}) = \sigma_{\text{el}}(x^1, x^2)\delta(x^3)$ angenähert werden, wobei σ_{el} eine Flächenladungsdichte ist. Somit gilt

$$\int_v \vec{\nabla} \cdot \vec{E}(\vec{r}) d^3\vec{r} = \int_v \frac{\rho_{\text{el}}(\vec{r})}{\epsilon_0} d^3\vec{r} = \int_{d\mathcal{S}} \frac{\sigma_{\text{el}}(x^1, x^2)}{\epsilon_0} dx^1 dx^2.$$

Der Vergleich mit dem oben erhaltenen Ausdruck für das gleiche Integral führt zu

$$E^3(x^1, x^2, x^3=0^+) - E^3(x^1, x^2, x^3=0^-) = \frac{\sigma_{\text{el}}(x^1, x^2)}{\epsilon_0}$$

in einem Punkt der Oberfläche, denn man kann $d\mathcal{S}$ so klein wie möglich nehmen. Die linke Seite dieser Gleichung ist der *Sprung* der Normalkomponenten des elektrischen Feldes an der Oberfläche.

⁽⁶³⁾Es wird angenommen, dass alle Felder im Unendlichen verschwinden.

⁽⁶⁴⁾Die Richtungen der Koordinatenachsen ändern sich in jedem Punkt \vec{r} der Oberfläche, so dass man \vec{e}_1 , \vec{e}_2 , \vec{e}_3 mit \vec{r} kennzeichnen sollte. \vec{e}_3 ($= \vec{e}_n$) ist in jedem Punkt der nach „außen“ gerichtete Normaleinheitsvektor zur Oberfläche, so dass $x^3=0^+$ bzw. $x^3=0^-$ einen Punkt im Äußeren bzw. im Inneren des Volumens \mathcal{V} bezeichnet.

Bezeichnet man diesen Sprung mit doppelten Klammern, so gilt $[[E^3(\vec{r})]] = \sigma_{\text{el.}}(\vec{r})/\epsilon_0$ oder äquivalent $[[\vec{E}_\perp(\vec{r})]] = [\sigma_{\text{el.}}(\vec{r})/\epsilon_0] \vec{e}_3$, d.h. unter Berücksichtigung der oben gefundenen Stetigkeit der Tangentialkomponente $\vec{E}_\parallel(\vec{r})$

$$\boxed{[[\vec{E}(\vec{r})]] = \frac{\sigma(\vec{r})}{\epsilon_0} \vec{e}_n(\vec{r}) \text{ in einem Punkt } \vec{r} \text{ einer Oberfläche,}} \quad (\text{VII.24})$$

wobei der Normaleinheitsvektor zur Oberfläche im Punkt \vec{r} nun mit $\vec{e}_n(\vec{r})$ bezeichnet wird.

Bemerkung: Eigentlich gilt dieses Ergebnis nur, wenn die „Oberfläche“ $\partial\mathcal{V}$ entweder eine (geladene) Fläche zwischen zwei leeren Bereichen ist, oder wenn sie die Grenzfläche zwischen einerseits einem leeren Hohlraum und andererseits einem elektrischen Leiter — in welchem im Gleichgewicht $\vec{E} = \vec{0}$ herrscht, wie bei der Beschreibung des Faraday-Käfigs argumentiert wurde — darstellt.

Wenn $\partial\mathcal{V}$ die Grenzfläche zwischen zwei unterschiedlichen Isolatoren (zwei „Dielektrika“) ist, so ist die Tangentialkomponente $\vec{E}_\parallel(\vec{r})$ noch für $\vec{r} \in \partial\mathcal{V}$ stetig. Dagegen ändert sich der Sprung der Normalkomponente $\vec{E}_\perp(\vec{r})$, weil die Maxwell-Gauß-Gleichung (VII.1a) innerhalb der Dielektrika nicht mehr relevant ist — anstelle des \vec{E} -Feldes soll man die elektrische Flussdichte, das \vec{D} -Feld, ins Betracht ziehen.

VII.3 Multipolentwicklung

Eine oft auftretende Fragestellung ist die Bestimmung des Einflusses einer gegebenen elektrischen Ladungsverteilung $\rho_{\text{el.}}$ mit typischer Ausdehnung R auf eine Punktladung — oder eine zweite Ladungsverteilung — in einem weit entfernten Punkt.

Wie wir schon gesehen haben, erzeugt die Ladungsverteilung das elektrische Skalarpotential in einem Punkt \vec{r}

$$\Phi(\vec{r}) = \int_{\mathcal{V}} \frac{\rho_{\text{el.}}(\vec{r}')}{4\pi\epsilon_0 |\vec{r} - \vec{r}'|} d^3\vec{r}'. \quad (\text{VII.12a})$$

wobei \mathcal{V} das durch die Verteilung besetzte Raumgebiet bezeichnet. Daraus lässt sich das elektrische Feld dann ableiten. Im Folgenden wird dieses Potential als eine Summe von sukzessiven Beiträgen geschrieben, die für $|\vec{r}| \gg R$ immer kleiner werden. Diese sog. *Multipolentwicklung* wird erstens in kartesischen Koordinaten (§ VII.3.1) und dann in Kugelkoordinaten (§ VII.3.4) durchgeführt.

Um die Berechnungen zu vereinfachen wird hiernach angenommen, dass die Ladungsverteilung um den Ursprung des Bezugssystems sitzt, d.h. $|\vec{r}'| \lesssim R$ für $\vec{r}' \in \mathcal{V}$. Dagegen soll der Punkt \vec{r} „weit entfernt“ sein, entsprechend $|\vec{r} - \vec{r}'| \gg R$ für $\vec{r}' \in \mathcal{V}$ oder äquivalent $|\vec{r}| \gg R$.

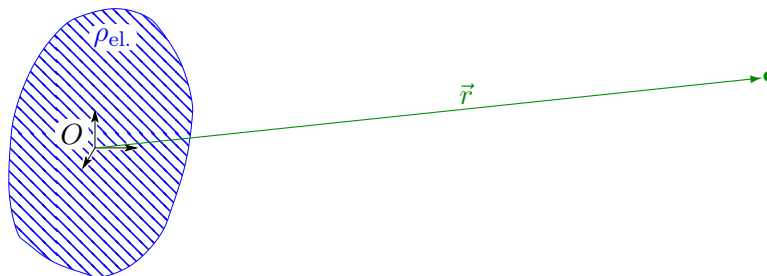


Abbildung VII.5

VII.3.1 Kartesische Multipolmomente

Aus den Ungleichungen $||\vec{r}| - |\vec{r}'|| \leq |\vec{r} - \vec{r}'| \leq |\vec{r}| + |\vec{r}'|$ und $|\vec{r}'| \lesssim R \ll |\vec{r}|$ folgt, dass der Abstand $|\vec{r} - \vec{r}'|$ im Nenner des Integranden in Gl. (VII.12a) ungefähr gleich $r \equiv |\vec{r}|$ ist. Dabei kann

die genaue Abweichung mithilfe einer Taylor-Entwicklung angenähert werden, wobei \vec{r}' die Rolle der „kleinen“ Verschiebung um den Bezugspunkt \vec{r} spielt.

Demgemäß lautet eine Taylor-Entwicklung bis zur zweiten Ordnung

$$\frac{1}{|\vec{r} - \vec{r}'|} = \frac{1}{|\vec{r} - \vec{r}'|} \Big|_{\vec{r}'=\vec{0}} + \sum_{i=1}^3 x'^i \frac{\partial}{\partial x'^i} \left(\frac{1}{|\vec{r} - \vec{r}'|} \right) \Big|_{\vec{r}'=\vec{0}} + \frac{1}{2} \sum_{i,j=1}^3 x'^i x'^j \frac{\partial^2}{\partial x'^i \partial x'^j} \left(\frac{1}{|\vec{r} - \vec{r}'|} \right) \Big|_{\vec{r}'=\vec{0}} + \dots,$$

wobei x^i bzw. x'^i mit $i \in \{1, 2, 3\}$ die kartesischen Komponenten von \vec{r} bzw. \vec{r}' in einem Koordinatensystem bezeichnet. Dabei ist der erste Term auf der rechten Seite, $1/|\vec{r} - \vec{r}'|$ ausgewertet in $\vec{r}' = \vec{0}$, einfach $1/r$.

Dann folgt aus $|\vec{r} - \vec{r}'| = \sqrt{(x^1 - x'^1)^2 + (x^2 - x'^2)^2 + (x^3 - x'^3)^2}$ die erste Ableitung

$$\frac{\partial}{\partial x'^i} \left(\frac{1}{|\vec{r} - \vec{r}'|} \right) = \frac{x^i - x'^i}{|\vec{r} - \vec{r}'|^3},$$

d.h. nach Auswertung in $\vec{r}' = \vec{0}$

$$\frac{\partial}{\partial x'^i} \left(\frac{1}{|\vec{r} - \vec{r}'|} \right) \Big|_{\vec{r}'=\vec{0}} = \frac{x^i}{r^3}.$$

Eine zweite Ableitung unter Berücksichtigung der Produktregel gibt

$$\frac{\partial}{\partial x'^j} \frac{\partial}{\partial x'^i} \left(\frac{1}{|\vec{r} - \vec{r}'|} \right) = -\frac{\delta^{ij}}{|\vec{r} - \vec{r}'|^3} + \frac{3(x^i - x'^i)(x^j - x'^j)}{|\vec{r} - \vec{r}'|^5},$$

d.h.

$$\frac{\partial^2}{\partial x'^i \partial x'^j} \left(\frac{1}{|\vec{r} - \vec{r}'|} \right) \Big|_{\vec{r}'=\vec{0}} = -\frac{\delta^{ij}}{r^3} + \frac{3x^i x^j}{r^5}.$$

Somit gilt insgesamt

$$\frac{1}{|\vec{r} - \vec{r}'|} = \frac{1}{r} + \sum_{i=1}^3 \frac{x^i}{r^3} x'^i + \sum_{i,j=1}^3 \frac{3x^i x^j - \delta^{ij} r^2}{2r^5} x'^i x'^j + \dots \quad \text{für } |\vec{r}'| \ll r \equiv |\vec{r}|. \quad (\text{VII.25})$$

Die Kombination $3x^i x^j - \delta^{ij} r^2$ ist *spurlos*, d.h. deren Multiplikation mit δ^{ij} gefolgt von einer Summe über alle Werte $i, j = 1, 2, 3$ ergibt Null.⁽⁶⁵⁾ Daher kann der quadratische Term in den x'^i gemäß

$$\sum_{i,j=1}^3 \frac{3x^i x^j - \delta^{ij} r^2}{2r^5} x'^i x'^j = \sum_{i,j=1}^3 \frac{3x^i x^j - \delta^{ij} r^2}{2r^5} \left(x'^i x'^j - \frac{\delta^{ij}}{3} |\vec{r}'|^2 \right)$$

erweitert werden. Dies führt zur Entwicklung

$$\frac{1}{|\vec{r} - \vec{r}'|} = \frac{1}{r} + \sum_{i=1}^3 \frac{x^i}{r^3} x'^i + \sum_{i,j=1}^3 \frac{3x^i x^j - \delta^{ij} r^2}{2r^5} \left(x'^i x'^j - \frac{\delta^{ij}}{3} |\vec{r}'|^2 \right) + \dots \quad (\text{VII.26})$$

Da die Komponente x^i bzw. x'^i typischerweise der Größenordnung $\mathcal{O}(r)$ bzw. $\mathcal{O}(R)$ ist, ist der erste Term der Ordnung (sogar gleich!) $1/r$, der zweite der Ordnung $\mathcal{O}(R/r^2)$ und der dritte der Ordnung $\mathcal{O}(R^2/r^3)$. Somit ist jeder neue Term in der Entwicklung um einen Faktor der Ordnung $R/r \ll 1$ kleiner als der vorige.

Nach Einsetzen der Gl. (VII.26) in den Ausdruck (VII.12a) des Skalarpotentials ergibt sich

$$\Phi(\vec{r}) = \frac{Q}{4\pi\epsilon_0 r} + \frac{\vec{r} \cdot \vec{P}}{4\pi\epsilon_0 r^3} + \sum_{i,j=1}^3 \frac{3x^i x^j - \delta^{ij} r^2}{8\pi\epsilon_0 r^5} Q^{ij} + \dots, \quad (\text{VII.27a})$$

⁽⁶⁵⁾Betrachtet man die Matrix \mathcal{A} mit Elementen $a_{ij} \equiv 3x^i x^j - \delta^{ij} r^2$, so ist die betrachtete Operation genau das Bilden der Spur $\text{Tr } \mathcal{A} \equiv \sum_i a_{ii}$ der Matrix, daher die Bezeichnung.

wobei die ersten *kartesischen elektrischen Multipolmomente* durch

$$Q \equiv \int_{\mathcal{V}} \rho_{\text{el.}}(\vec{r}') d^3\vec{r}' \quad (\text{VII.27b})$$

$$\vec{P} \equiv \int_{\mathcal{V}} \vec{r}' \rho_{\text{el.}}(\vec{r}') d^3\vec{r}' \quad (\text{VII.27c})$$

$$Q^{ij} \equiv \int_{\mathcal{V}} \left(x'^i x'^j - \frac{\delta^{ij}}{3} |\vec{r}'|^2 \right) \rho_{\text{el.}}(\vec{r}') d^3\vec{r}' \quad (\text{VII.27d})$$

definiert sind: Q ist die schon bekannte Gesamtladung. Dann heißt \vec{P} *elektrisches Dipolmoment* der Ladungsverteilung $\rho_{\text{el.}}$. Die Zahlen Q^{ij} — die *elektrischen Quadrupolmomente* — sind die Komponenten eines Tensors zweiter Stufe \mathbf{Q} , des *elektrischen Quadrupoltensors*.

Man sieht sofort, dass der Quadrupoltensor symmetrisch — d.h. $Q^{ij} = Q^{ji}$ für alle i, j — und spurlos — d.h. $\sum_i Q^{ii} = 0$ — ist, wobei die letztere Eigenschaft schon beim Integranden auftritt. Demzufolge sind nur 5 der 9 Komponenten unabhängig.

Bemerkungen:

* Der erste Term im Skalarpotential (VII.27a) ist proportional zu $1/r$, der zweite proportional zu $1/r^2$, der dritte proportional zu (oder zumindest der Ordnung) $1/r^3$, usw. Somit tragen die niedrigeren Multipolmomente am meisten zu $\Phi(\vec{r})$ bei.⁽⁶⁶⁾

* Das elektrische Dipolmoment bzw. Quadrupolmoment hat physikalische Dimension $[\vec{P}] = \text{L T I}$ bzw. $[Q^{ij}] = \text{L}^2 \text{T I}$, entsprechend der SI-Einheit $\text{C} \cdot \text{m}$ bzw. $\text{C} \cdot \text{m}^2$.

* Im dritten Summanden der Multipolentwicklung (VII.27a) wird oft der Term $-\delta^{ij} r^2$ im Zähler des Vorfaktors von Q^{ij} weggelassen: wegen der Spurlosigkeit von Q^{ij} ist der Beitrag dieses Terms in der Tat Null.

VII.3.2 Beispiele von elektrischen Multipolmomenten

VII.3.2a Kugelsymmetrische Ladungsverteilung

Ein einfaches Beispiel von Ladungsverteilung ist jenes einer Verteilung $\rho_{\text{el.}}(\vec{r}')$, die nur von $|\vec{r}'|$ abhängt. Dementsprechend ist $\rho_{\text{el.}}$ eine gerade Funktion von \vec{r}' — d.h. $\rho_{\text{el.}}(\vec{r}') = \rho_{\text{el.}}(-\vec{r}')$ — bzw. von jeder der Komponenten x'^i .

Abhängig von der Form von $\rho_{\text{el.}}(\vec{r}')$ wird sich über Gl. (VII.27b) eine bestimmte Gesamtladung Q ergeben, die ohne zusätzliche Information nicht präzisiert werden kann.

Bezüglich des Dipolmoments kann man bemerken, dass das Integrand $\vec{r}' \rho_{\text{el.}}(\vec{r}')$ der definierenden Gleichung (VII.27c) jetzt eine ungerade Funktion von \vec{r}' , die über ein symmetrisches Gebiet integriert wird, woraus sich automatisch Null ergibt, unabhängig von der genauen funktionalen Form von $\rho_{\text{el.}}$.

Ein detaillierterer Beweis besteht in das Ersetzen von \vec{r}' durch $-\vec{r}'$ im zweiten Integral von

$$\vec{P} = \frac{1}{2} \left[\int \vec{r}' \rho_{\text{el.}}(\vec{r}') d^3\vec{r}' + \int \vec{r}' \rho_{\text{el.}}(\vec{r}') d^3\vec{r}' \right].$$

Diese Substitution gibt dann genau das Negative des ersten Integrals, voraus $\vec{P} = \vec{0}$ folgt. \square

Mithilfe einer ähnlichen Argumentation findet man, dass die Quadrupolmomente Q^{ij} für $i \neq j$ verschwinden. In der Tat lauten sie [Gl. (VII.27d)]

$$Q^{ij} = \int_{\mathcal{V}} x'^i x'^j \rho_{\text{el.}}(\vec{r}') d^3\vec{r}' \quad \text{für } i \neq j.$$

⁽⁶⁶⁾... wenn sie nicht Null sind!

Dabei ist das Integrand eine ungerade Funktion von x^i , die über ein symmetrisches Intervall integriert wird, was zu $Q^{ij} = 0$ führt. Somit ist Q^{ij} diagonal.

Schließlich sind die drei Koordinatenachsen wegen der angenommenen Kugelsymmetrie äquivalent, so dass $Q^{11} = Q^{22} = Q^{33}$. Da der Quadrupoltensor spurlos sein soll, ergibt sich $Q^{ii} = 0$ für jedes $i = 1, 2, 3$, d.h. insgesamt $\mathbf{Q} = \mathbf{0}$.

Bemerkung: Einige der obigen Ergebnisse lassen sich auch „geometrisch“ finden. Die angenommene Symmetrie bedeutet, dass keine Richtung bzw. kein Vektor von \mathbb{R}^3 eine besondere Rolle spielt, sondern dass sie alle äquivalent sind. Dementsprechend gibt es keine bevorzugte Richtung für das Dipolmoment \vec{P} , das deshalb verschwinden muss. Auf ähnlicher Weise darf der Tensor \mathbf{Q} keine eindeutig (bis auf einen multiplikativen Faktor) definierte Eigenvektoren haben, sondern müssen alle Vektoren von \mathbb{R}^3 Eigenvektoren sein: somit muss \mathbf{Q} proportional zur Identität sein, d.h. $Q^{ij} \propto \delta^{ij}$. Dann liefert die Spurlosigkeit den Proportionalitätsfaktor.

Da kugelsymmetrische Ladungsverteilungen keine elektrische Dipol- oder Quadrupolmomente haben — und eigentlich auch keine höhere Momente —, kann man reziprok folgern, dass solche Multipolmomente Abweichungen von der Kugelsymmetrie signalisieren.

VII.3.2b Einfacher Dipol

Der einfachste mögliche elektrische Dipol besteht aus zwei um einen Abstandsvektor \vec{a} getrennten Punktladungen q und $-q$.

Bezeichnet man mit \vec{r}_0 den Ortsvektor des Mittelpunkts, so lautet die zugehörige Ladungsdichte

$$\rho_{\text{el.}}(\vec{r}') = q\delta^{(3)}\left(\vec{r}' - \vec{r}_0 - \frac{\vec{a}}{2}\right) - q\delta^{(3)}\left(\vec{r}' - \vec{r}_0 + \frac{\vec{a}}{2}\right).$$



Abbildung VII.6

Eingesetzt in die Gl. (VII.27b)–(VII.27d) führt diese Verteilung zur Gesamtladung $Q = 0$, zum elektrischen Dipolmoment

$$\vec{P} = q\vec{a} \quad (\text{VII.28})$$

und zu einem verschwindenden Quadrupoltensor $\mathbf{Q} = \mathbf{0}$.

VII.3.2c Beispiel eines Quadrupols

Ein einfaches Beispiel von einer Ladungsverteilung mit nicht-trivialem elektrischem Quadrupoltensor besteht aus zwei Punktladungen q in den Punkten $\vec{r}_0 \pm \vec{a}$ und einer dritten Punktladung $-2q$ im Mittelpunkt \vec{r}_0 . Die zugehörige Ladungsdichte ist

$$\rho_{\text{el.}}(\vec{r}') = q\left[\delta^{(3)}(\vec{r}' - \vec{r}_0 - \vec{a}) + \delta^{(3)}(\vec{r}' - \vec{r}_0 + \vec{a}) - 2\delta^{(3)}(\vec{r}' - \vec{r}_0)\right],$$

was sofort $Q = 0$ und $\vec{P} = \vec{0}$ für die ersten zwei Multipolmomente ergibt.

Die x^3 -Achse sei entlang \vec{a} .

Die Ladungsdichte ist eine gerade Funktion von x'^1 und x'^2 . Wenn $i \neq j$ mit $i, j \in \{1, 2, 3\}$ muss entweder i oder j gleich 1 oder 2 sein, z.B. $i = 1$: dann ist das Integrand $x'^1 x'^j \rho_{\text{el.}}(\vec{r}')$ von Q^{1j} ungerade in x'^1 , was nach Integration über \mathbb{R} das Quadrupolmoment $Q^{1j} = 0$ ergibt.

Wegen der Symmetrie um die x^3 -Achse sind die x^1 - und x^2 -Richtungen äquivalent, so dass $Q^{11} = Q^{22}$. Aus der Spurlosigkeit von \mathbf{Q} folgt dann

$$Q^{11} = Q^{22} = -\frac{Q^{33}}{2}.$$

Schließlich kann man Q^{33} direkt berechnen:

$$Q^{33} = \int \left[(x'^3)^2 - \frac{1}{3} |\vec{r}'|^2 \right] \rho_{\text{el.}}(\vec{r}') d^3\vec{r}' = \int \left[\frac{2(x'^3)^2}{3} - \frac{(x'^1)^2 + (x'^2)^2}{3} \right] \rho_{\text{el.}}(\vec{r}') d^3\vec{r}'$$

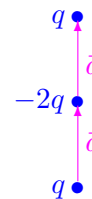


Abbildung VII.7

und nach Einsetzen der Ladungsdichte kommt

$$Q^{33} = \frac{2}{3} \left[q(x_0^3 - a)^2 + q(x_0^3 + a)^2 - 2q(x_0^3)^2 \right] = \frac{4}{3}qa^2$$

wobei x_0^3 die relevante Komponente von \vec{r}_0 bezeichnet.

Bemerkung: Eine andere geometrische Anordnung, die zu einem nicht-verschwindenden Quadrupol-tensor bei $Q = 0$ und $\vec{P} = \vec{0}$ führt, besteht aus 4 Punktladungen $q, -q, q, -q$ an den Ecken eines Quadrats. In diesem Fall ist \mathbf{Q} aber nicht diagonal.

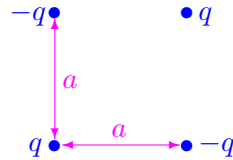


Abbildung VII.8

VII.3.3 Wechselwirkung zwischen zwei Ladungsverteilungen

Die Wechselwirkung zwischen zwei elektrischen Punktladungen ist durch die bekannte Coulomb-Kraft (VII.14a) gegeben: sie ist proportional zum Produkt der Ladungen und lässt sich aus der Potentialenergie (VII.14b) ableiten. In diesem Paragraphen wird die Wechselwirkung zwischen zwei beliebigen statischen Ladungsverteilungen $\rho_{\text{el.},a}$, $\rho_{\text{el.},b}$ untersucht, insbesondere für den Fall, wenn diese weit von einander sind — d.h. falls deren Abstand viel größer als ihre typische Größe ist.

VII.3.3 a Wechselwirkungsenergie zweier Ladungsverteilungen

Sei Φ_a bzw. Φ_b das von der Verteilung $\rho_{\text{el.},a}$ bzw. $\rho_{\text{el.},b}$ gemäß der Poisson-Gleichung (VII.4) herrührende elektrische Skalarpotential. Laut dem schon in § VII.1.3 b benutzten Superpositionsprinzip erzeugt die Ladungsverteilung $\rho_{\text{el.},a} + \rho_{\text{el.},b}$ das Potential $\Phi_a + \Phi_b$. Unter Verwendung der Gl. (VII.16) lautet dann die gesamte Potentialenergie des Systems aus den zwei Verteilungen

$$V = \frac{1}{2} \int [\rho_{\text{el.},a}(\vec{r}) + \rho_{\text{el.},b}(\vec{r})] [\Phi_a(\vec{r}) + \Phi_b(\vec{r})] d^3\vec{r}. \quad (\text{VII.29})$$

Nach trivialem Ausmultiplizieren des Integranden wird diese Energie zu

$$\begin{aligned} V &= \frac{1}{2} \int \rho_{\text{el.},a}(\vec{r}) \Phi_a(\vec{r}) d^3\vec{r} + \frac{1}{2} \int \rho_{\text{el.},b}(\vec{r}) \Phi_b(\vec{r}) d^3\vec{r} + \frac{1}{2} \int [\rho_{\text{el.},a}(\vec{r}) \Phi_b(\vec{r}) + \rho_{\text{el.},b}(\vec{r}) \Phi_a(\vec{r})] d^3\vec{r} \\ &\equiv V_a + V_b + V_W \end{aligned} \quad (\text{VII.30})$$

Die zwei ersten Terme V_a , V_b sind der Art (VII.16) mit im Integranden dem Produkt aus einer Ladungsdichte und dem dadurch verursachten Skalarpotential. Diese Terme stehen für die jeweiligen „Selbstenergien“ der Verteilungen $\rho_{\text{el.},a}$ und $\rho_{\text{el.},b}$, und entsprechen ihren potentiellen Energien in Abwesenheit der anderen Verteilung.

Dagegen koppelt der dritte Term V_W jede Ladungsdichte mit dem Potential, das die andere Verteilung erzeugt. Dieser Beitrag stellt die *Wechselwirkungsenergie* der Ladungsverteilungen dar. Er lässt sich mithilfe eines Tricks umschreiben.

Aus der Kettenregel folgen nämlich (die Variablen werden nicht geschrieben) die Identitäten

$$\vec{\nabla} \cdot (\Phi_a \vec{\nabla} \Phi_b) = \vec{\nabla} \Phi_a \cdot \vec{\nabla} \Phi_b + \Phi_a \Delta \Phi_b \quad \text{und} \quad \vec{\nabla} \cdot (\Phi_b \vec{\nabla} \Phi_a) = \vec{\nabla} \Phi_b \cdot \vec{\nabla} \Phi_a + \Phi_b \Delta \Phi_a.$$

Dabei sind die Skalarprodukte auf der rechten Seite beider Gleichungen gleich. Daraus ergibt sich

$$\Phi_a \Delta \Phi_b = \Phi_b \Delta \Phi_a + \vec{\nabla} \cdot [(\Phi_a \vec{\nabla} \Phi_b) - (\Phi_b \vec{\nabla} \Phi_a)].$$

Integriert man diese Gleichung über ein Volumen, so lässt sich das Integral des letzten Terms dank dem Gaußschen Integralsatz in ein Oberflächenintegral transformieren, und zwar

$$\oint_{\partial V} [(\Phi_a \vec{\nabla} \Phi_b) - (\Phi_b \vec{\nabla} \Phi_a)] \cdot d^2\vec{S}.$$

Sei \mathcal{V} eine um den Nullpunkt zentrierte Kugel mit Radius L , welche die zwei Ladungsverteilungen einschließt. Ist die Kugeloberfläche $\partial\mathcal{V}$ weit entfernt von den Verteilungen, d.h. für L groß genug, so lässt sich das Potential Φ_i mit $i = a$ oder b in einem Punkt von $\partial\mathcal{V}$ durch die Multipolentwicklung (VII.27a) annähern. Insbesondere nimmt Φ_i in einem solchen Punkt wie $1/L$ ab, möglicherweise noch schneller falls $Q_i = 0$. Dementsprechend ist $|\vec{\nabla}\Phi_i|$ proportional zu $1/L^2$ oder kleiner, und $|\Phi_j\vec{\nabla}\Phi_i|$ proportional zu $1/L^3$. Da sich das Flächenelement $d^2\mathcal{S}$ als $L^2 d^2\Omega$ mit dem Raumwinkelelement $d^2\Omega$ schreiben lässt, findet man, dass das obige Oberflächenintegral proportional zu $1/L^3 \cdot L^2 = 1/L$ ist: es verschwindet im Grenzwert $L \rightarrow \infty$, d.h. wenn $\mathcal{V} \rightarrow \mathbb{R}^3$.⁽⁶⁷⁾ Somit gilt

$$\int_{\mathbb{R}^3} \Phi_a(\vec{r}) \Delta \Phi_b(\vec{r}) d^3\vec{r} = \int_{\mathbb{R}^3} \Phi_b(\vec{r}) \Delta \Phi_a(\vec{r}) d^3\vec{r}.$$

Multipliziert mit $-\epsilon_0$ und unter Verwendung von $-\epsilon_0 \Delta \Phi_i(\vec{r}) = \rho_{\text{el.},i}(\vec{r})$ wird dies zu

$$\int \Phi_a(\vec{r}) \rho_{\text{el.},b}(\vec{r}) d^3\vec{r} = \int \Phi_b(\vec{r}) \rho_{\text{el.},a}(\vec{r}) d^3\vec{r},$$

so dass die in Gl. (VII.30) definierte Wechselwirkungsenergie als

$$V_W = \int \rho_{\text{el.},a}(\vec{r}) \Phi_b(\vec{r}) d^3\vec{r} = \int \rho_{\text{el.},b}(\vec{r}) \Phi_a(\vec{r}) d^3\vec{r} \quad (\text{VII.31})$$

geschrieben werden kann.

VII.3.3b Wechselwirkungsenergie zweier entfernter Ladungsverteilungen

Sei jetzt angenommen, dass die zwei Ladungsverteilungen $\rho_{\text{el.},a}$, $\rho_{\text{el.},b}$ weit entfernt von einander sind, wobei die Verteilung $\rho_{\text{el.},a}$ in einer Umgebung des Koordinatennullpunkts lokalisiert ist.



Abbildung VII.9

Die Taylor-Entwicklung des von der Verteilung $\rho_{\text{el.},b}$ herrührenden Skalarpotentials $\Phi_b(\vec{r})$ bis zur zweiten Ordnung um den Wert $\vec{r} = \vec{0}$ lautet

$$\Phi_b(\vec{r}) = \Phi_b(\vec{0}) + \sum_{i=1}^3 x^i \frac{\partial \Phi_b(\vec{0})}{\partial x^i} + \frac{1}{2} \sum_{i,j=1}^3 x^i x^j \frac{\partial^2 \Phi_b(\vec{0})}{\partial x^i \partial x^j} + \dots$$

Im linearen Term sind die Ableitungen von Φ_b genau die Komponenten dessen Gradienten; dann erkennt man das Skalarprodukt aus dem letzteren mit dem Vektor \vec{r} . Dazu kann der quadratische Term erweitert werden, um $x^i x^j$ als die Summe eines spurlosen und eines diagonalen Beitrags zu schreiben:

$$\Phi_b(\vec{r}) = \Phi_b(\vec{0}) + \vec{r} \cdot \vec{\nabla} \Phi_b(\vec{0}) + \frac{1}{2} \sum_{i,j=1}^3 \left(x^i x^j - \frac{\vec{r}^2}{3} \delta^{ij} \right) \frac{\partial^2 \Phi_b(\vec{0})}{\partial x^i \partial x^j} + \frac{\vec{r}^2}{6} \sum_{i,j=1}^3 \delta^{ij} \frac{\partial^2 \Phi_b(\vec{0})}{\partial x^i \partial x^j} + \dots$$

Im letzten Term ist die Summe genau gleich dem Laplace-Operator $\Delta \Phi_b(\vec{0})$: unter Verwendung der Poisson-Gleichung ist dieser Beitrag proportional zur Ladungsdichte $\rho_{\text{el.},b}(\vec{0})$, die per Annahme Null

⁽⁶⁷⁾Die hier verwendete Argumentation ist genau die gleiche wie für den Beweis von Gl. (VII.17a) in § VII.1.4 c.

ist — die Verteilung $\rho_{\text{el.},b}$ ist anderswo lokalisiert. Ersetzt man dazu $\vec{\nabla}\Phi_b$ durch $-\vec{E}_b$, und drückt man dementsprechend die zweiten Ableitungen von Φ_b durch erste Ableitungen von \vec{E}_b aus, so ergibt sich

$$\Phi_b(\vec{r}) = \Phi_b(\vec{0}) - \vec{r} \cdot \vec{E}_b(\vec{0}) - \frac{1}{2} \sum_{i,j=1}^3 \left(x^i x^j - \frac{\vec{r}^2}{3} \delta^{ij} \right) \frac{\partial E_b^i(\vec{0})}{\partial x^j} + \dots$$

Dieser Ausdruck des Skalarpotentials kann dann in die Wechselwirkungsenergie (VII.31) — in der Form mit dem Integranden $\rho_{\text{el.},a}(\vec{r})\Phi_b(\vec{r})$ — eingesetzt werden. Dabei sind $\Phi_b(\vec{0})$, $\vec{E}_b(\vec{0})$ und $\partial E_b^i(\vec{0})/\partial x^j$ und unabhängig von der Integrationsvariablen \vec{r} , und können deshalb aus den jeweiligen Integralen herausgezogen werden. Unter Verwendung der kartesischen elektrischen Multipolmomente (VII.27b)–(VII.27d) der Verteilung a ergibt sich dann

$$V_W = Q_a \Phi_b(\vec{0}) - \vec{P}_a \cdot \vec{E}_b(\vec{0}) - \frac{1}{2} \sum_{i,j=1}^3 Q_a^{ij} \frac{\partial E_b^i(\vec{0})}{\partial x^j} + \dots \quad (\text{VII.32})$$

Bemerkungen:

* Wegen der Symmetrie der Wechselwirkungsenergie (VII.31) gegenüber dem Austausch der beiden Ladungsverteilungen kann V_W auch ausgedrückt werden durch die elektrischen Multipolmomente der Verteilung $\rho_{\text{el.},b}$ und das von $\rho_{\text{el.},a}$ erzeugte Skalarpotential $\Phi_a(\vec{r}_0)$ und dessen Ableitungen ($\vec{E}_a(\vec{r}_0)$, $\partial E_a^i(\vec{r}_0)/\partial x^j$, ...) in einem Punkt \vec{r}_0 des durch die Verteilung $\rho_{\text{el.},b}$ besetzten Raumbereichs:

$$V_W = Q_b \Phi_a(\vec{r}_0) - \vec{P}_b \cdot \vec{E}_a(\vec{r}_0) - \frac{1}{2} \sum_{i,j=1}^3 Q_b^{ij} \frac{\partial E_a^i(\vec{r}_0)}{\partial x^j} + \dots \quad (\text{VII.33})$$

Dabei sollen die Multipolmomente relativ zum Punkt \vec{r}_0 berechnet werden, d.h. mit $\vec{r}' - \vec{r}_0$ bzw. $x'^i - x_0^i$ statt \vec{r}' bzw. x'^i in den Vorfaktoren von $\rho_{\text{el.}}(\vec{r}')$ in Gl. (VII.27c)–(VII.27d).

* Die sukzessiven Terme in der Formel (VII.32) werden im Prinzip — d.h. wenn das zugehörige Multipolmoment nicht Null ist — immer kleiner mit wachsender Ordnung der darin auftretenden Ableitung von Φ .

Somit ist der führende Beitrag jener der Gesamtladung, der genau gleich der Potentialenergie einer Punktladung Q_a in einem äußeren (d.h. durch andere Ladungen erzeugten) Skalarpotential Φ_b ist.

* Wenn die Verteilung a sich drehen kann, dann wird sie es so machen, dass die Wechselwirkungsenergie minimiert wird, entsprechend einem stabilen Gleichgewicht. Falls $\vec{P}_a \neq \vec{0}$ ist, wird man deshalb zu einer Situation tendieren, in der \vec{P}_a parallel zum elektrischen Feld $\vec{E}_b(\vec{0})$ ist.

VII.3.3c Dipol-Dipol-Wechselwirkung

Sei jetzt angenommen, dass beide Verteilungen „reine Dipole“ sind, d.h. nur deren elektrische Dipolmomente sind nicht Null, während alle anderen Momente, einschließlich der Gesamtladung, verschwinden. Der elektrische Dipol a befindet sich im Nullpunkt, der Dipol b in einem Punkt \vec{r} .

Unter einem „reinen Dipol“ versteht man eigentlich ein punktförmiges Objekt, versehen mit einem elektrischen Dipolmoment, jedoch ohne Gesamtladung oder höhere Multipolmomente. Wegen der verschwindenden Ausdehnung gilt das Skalarpotential (VII.34a) überall im Raum (bis auf $\vec{r} = \vec{0}$, wo der Dipol sitzt), anstatt nur in großem Abstand vom Dipol, wie es beim einfachen „physikalischen“ Dipol des § VII.3.2 b der Fall ist.

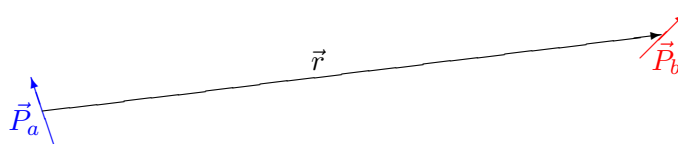


Abbildung VII.10 – Wechselwirkung zweier Dipole.

Dann ist das durch den Dipol a erzeugte Skalarpotential $\Phi_a(\vec{r})$ laut Gl. (VII.27a) durch

$$\Phi_a(\vec{r}) = \frac{\vec{r} \cdot \vec{P}_a}{4\pi\epsilon_0 r^3} \quad \text{mit } r \equiv |\vec{r}| \neq 0 \quad (\text{VII.34a})$$

gegeben. Dieser Ausdruck führt nach Ableitung zum *elektrischen Dipolfeld*

$$\vec{E}_a(\vec{r}) = -\vec{\nabla}\Phi_a(\vec{r}) = -\frac{\vec{P}_a}{4\pi\epsilon_0 r^3} + \frac{3(\vec{r} \cdot \vec{P}_a)\vec{r}}{4\pi\epsilon_0 r^5}$$

d.h.

$$\vec{E}_a(\vec{r}) = \frac{3(\vec{r} \cdot \vec{P}_a)\vec{r} - r^2\vec{P}_a}{4\pi\epsilon_0 r^5}. \quad (\text{VII.34b})$$

Das Einsetzen dieses Feldes in die Wechselwirkungsenergie (VII.33), in der nur der Dipolterm betrachtet wird, ergibt

$$V_W = \frac{\vec{P}_a \cdot \vec{P}_b - 3(\vec{e}_r \cdot \vec{P}_a)(\vec{e}_r \cdot \vec{P}_b)}{4\pi\epsilon_0 r^3} \quad (\text{VII.35})$$

wobei $\vec{e}_r \equiv \vec{r}/|\vec{r}|$ den Einheitsvektor in Richtung \vec{r} bezeichnet. Dies stellt die Potentialenergie der Dipol-Dipol-Wechselwirkung dar.

Bemerkung: Das elektrische Skalarpotential (VII.34a) bzw. Dipolfeld (VII.34b) nimmt mit dem Abstand r zur Quelle wie $1/r^2$ bzw. $1/r^3$ ab, d.h. schneller als das durch eine Punktladung erzeugte Potential bzw. Feld ($1/r$ bzw. $1/r^2$, § VII.1.3 a).

VII.3.4 Multipolentwicklung in Kugelkoordinaten

Die Multipolentwicklung kann auch in Kugelkoordinaten (r, θ, φ) formuliert werden, wobei die Ladungsverteilung, die das Skalarpotential verursacht, im Nullpunkt des Bezugssystems sitzt.

Diese Entwicklung basiert auf die folgende Formel, die für beliebige Vektoren \vec{r}, \vec{r}' gilt, mit den Bezeichnungen $r_{<} \equiv \min(r, r')$ und $r_{>} \equiv \max(r, r')$ für $r \equiv |\vec{r}|$, $r' \equiv |\vec{r}'|$:

$$\frac{1}{|\vec{r} - \vec{r}'|} = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \frac{4\pi}{2\ell+1} \frac{r_{<}^{\ell}}{r_{>}^{\ell+1}} Y_{\ell m}^*(\theta', \varphi') Y_{\ell m}(\theta, \varphi). \quad (\text{VII.36})$$

Dabei sind die $Y_{\ell m}$ die *Kugelflächenfunktionen*,⁽⁶⁸⁾ und (θ, φ) bzw. (θ', φ') die Polar- und Azimutwinkel des Ortsvektors \vec{r} bzw. \vec{r}' . Setzt man diese Identität in den Ausdruck (VII.12a) des Skalarpotentials in einem Punkt \vec{r} ein, wobei $|\vec{r}| \geq |\vec{r}'|$ für jeden \vec{r}' in der Verteilung, so ergibt sich

$$\Phi(\vec{r}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \sqrt{\frac{4\pi}{2\ell+1}} q_{\ell m} \frac{Y_{\ell m}(\theta, \varphi)}{4\pi\epsilon_0 r^{\ell+1}}, \quad (\text{VII.37a})$$

wobei die elektrischen Multipolmomenten $q_{\ell m}$ durch

$$q_{\ell m} \equiv \sqrt{\frac{4\pi}{2\ell+1}} \int (r')^{\ell} Y_{\ell m}^*(\theta', \varphi') \rho_{\text{el.}}(\vec{r}') d^3\vec{r}' \quad (\text{VII.37b})$$

definiert sind.

Jeder ganzen Zahl $\ell \in \mathbb{N}$ sind genau $2\ell + 1$ mögliche Werte von m zugeordnet, d.h. $2\ell + 1$ Multipolmomente $q_{\ell m}$. Für $\ell = 0$ prüft man mit $Y_{00}(\theta, \varphi) = 1/\sqrt{4\pi}$ einfach nach, dass q_{00} gleich der Gesamtladung Q ist. Dann sind die 3 bzw. 5 Multipolmomente q_{1m} bzw. q_{2m} Linearkombinationen der 3 bzw. 5 unabhängigen Komponenten des Dipolmoments \vec{P} bzw. des Quadrupoltensors \mathbf{Q} .

Allgemeiner entsprechen die $2\ell + 1$ Momente $q_{\ell m}$ den $2\ell + 1$ unabhängigen Komponenten des symmetrischen, spurlosen elektrischen 2^{ℓ} -poltensors, wobei der letztere ℓ -ter Stufe ist.

⁽⁶⁸⁾Vgl. Anhang F.

Bemerkungen:

* Für eine kugelsymmetrische Verteilung ist die Ladungsdichte $\rho_{\text{el.}}(\vec{r}')$ winkelunabhängig, d.h. der Form $Y_{00}(\theta', \varphi')f(r')$. Aus der Orthogonalität der Kugelflächenfunktionen folgt dann, dass jedes Moment $q_{\ell m}$ mit $\ell \neq 0$ Null ist, was sich hier viel schneller als in § VII.3.2 a beweisen lässt.

* Die Entwicklung in Kugelflächenfunktionen (VII.37a) ist zwar eine exakte Gleichung, während die kartesische Multipolentwicklung (VII.27a) eine Näherung für große Abstände darstellt. Die erstere ist in der Praxis aber nur nützlich, falls die Anzahl der signifikanten Multipolmomente endlich (und klein) ist, wodurch die Näherung eines großen Abstands wieder auftaucht.

Literatur zum Kapitel VII

- Fließbach, *Elektrodynamik* [3] Teil II.
- Greiner, *Klassische Elektrodynamik* [8] Teil I.
- Griffiths, *Elektrodynamik* [9] = *Introduction to Electrodynamics* [10] Kap 2.1–2.4 & 3.
- Jackson, *Klassische Elektrodynamik* [11] = *Classical Electrodynamics* [12] Kap. 1.1–1.5, 1.7–1.11, 2.1, 3.5–3.6 & 4.1–4.2.
- Landau & Lifschitz, *Klassische Feldtheorie* [14] = *The classical theory of fields* [28] Kap. 5 § 36–38 & 41–42
- Nolting, *Elektrodynamik* [17] Kap. 2.1–2.3.
- Scheck, *Klassische Feldtheorie* [19] Kap. 1.5 & 1.7.

KAPITEL VIII

Magnetostatik

VIII.1	Grundbegriffe und -ergebnisse der Magnetostatik	159
VIII.1.1	Vektorpotential	160
VIII.1.2	Poisson-Gleichungen der Magnetostatik	161
VIII.1.3	Integrale Formulierung der Grundgleichungen der Magnetostatik	163
VIII.1.4	Magnetisches Feld induziert durch einfache Ladungsströme	165
VIII.1.5	Kraft zwischen zwei Stromkreisen	168
VIII.2	Multipolentwicklung	169
VIII.2.1	Multipolmomente einer Ladungsstromverteilung	169
VIII.2.2	Magnetisches Dipolmoment einer Leiterschleife	172
VIII.2.3	Magnetischer Dipol in einem äußeren magnetischen Feld	172

Die Magnetostatik ist die Lehre der magnetischen Felder, die von zeitlich konstanten elektrischen Strömen herrühren. Im entsprechenden stationären Regime vereinfachen sich die Maxwell–Thomson- und die Maxwell–Ampère-Gleichung (VI.1b), (VI.1d) für die magnetische Flussdichte $\vec{B}(\vec{r})$ — die jetzt zeitunabhängig ist — zu

$$\vec{\nabla} \cdot \vec{B}(\vec{r}) = 0 \quad \text{(VIII.1a)}$$

$$\vec{\nabla} \times \vec{B}(\vec{r}) = \mu_0 \vec{j}_{\text{el.}}(\vec{r}). \quad \text{(VIII.1b)}$$

Somit entkoppeln die Gleichungen für das Magnetfeld von denen für das elektrische Feld — so dass das letztere hiernach ignoriert wird.

Den bewegten Ladungen, welche die Stromdichte $\vec{j}_{\text{el.}}$ bilden, kann auch eine Ladungsdichte $\rho_{\text{el.}}$ zugeordnet werden, die ebenfalls zeitlich konstant ist. Über die Maxwell–Gauß-Gleichung (VI.1a) erzeugt diese Ladungsverteilung ein stationäres elektrisches Feld — es sei denn, es existiert eine zweite, statische Ladungsverteilung,⁽⁶⁹⁾ deren Effekt jenen von $\rho_{\text{el.}}$ kompensiert, so dass das resultierende elektrische Feld Null ist.

In Abschn. VIII.1 werden zunächst einige Folgerungen der zwei Grundgleichungen (VIII.1) der Magnetostatik hergeleitet, und zwar die Existenz eines Vektorpotentials für das Magnetfeld sowie verschiedene Beziehungen zwischen der Ladungsstromdichte $\vec{j}_{\text{el.}}$ und entweder dem magnetischen Feld oder dem Vektorpotential. Dann befasst sich Abschn. VIII.2 mit der Multipolentwicklung der Magnetostatik, wobei meistens die Dipol-Näherung betrachtet wird.

Im ganzen Kapitel werden Analogien mit dem Formalismus der Elektrostatik (Kap. VII) stark benutzt, um einige Ergebnisse ohne Herleitung anzugeben.

VIII.1 Grundbegriffe und -ergebnisse der Magnetostatik

Dieser Abschnitt befasst sich mit einigen elementaren Folgerungen der Grundgleichungen (VIII.1) der Magnetostatik. Ähnlich wie das elektrostatische Feld aus einem Skalarpotential abgeleitet wer-

⁽⁶⁹⁾... die ungefähr gleich $-\rho_{\text{el.}}$ sein soll!

den kann, lässt sich dem magnetischen Feld ein Vektorpotential zuordnen (§ VIII.1.1). Sowohl das Magnetfeld als auch dieses Vektorpotential genügen jeweiligen Poisson-Gleichungen (§ VIII.1.2), deren Lösungen auf \mathbb{R}^3 sich in Analogie mit den Ergebnissen des Abschn. VII.2 schreiben lassen.

Äquivalent zu den „lokalen“ (Differential-)Gleichungen (VIII.1) können „globale“ Formulierungen der Grundgleichungen gefunden werden (§ VIII.1.3). Schließlich werden die Resultate der ersten Paragraphen auf die Berechnung der von verschiedenen einfachen Stromverteilungen herrührenden Magnetfelder angewandt (§ VIII.1.4).

VIII.1.1 Vektorpotential

Aus der Maxwell–Thomson-Gleichung (VIII.1a) folgt die Existenz eines differenzierbaren Vektorfeldes $\vec{A}(\vec{r})$ derart, dass die Beziehung

$$\vec{B}(\vec{r}) = \vec{\nabla} \times \vec{A}(\vec{r}) \quad (\text{VIII.2})$$

in jedem Punkt \vec{r} erfüllt wird. Das Feld $\vec{A}(\vec{r})$ heißt (*magnetisches*) *Vektorpotential*, mit der physikalischen Dimension $[A] = \text{M L T}^{-2} \text{I}^{-1}$.⁽⁷⁰⁾

Falls das magnetische Feld \vec{B} auf \mathbb{R}^3 definiert ist, wird die Existenz von $\vec{A}(\vec{r})$ unten bewiesen, indem es explizit konstruiert wird [Gl. (VIII.11)].

Bemerkung: Das Produkt aus dem Vektorpotential $\vec{A}(\vec{r})$ und einer elektrischen Ladung q hat die physikalische Dimension M L T^{-1} eines Impulses.

Eichfreiheit

Das Vektorpotential $\vec{A}(\vec{r})$, das zu einem gegebenen magnetischen Feld $\vec{B}(\vec{r})$ führt, wird durch Gl. (VIII.2) nicht eindeutig festgestellt. Definiert man nämlich ein Vektorfeld \vec{A}' über

$$\vec{A}'(\vec{r}) \equiv \vec{A}(\vec{r}) + \vec{\nabla}\chi(\vec{r}) \quad (\text{VIII.3a})$$

mit einer beliebigen kontinuierlich differenzierbaren skalaren Funktion χ der Position, so ist die Rotation von \vec{A}'

$$\vec{B}'(\vec{r}) \equiv \vec{\nabla} \times \vec{A}'(\vec{r}) = \vec{\nabla} \times \vec{A}(\vec{r}) + \vec{\nabla} \times \vec{\nabla}\chi(\vec{r}) = \vec{B}(\vec{r}), \quad (\text{VIII.3b})$$

d.h. \vec{A}' ist auch ein geeignetes Vektorpotential für das Feld \vec{B} . Die Mehrdeutigkeit in der Wahl des Vektorpotentials wird als *Eichfreiheit* bezeichnet, und eine bestimmte Wahl als *Eichung*.

Um das Vektorpotential eindeutig — möglicherweise bis auf einen konstanten additiven Vektor — festzulegen, kann man eine zusätzliche Bedingung über $\vec{A}'(\vec{r})$ fordern. Dabei soll diese *Eichbedingung* so gewählt sein, dass einige Gleichungen damit einfacher werden, wie es in § VIII.1.2 c der Fall sein wird.

In der Magnetostatik wird meistens die *Coulomb-Eichung* gewählt, entsprechend der *Coulomb-Eichbedingung*

$$\vec{\nabla} \cdot \vec{A}'(\vec{r}) = 0 \quad (\text{VIII.4})$$

über das Vektorpotential.

Sei $\vec{A}_0(\vec{r})$ ein Vektorpotential für die magnetische Induktion $\vec{B}(\vec{r})$. Falls die Divergenz von \vec{A}_0 Null ist, dann erfüllt es schon die Coulomb-Eichbedingung. Wenn $\vec{\nabla} \cdot \vec{A}_0(\vec{r})$ nicht identisch verschwindet, dann hat die Poisson-Gleichung

$$\Delta\chi(\vec{r}) = -\vec{\nabla} \cdot \vec{A}_0(\vec{r})$$

eine nicht-konstante Lösung $\chi(\vec{r})$ auf \mathbb{R}^3 . Definiert man dann $\vec{A} \equiv \vec{A}_0 + \vec{\nabla}\chi$, so findet man sofort, dass \vec{A} die Bedingung (VIII.4) erfüllt.

⁽⁷⁰⁾Die zugehörige SI-Einheit wird nie benutzt.

VIII.1.2 Poisson-Gleichungen der Magnetostatik

Aus den Grundgleichungen (VIII.1) können Poisson-Gleichungen für die magnetische Induktion $\vec{B}(\vec{r})$ oder das Vektorpotential $\vec{A}(\vec{r})$ hergeleitet werden, wobei sich diese Vektorfelder durch die elektrische Stromdichte $\vec{j}_{\text{el.}}(\vec{r})$ ausdrücken lassen.

VIII.1.2a Poisson-Gleichung für die magnetische Induktion

Die Rotationsbildung der stationären Maxwell–Ampère-Gleichung (VIII.1b) lautet

$$\vec{\nabla} \times [\vec{\nabla} \times \vec{B}(\vec{r})] = \frac{1}{\epsilon_0 c^2} \vec{\nabla} \times \vec{j}_{\text{el.}}(\vec{r}).$$

Dabei kann auf der linken Seite die Identität

$$\vec{\nabla} \times [\vec{\nabla} \times \vec{B}(\vec{r})] = \vec{\nabla} [\vec{\nabla} \cdot \vec{B}(\vec{r})] - \Delta \vec{B}(\vec{r})$$

benutzt werden, wobei der erste Term im rechten Glied wegen der Maxwell–Thomson-Gleichung Null ist. Insgesamt ergibt sich somit

$$\Delta \vec{B}(\vec{r}) = -\mu_0 \vec{\nabla} \times \vec{j}_{\text{el.}}(\vec{r}) \quad (\text{VIII.5})$$

d.h. jede Komponente der magnetischen Induktion genügt der Poisson-Gleichung mit der entsprechenden Komponente von $-\mu_0 \vec{\nabla} \times \vec{j}_{\text{el.}}$ als „Quelle“ auf der rechten Seite der Gleichung. Die Lösung dieser Differentialgleichung wurde in schon Abschn. VII.2 studiert. Ist die elektrische Ladungsstromdichte $\vec{j}_{\text{el.}}$ bekannt auf \mathbb{R}^3 — und wenn deren Rotation schnell genug im Unendlichen nach Null geht, damit das untere Integral definiert ist —, so gilt⁽⁷¹⁾

$$\vec{B}(\vec{r}) = \int \frac{\mu_0}{4\pi} \frac{\vec{\nabla}_{\vec{r}'} \times \vec{j}_{\text{el.}}(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3\vec{r}'. \quad (\text{VIII.6})$$

Dabei bezeichnet $\vec{\nabla}_{\vec{r}'}$ den Gradienten bezüglich der Komponenten von \vec{r}' .

In der Praxis wird das Integral oft auf ein Raumgebiet \mathcal{V} beschränkt, wenn die Randbedingungen am Rand $\partial\mathcal{V}$ von \mathcal{V} nicht weiter präzisiert wurden.

VIII.1.2b Biot–Savart-Gesetz

Unter Einführung eines kartesischen Koordinatensystems mit Basis $\{\vec{e}_i\}$ lässt sich die i -te Komponente der Beziehung (VIII.6) als

$$B^i(\vec{r}) = \frac{\mu_0}{4\pi} \int_{\mathcal{V}} \frac{\epsilon^{ijk}}{|\vec{r} - \vec{r}'|} \frac{\partial j_{\text{el.}}^k(\vec{r}')}{\partial x'^j} d^3\vec{r}'$$

umschreiben, mit Summen über die doppelt auftretenden Indizes j und k . Das Integral auf der rechten Seite kann mithilfe einer partiellen Integration transformiert werden. Dabei wird angenommen, dass das Integrationsvolumen \mathcal{V} groß genug ist, damit die Stromdichte $\vec{j}_{\text{el.}}$ am Rand $\partial\mathcal{V}$ verschwindet. Dann ist der integrierte Term aus der partiellen Integration Null, und es bleibt

$$B^i(\vec{r}) = -\frac{\mu_0}{4\pi} \int_{\mathcal{V}} \epsilon^{ijk} j_{\text{el.}}^k(\vec{r}') \frac{\partial}{\partial x'^j} \left(\frac{1}{|\vec{r} - \vec{r}'|} \right) d^3\vec{r}'.$$

Das Ersetzen von $-\epsilon^{ijk}$ durch ϵ^{ikj} und die Berechnung der Ableitung liefern

$$B^i(\vec{r}) = \frac{\mu_0}{4\pi} \int_{\mathcal{V}} \epsilon^{ikj} j_{\text{el.}}^k(\vec{r}') \frac{x^j - x'^j}{|\vec{r} - \vec{r}'|^3} d^3\vec{r}'$$

⁽⁷¹⁾Der Übergang von der Poisson-Differentialgleichung (VIII.5) zur Lösung (VIII.6) ist ähnlich dem von Gl. (VII.4) zur zugehörigen Lösung (VII.22).

für $i = 1, 2, 3$, das heißt

$$\vec{B}(\vec{r}) = \int_{\mathcal{V}} \frac{\mu_0}{4\pi} \frac{\vec{j}_{\text{el.}}(\vec{r}') \times (\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} d^3\vec{r}'. \quad (\text{VIII.7})$$

Diese Beziehung heißt *Biot^(ah)–Savart^(ai)-Gesetz*.

VIII.1.2c Poisson-Gleichung für das Vektorpotential

Ersetzt man die magnetische Induktion $\vec{B}(\vec{r})$ durch $\vec{\nabla} \times \vec{A}(\vec{r})$ in der stationären Maxwell–Ampère-Gleichung (VIII.1b), so ergibt sich

$$\vec{\nabla} \times [\vec{\nabla} \times \vec{A}(\vec{r})] = \mu_0 \vec{j}_{\text{el.}}(\vec{r}),$$

d.h. unter Berücksichtigung der Formel für die Rotation einer Rotation

$$\Delta \vec{A}(\vec{r}) = -\mu_0 \vec{j}_{\text{el.}}(\vec{r}) + \vec{\nabla} [\vec{\nabla} \cdot \vec{A}(\vec{r})]. \quad (\text{VIII.8})$$

Unter Verwendung der Coulomb-Eichbedingung (VIII.4) vereinfacht sich diese Gleichung erneut zu einer Poisson-Gleichung

$$\Delta \vec{A}(\vec{r}) = -\mu_0 \vec{j}_{\text{el.}}(\vec{r}) \quad \text{in der Coulomb-Eichung.} \quad (\text{VIII.9})$$

Wie oben kann man die Lösung dieser Gleichung direkt schreiben:

$$\vec{A}(\vec{r}) = \int_{\mathcal{V}} \frac{\mu_0}{4\pi} \frac{\vec{j}_{\text{el.}}(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3\vec{r}'. \quad (\text{VIII.10})$$

Falls $\vec{j}_{\text{el.}}$ am Rand des Integrationsgebiets \mathcal{V} verschwindet — wenn es keinen Ladungsstrom im Unendlichen gibt, kann man immer \mathcal{V} groß genug wählen, damit es keinen Strom auf $\partial\mathcal{V}$ gibt —, so erfüllt das Vektorpotential (VIII.10) die Coulomb-Eichbedingung (VIII.4).

Beweis: Bezeichnet man mit $\vec{\nabla}_{\vec{r}}$ bzw. $\vec{\nabla}_{\vec{r}'}$ den Nabla-Operator bezüglich der Koordinaten des Vektors \vec{r} bzw. \vec{r}' , so gilt

$$\vec{\nabla} \cdot \vec{A}(\vec{r}) = \frac{\mu_0}{4\pi} \int_{\mathcal{V}} \vec{\nabla}_{\vec{r}} \cdot \frac{\vec{j}_{\text{el.}}(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3\vec{r}' = \frac{\mu_0}{4\pi} \int_{\mathcal{V}} \vec{j}_{\text{el.}}(\vec{r}') \cdot \vec{\nabla}_{\vec{r}} \left(\frac{1}{|\vec{r} - \vec{r}'|} \right) d^3\vec{r}'.$$

Im Integranden kann man $\vec{\nabla}_{\vec{r}}(1/|\vec{r} - \vec{r}'|) = -\vec{\nabla}_{\vec{r}'}(1/|\vec{r} - \vec{r}'|)$ schreiben. Dann lässt sich das daraus folgende Integral über partielle Integration berechnen, wobei der integrierte Term, proportional zur Ladungsstromdichte ausgewertet am Rand des Integrationsgebiets, Null ist:

$$\vec{\nabla} \cdot \vec{A}(\vec{r}) = -\frac{\mu_0}{4\pi} \int_{\mathcal{V}} \vec{j}_{\text{el.}}(\vec{r}') \cdot \vec{\nabla}_{\vec{r}'} \left(\frac{1}{|\vec{r} - \vec{r}'|} \right) d^3\vec{r}' = \frac{\mu_0}{4\pi} \int_{\mathcal{V}} \frac{1}{|\vec{r} - \vec{r}'|} \vec{\nabla}_{\vec{r}'} \cdot \vec{j}_{\text{el.}}(\vec{r}') d^3\vec{r}'.$$

Aus der stationären Maxwell–Ampère-Gleichung (VIII.1b) folgt sofort $\vec{\nabla}_{\vec{r}'} \cdot \vec{j}_{\text{el.}}(\vec{r}') = 0$, was zu $\vec{\nabla} \cdot \vec{A}(\vec{r}) = 0$ führt. \square

Das Biot–Savart-Gesetz (VIII.7) kann auch aus der Beziehung (VIII.10) wiedergefunden werden. Bildet man deren Rotation, so ergibt sich

$$\vec{\nabla} \times \vec{A}(\vec{r}) = \frac{\mu_0}{4\pi} \int_{\mathcal{V}} \vec{\nabla}_{\vec{r}} \times \left[\frac{\vec{j}_{\text{el.}}(\vec{r}')}{|\vec{r} - \vec{r}'|} \right] d^3\vec{r}'.$$

In einem kartesischen Koordinatensystem lautet die i -te Komponente dieser Gleichung

$$B^i(\vec{r}) = \frac{\mu_0}{4\pi} \int_{\mathcal{V}} \epsilon^{ijk} \frac{\partial}{\partial x^j} \frac{j_{\text{el.}}^k(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3\vec{r}' = \frac{\mu_0}{4\pi} \int_{\mathcal{V}} \epsilon^{ijk} j_{\text{el.}}^k(\vec{r}') \frac{\partial}{\partial x^j} \left(\frac{1}{|\vec{r} - \vec{r}'|} \right) d^3\vec{r}'.$$

^(ah)J.-B. BIOT, 1774–1862 ^(ai)F. SAVART, 1791–1841

Dabei gibt die Ableitung $-(x^j - x'^j)/|\vec{r} - \vec{r}'|^3$. Mit $-\epsilon^{ijk} = \epsilon^{ikj}$ ergibt sich

$$B^i(\vec{r}) = \frac{\mu_0}{4\pi} \int_{\mathcal{V}'} \epsilon^{ikj} j_{\text{el.}}^k(\vec{r}') \frac{x^j - x'^j}{|\vec{r} - \vec{r}'|^3} d^3\vec{r}',$$

entsprechend genau der i -ten Komponente der Gl. (VIII.7).

Bemerkung: Unter Verwendung der stationären Maxwell–Ampère-Gleichung wird Gl. (VIII.10) zu einem Zusammenhang zwischen dem Vektorpotential (in Coulomb-Eichung) und der magnetischen Induktion:

$$\vec{A}(\vec{r}) = \int_{\mathcal{V}'} \frac{1}{4\pi} \frac{\vec{\nabla} \times \vec{B}(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3\vec{r}'. \quad (\text{VIII.11})$$

Dabei handelt es sich um einen Sonderfall der *Helmholtz*^(aj)-Zerlegung, laut welcher jedes Vektorfeld $\vec{V}(\vec{r})$ auf \mathbb{R}^3 , das sich im Unendlichen „gut verhält“ — d.h. derart, dass die Integrale (VIII.12b)–(VIII.12c) existieren⁽⁷²⁾ —, als Summe eines Gradientenfeldes $\vec{V}_{\parallel}(\vec{r})$ und eines Rotationsfeldes $\vec{V}_{\perp}(\vec{r})$ geschrieben werden kann:

$$\vec{V}(\vec{r}) = \vec{V}_{\parallel}(\vec{r}) + \vec{V}_{\perp}(\vec{r}), \quad \text{wobei} \quad (\text{VIII.12a})$$

$$\vec{V}_{\parallel}(\vec{r}) = -\vec{\nabla}\Phi(\vec{r}) \quad \text{mit} \quad \Phi(\vec{r}) \equiv \int_{\mathcal{V}'} \frac{\vec{\nabla}_{\vec{r}'} \cdot \vec{V}(\vec{r}')}{4\pi |\vec{r} - \vec{r}'|} d^3\vec{r}' \quad (\text{VIII.12b})$$

$$\vec{V}_{\perp}(\vec{r}) = \vec{\nabla} \times \vec{A}(\vec{r}) \quad \text{mit} \quad \vec{A}(\vec{r}) \equiv \int_{\mathcal{V}'} \frac{\vec{\nabla}_{\vec{r}'} \times \vec{V}(\vec{r}')}{4\pi |\vec{r} - \vec{r}'|} d^3\vec{r}'. \quad (\text{VIII.12c})$$

Dabei gelten

$$\vec{\nabla} \cdot \vec{V}_{\perp}(\vec{r}) = 0 \quad \text{und} \quad \vec{\nabla} \times \vec{V}_{\parallel}(\vec{r}) = \vec{0}, \quad (\text{VIII.12d})$$

d.h.

$$\vec{\nabla} \cdot \vec{V}(\vec{r}) = \vec{\nabla} \cdot \vec{V}_{\parallel}(\vec{r}) \quad \text{und} \quad \vec{\nabla} \times \vec{V}(\vec{r}) = \vec{\nabla} \times \vec{V}_{\perp}(\vec{r}). \quad (\text{VIII.12e})$$

In der Magnetostatik ist $\vec{V}(\vec{r}) = \vec{B}(\vec{r})$, und $\vec{\nabla} \cdot \vec{B}(\vec{r})$ bzw. $\vec{\nabla} \times \vec{B}(\vec{r})$ wird durch Gl. (VIII.1a) bzw. (VIII.1b) gegeben. Für die Elektrostatik soll man $\vec{V}(\vec{r}) = \vec{E}(\vec{r})$ betrachten, während $\vec{\nabla} \cdot \vec{E}(\vec{r})$ bzw. $\vec{\nabla} \times \vec{E}(\vec{r})$ durch Gl. (VII.1a) bzw. (VII.1b) gegeben sind.

VIII.1.3 Integrale Formulierung der Grundgleichungen der Magnetostatik

In diesem Paragraphen werden die integralen Formulierungen dargestellt, die den lokalen Grundgleichungen (VIII.1) der Magnetostatik entsprechen.

VIII.1.3a Erhaltung des elektrischen Stroms

Die Divergenzbildung der stationären Maxwell–Ampère-Gleichung (VIII.1b) gibt

$$\vec{\nabla} \cdot \vec{j}_{\text{el.}}(\vec{r}) = 0. \quad (\text{VIII.13})$$

Unter Verwendung des Integralsatzes von Gauß ergibt sich dann

$$0 = \int_{\mathcal{V}} \vec{\nabla} \cdot \vec{j}_{\text{el.}}(\vec{r}) d^3\vec{r} = \oint_{\partial\mathcal{V}} \vec{j}_{\text{el.}}(\vec{r}) \cdot d^2\vec{S} \quad (\text{VIII.14})$$

d.h. der resultierende Ladungsstrom durch die Oberfläche $\partial\mathcal{V}$ eines Volumens \mathcal{V} verschwindet: der einströmende Ladungsstrom muss wieder herausströmen, entsprechend der Abwesenheit von Quellen des elektrischen Stroms.

Bemerkung: Dementsprechend wird ein allgemeines Vektorfeld $\vec{V}(\vec{r})$, dessen Divergenz überall verschwindet — wie z.B. die magnetische Induktion $\vec{B}(\vec{r})$ —, als *quellfrei* bezeichnet.

⁽⁷²⁾In Lehrbüchern findet man oft die Forderung, dass der Betrag $|\vec{V}(\vec{r})|$ schneller als $1/r$ für $r \rightarrow \infty$ abnehmen soll. Mathematisch gilt das Ergebnis auch mit schwächeren Annahmen über \vec{V} , vgl. z.B. Ref. [29] für eine Diskussion.

^(aj)H. VON HELMHOLTZ, 1821–1894

Aus dieser Erhaltung des Ladungsstroms folgt auch, dass die elektrische Stromstärke I durch den Querschnitt eines leitenden Drahts konstant entlang des Drahts bleibt.

Eine zweite Folgerung der Ladungsstromerhaltung ist die *Kirchhoffsche*^(ak) *Knotenregel*,⁽⁷³⁾ laut der die Summe der an einem Leiterknoten zufließenden elektrischen Ströme gleich der Summe der daraus abfließenden Ströme ist.

Beweis: Betrachtet man ein Volumen \mathcal{V} um den Knoten, so ist der nach außen gerichtete Fluss der Stromdichte $\vec{j}_{\text{el.}}$ durch die Fläche $\partial\mathcal{V}$ gleich der algebraischen Summe der Stromstärken in den Leitern, wobei aus- bzw. einfließende Ströme positiv bzw. negativ gezählt werden. Für das Beispiel der Abb. VIII.1

$$\oint_{\partial\mathcal{V}} \vec{j}_{\text{el.}}(\vec{r}) \cdot d^2\vec{S} = I_1 - I_2 - I_3 - I_4 + I_5 = 0.$$

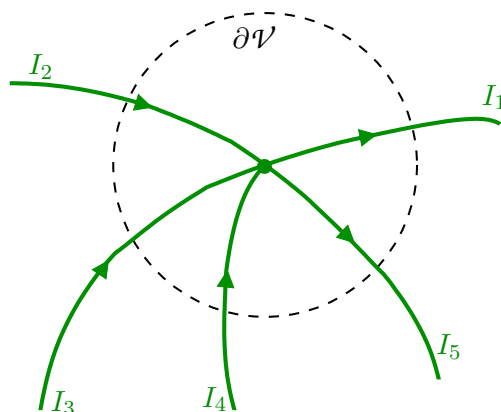


Abbildung VIII.1 – Knotenregel

VIII.1.3 b Fluss der magnetischen Induktion

Sei $\partial\mathcal{V}$ eine geschlossene Fläche, die ein Gebiet \mathcal{V} einschließt. Der magnetische Fluss durch die Oberfläche $\partial\mathcal{V}$ wird durch

$$\Phi_B \equiv \oint_{\partial\mathcal{V}} \vec{B}(\vec{r}) \cdot d^2\vec{S} \quad (\text{VIII.15})$$

definiert. Dabei ist $d^2\vec{S} = d^2S \vec{e}_n$ der vektorielle Oberflächenelement, wobei \vec{e}_n den nach außen ausgerichteten Normaleinheitsvektor zu d^2S bezeichnet.

Unter Verwendung des Integralsatzes von Gauß ist das Oberflächenintegral auf der rechten Seite gleich dem Volumenintegral der Divergenz des Integranden, d.h.

$$\oint_{\partial\mathcal{V}} \vec{B}(\vec{r}) \cdot d^2\vec{S} = \int_{\mathcal{V}} \vec{\nabla} \cdot \vec{B}(\vec{r}) d^3\vec{r}.$$

Wegen der Maxwell–Thomson-Gleichung (VIII.1a) ist die rechte Seite dieser Gleichung immer Null, unabhängig von \mathcal{V}

$$\Phi_B = 0. \quad (\text{VIII.16})$$

Das heißt, der magnetische Fluss durch jede geschlossene Fläche ist immer Null.

VIII.1.3 c Ampère-Gesetz

Sei \mathcal{S} eine Fläche und $\partial\mathcal{S}$ deren Rand. Das Flächenintegral über \mathcal{S} der stationären Maxwell–Ampère-Gleichung (VIII.1b) lautet

$$\int_{\mathcal{S}} \vec{\nabla} \times \vec{B}(\vec{r}) \cdot d^2\vec{S} = \mu_0 \int_{\mathcal{S}} \vec{j}_{\text{el.}}(\vec{r}) \cdot d^2\vec{S}.$$

⁽⁷³⁾ ... auch bekannt als *erstes Kirchhoffsches Gesetz*.

^(ak) G. KIRCHHOFF, 1824–1887

Dabei ist das Integral auf der rechten Seite genau gleich der Stärke I des elektrischen Stroms durch \mathcal{S} . Wiederum lässt sich auf der linken Seite das Flächenintegral einer Rotation über \mathcal{S} mit dem Integralsatz von Stokes in ein Linienintegral entlang des geschlossenen Rands $\partial\mathcal{S}$ transformieren. Somit ergibt sich das *Ampère-Gesetz*

$$\oint_{\partial\mathcal{S}} \vec{B}(\vec{r}) \cdot d\vec{\ell} = \mu_0 \int_{\mathcal{S}} \vec{j}_{\text{el.}}(\vec{r}) \cdot d^2\vec{S} = \mu_0 I. \quad (\text{VIII.17})$$

Bemerkung: Das Linienintegral eines Vektorfeldes $\vec{V}(\vec{r})$ entlang einer geschlossenen Kurve \mathcal{C} wird *Zirkulation* des Feldes längs \mathcal{C} genannt. Dementsprechend tritt auf der linken Seite des Ampère-Gesetzes (VIII.17) die Zirkulation des Magnetfeldes längs des Rands $\partial\mathcal{S}$.

Dann besagt das Ampère-Gesetz, dass die Zirkulation der magnetischen Induktion längs einer geschlossenen Kurve $\partial\mathcal{S}$ gleich μ_0 mal dem elektrischen Strom durch eine von $\partial\mathcal{S}$ aufgespannte Fläche ist.

VIII.1.4 Magnetisches Feld induziert durch einfache Ladungsströme

Im Fall eines Ladungsstroms durch einen dünnen Draht vereinfacht sich das Integral im Biot-Savart-Gesetz (VIII.7) oder in den äquivalenten Formeln (VIII.6) und (VIII.10). Betrachtet man nämlich ein infinitesimales Volumenelement $d^3\vec{r}'$ um einen Teil des Drahts, so lässt sich dieses als $d^2\mathcal{S}' d\ell'$ schreiben, mit $d\ell'$ der Länge von $d^3\vec{r}'$ entlang des Drahts und $d^2\mathcal{S}'$ die Fläche des Querschnitts von $d^3\vec{r}'$ senkrecht dazu.

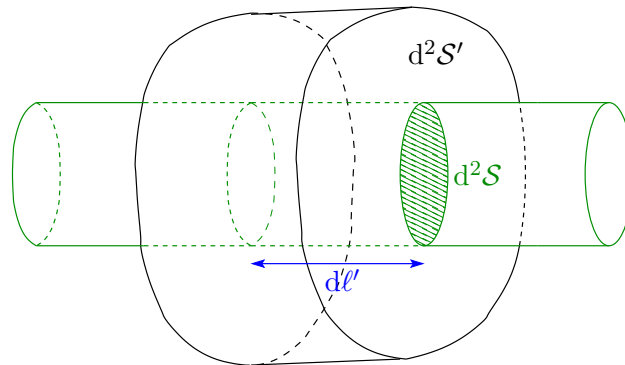


Abbildung VIII.2

Sei \vec{e}_{\parallel} der Einheitsvektor in Drahtrichtung — d.h. kollinear zur Stromdichte $\vec{j}_{\text{el.}}(\vec{r}')$ —, sowie Vektoren $d\vec{\ell}' \equiv d\ell' \vec{e}_{\parallel}$ und $d^2\vec{S}' \equiv d^2\mathcal{S}' \vec{e}_{\parallel}$, im Einklang mit der üblichen Definition des vektoriellen Flächenelements. Dann gilt

$$\vec{j}_{\text{el.}}(\vec{r}') d^3\vec{r}' = \vec{j}_{\text{el.}}(\vec{r}') d^2\mathcal{S}' d\ell' = [\vec{j}_{\text{el.}}(\vec{r}') \cdot d^2\vec{S}'] d\vec{\ell}'.$$

Daraus folgt

$$\int \vec{j}_{\text{el.}}(\vec{r}') d^3\vec{r}' = \int \left(\int [\vec{j}_{\text{el.}}(\vec{r}') \cdot d^2\vec{S}'] \right) d\vec{\ell}'.$$

Zum Flächenintegral trägt nur der in Abb. (VIII.2) Querschnitt $d^2\mathcal{S}$ bei, was die Stromstärke I des Ladungsstroms ergibt. Insgesamt können Volumenintegrale durch Linienintegrale entlang des Drahts ersetzt werden, mit der Substitution

$$\vec{j}_{\text{el.}}(\vec{r}') d^3\vec{r}' = I d\vec{\ell}' \quad (\text{VIII.18})$$

im Integranden, wie z.B. für das Biot-Savart-Gesetz (VIII.7) in Gl. (VIII.21) gemacht wird.

Wem diese Substitution nicht gefällt, kann stattdessen die Ladungsstromdichte durch den Draht mithilfe einer zweidimensionalen Delta-Distribution schreiben:

$$\vec{j}_{\text{el.}}(\vec{r}) = I \delta(x^1) \delta(x^2) \vec{e}_3,$$

wobei ein (lokales) Koordinatensystem gewählt wurde, in dem der unendlich dünne Draht entlang der x^3 -Achse liegt. Nach Einsetzen in ein Integral ergibt sich das gleiche wie mit der Substitution (VIII.18).

VIII.1.4 a Magnetisches Feld eines dünnen geradlinigen Draht

Durch einen unendlich langen dünnen geradlinigen elektrischen Leiter fließt ein stationärer Ladungsstrom I . Dieser erzeugt in jedem Punkt \vec{r} außerhalb des Drahts eine magnetische Induktion $\vec{B}(\vec{r})$.

Wegen der Zylindersymmetrie um den Draht ist es sinnvoll, Zylinderkoordinaten (r, θ, z) mit den zugehörigen Basisvektoren $(\vec{e}_r, \vec{e}_\theta, \vec{e}_z)$ einzuführen. Dabei zeigt in jedem Punkt \vec{e}_r in die Radialrichtung weg vom Draht, während \vec{e}_θ orthogonal dazu ist.

Wegen der Zylindersymmetrie kann der Betrag des Magnetfeldes in einem Punkt nur von dessen Abstand r zur Achse des Leiters abhängen, sei $B(r)$.

Sei \mathcal{C} ein Kreis mit Radius R in einer Ebene senkrecht zum Draht, wobei der letztere durch das Zentrum des Kreises durchläuft; dann ist $|\vec{B}(\vec{r})|$ konstant entlang \mathcal{C} . Wegen der Symmetrie der Geometrie unter Spiegelungen bezüglich der Ebene von \mathcal{C} muss $\vec{B}(\vec{r})$ in einem Punkt \vec{r} des Kreises in der Ebene liegen; genauer ist $\vec{B}(\vec{r})$ tangential zum Kreis, denn es orthogonal zum Abstandsvektor zur Drahtachse ist. Die genaue Richtung, und zwar mit einer positiven Komponente entlang \vec{e}_θ , folgt aus der Rechte-Hand-Regel: $\vec{B}(\vec{r}) = B(R) \vec{e}_\theta$ für $\vec{r} \in \mathcal{C}$.

Da $\vec{B}(\vec{r})$ in jedem Punkt \vec{r} des Kreises tangential dazu ist, lautet die Zirkulation der magnetischen Induktion entlang \mathcal{C}

$$\oint_{\mathcal{C}} \vec{B}(\vec{r}) \cdot d\vec{\ell} = \oint_{\mathcal{C}} |\vec{B}(\vec{r})| d\ell = 2\pi R B(R),$$

wobei die zweite Gleichung den konstanten Wert von $|\vec{B}(\vec{r})|$ auf \mathcal{C} berücksichtigt. Laut dem Ampère-Gesetz ist diese Zirkulation gleich $\mu_0 I$, woraus sich $B(R) = \mu_0 I / 2\pi R$ und somit

$$\vec{B}(\vec{r}) = \frac{\mu_0 I}{2\pi r} \vec{e}_\theta \quad \text{für } r > 0 \quad (\text{VIII.19})$$

ergibt.

Man prüft schnell nach, dass dieses Magnetfeld aus dem Vektorpotential

$$\vec{A}(\vec{r}) = -\frac{\mu_0 I}{2\pi} \ln\left(\frac{r}{r_0}\right) \vec{e}_z \quad \text{für } r > 0 \quad (\text{VIII.20})$$

abgeleitet werden kann. Dabei bezeichnet r_0 eine beliebige positive Zahl, die eingeführt wurde, damit das Argument des Logarithmus dimensionslos ist. Dazu erfüllt dieses Vektorpotential die Coulomb-Eichbedingung (VIII.4).

Bemerkung: Anhand einer ähnlicher Herangehensweise kann man zeigen, dass ein von einer gleichförmigen Ladungsstromdichte $\vec{j}_{\text{el.}}(\vec{r}) = (I/\pi a^2)\Theta(r-a)\vec{e}_z$ durchflossener geradliniger Draht mit Radius a die magnetische Induktion

$$\vec{B}(\vec{r}) = \begin{cases} \frac{\mu_0 I r}{2\pi a^2} \vec{e}_\theta & \text{für } r \leq a \\ \frac{\mu_0 I}{2\pi r} \vec{e}_\theta & \text{für } r \geq a \end{cases}$$

erzeugt.

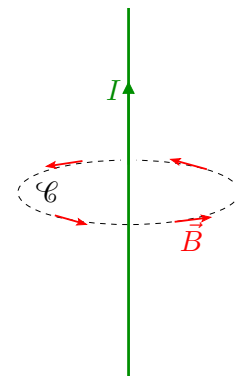


Abbildung VIII.3

VIII.1.4 b Magnetisches Feld einer Leiterschleife

Als Ladungsstrom wird jetzt ein stationärer Strom I durch eine kreisförmige Leiterschleife \mathcal{C} mit Radius R betrachtet. Gesucht ist die magnetische Induktion $\vec{B}(\vec{r})$ in einem Punkt P auf der Achse des Kreises. Sei O das Zentrum von \mathcal{C} und z der Abstand zwischen O und P .

Laut dem Biot-Savart-Gesetz (VIII.7) mit der Substitution (VIII.18) ist das durch die Leiterschleife erzeugte magnetische Feld durch

$$\vec{B}(\vec{r}) = \oint_{\mathcal{C}} \frac{\mu_0 I d\vec{\ell}' \times (\vec{r} - \vec{r}')}{4\pi |\vec{r} - \vec{r}'|^3} \quad (\text{VIII.21})$$

gegeben. Dabei ist \vec{r}' der Ortsvektor des Linienelements $d\vec{\ell}'$ entlang des Leiters.

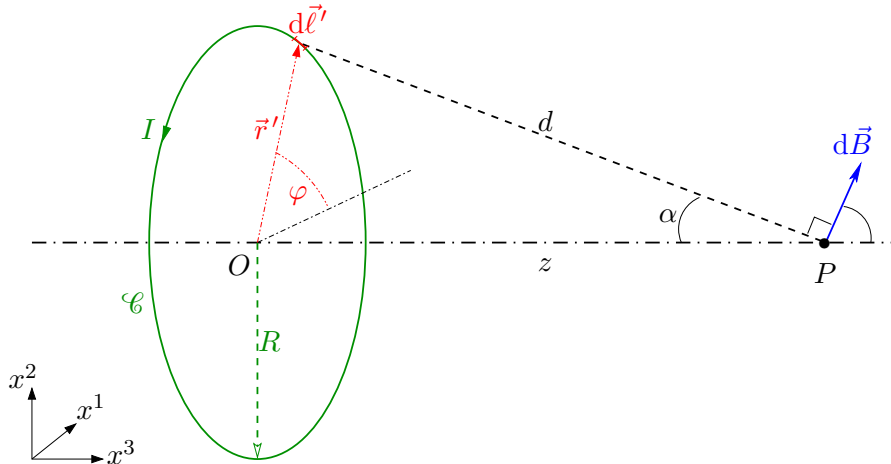


Abbildung VIII.4

Ein solches Leiterelement erzeugt eine infinitesimale magnetische Induktion $d\vec{B}$, die durch das Integrand des Kurvenintegrals gegeben ist. Dieses Feld ist senkrecht zum Abstandsvektor $\vec{r} - \vec{r}'$ von $d\ell'$ bis zum Punkt P , wie in Abb. VIII.4 dargestellt wird. Da das Problem Zylindersymmetrie um die Achse des Kreises besitzt, muss das durch den ganzen Stromkreis erzeugte Feld $\vec{B}(z)$ parallel dieser Achse sein. Dementsprechend wird hiernach nur die Komponente der Gl. (VIII.21) entlang der Kreisachse betrachtet.

Unabhängig von der Position des Linienelements $d\vec{\ell}'$ entlang \mathcal{C} bleibt dessen Abstand $|\vec{r} - \vec{r}'|$ zum Punkt P konstant, sei d , während der Abstandsvektor $\vec{r} - \vec{r}'$ und $d\vec{\ell}'$ immer orthogonal zu einander sind. Somit ist der Betrag des Integranden von Gl. (VIII.21) konstant gleich $\mu_0 I d\ell' / 4\pi d^2$.

Gleichfalls ist der Winkel zwischen dem von dem durch $d\vec{\ell}'$ fließenden Strom herrührenden Magnetfeld $d\vec{B}$ und der Kreisachse auch unabhängig von der Position von $d\vec{\ell}'$ entlang \mathcal{C} , und zwar gleich $\pi/2 - \alpha$, wobei α in Abb. VIII.4 definiert ist. Daraus folgert man für den Betrag von \vec{B}

$$|\vec{B}(z)| = \oint_{\mathcal{C}} \frac{\mu_0 I d\ell'}{4\pi d^2} \cos\left(\frac{\pi}{2} - \alpha\right) = \frac{\mu_0 I}{4\pi d^2} \cos\left(\frac{\pi}{2} - \alpha\right) \oint_{\mathcal{C}} d\ell' = \frac{\mu_0 I}{4\pi d^2} \cos\left(\frac{\pi}{2} - \alpha\right) 2\pi R.$$

Dabei gilt $\cos(\pi/2 - \alpha) = \sin \alpha = R/d$, mit $d = \sqrt{R^2 + z^2}$. Insgesamt kommt

$$|\vec{B}(z)| = \frac{\mu_0 I}{2} \frac{R^2}{(R^2 + z^2)^{3/2}} \quad (\text{VIII.22})$$

und \vec{B} ist gerichtet nach rechts in Abb. VIII.4.

Alternativ kann man ein kartesisches Koordinatensystem mit dem Nullpunkt im Zentrum der Kreisschleife einführen (vgl. Abb. VIII.4) und damit arbeiten.⁽⁷⁴⁾ Parametrisiert man die Posi-

⁽⁷⁴⁾ Auf erster Sicht können kartesische Koordinaten für ein Problem mit Zylindersymmetrie zwar überraschend wirken. Zur Berechnung des Kreuzprodukts im Integranden des Biot-Savart-Gesetzes sind sie aber die einfachsten.

tion eines Punkts von \mathcal{C} mit einem Winkel φ , so gelten

$$\vec{r} = \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix} \quad \text{und} \quad \vec{r}' = \begin{pmatrix} R \cos \varphi \\ R \sin \varphi \\ 0 \end{pmatrix} \quad \text{d.h.} \quad \vec{r} - \vec{r}' = \begin{pmatrix} -R \cos \varphi \\ -R \sin \varphi \\ z \end{pmatrix}.$$

Dazu kann man bemerken, dass das Linienelement $d\vec{\ell}'$ tangential zum Kreis genau die infinitesimale Variation von \vec{r}' darstellt; daher ist

$$d\vec{\ell}' = \frac{d\vec{r}'}{d\varphi} d\varphi \quad \text{mit} \quad \frac{d\vec{r}'}{d\varphi} = \begin{pmatrix} -R \sin \varphi \\ R \cos \varphi \\ 0 \end{pmatrix}$$

und das Kurvenintegral entlang \mathcal{C} wird zu einem Integral über $\varphi \in [0, 2\pi]$. Daraus folgt

$$d\vec{\ell}' \times (\vec{r} - \vec{r}') = \begin{pmatrix} zR \cos \varphi \\ zR \sin \varphi \\ R^2(\cos^2 \varphi + \sin^2 \varphi) \end{pmatrix} d\varphi.$$

Das Einsetzen in die Biot–Savart-Formel (VIII.21) unter Berücksichtigung von $|\vec{r} - \vec{r}'| = \sqrt{R^2 + z^2}$ ergibt schließlich

$$\vec{B}(\vec{r}) = \frac{\mu_0}{4\pi} \int_0^{2\pi} \frac{I}{(R^2 + z^2)^{3/2}} \begin{pmatrix} zR \cos \varphi \\ zR \sin \varphi \\ R^2 \end{pmatrix} d\varphi = \frac{\mu_0 I}{2} \frac{R^2}{(R^2 + z^2)^{3/2}} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

im Einklang mit dem schon oben gefundenen Ergebnis (VIII.22).

VIII.1.4 c Magnetisches Feld einer Zylinderspule

... **Aufgabe 78!**

VIII.1.5 Kraft zwischen zwei Stromkreisen

In Abwesenheit eines elektrischen Feldes lautet die Lorentz-Kraftdichte auf eine Ladungsstromdichte \vec{j}_{el} in einem magnetischen Feld

$$\vec{f}_{\text{L}}(\vec{r}) = \vec{j}_{\text{el}}(\vec{r}) \times \vec{B}(\vec{r}). \quad (\text{VIII.23})$$

Ausgehend aus dieser Formel kann man die Kraft $\vec{F}_{a \rightarrow b}$ bestimmen, die ein erster Stromkreis a mit Stromdichte $\vec{j}_{\text{el},a}$ auf einem zweiten Stromkreis b mit Stromdichte $\vec{j}_{\text{el},b}$ ausübt.

Sei \mathcal{V}_a bzw. \mathcal{V}_b ein Gebiet, das den Stromkreis a bzw. b beinhaltet, wobei $\vec{j}_{\text{el},i}$ in jedem Punkt des Rands von \mathcal{V}_i verschwindet. Die Stromdichte $\vec{j}_{\text{el},a}$ erzeugt eine magnetische Induktion \vec{B}_a im Raum, insbesondere in jedem Punkt $\vec{r}_b \in \mathcal{V}_b$. Dadurch erfährt der ganze Stromkreis b eine resultierende Kraft

$$\vec{F}_{a \rightarrow b} = \int_{\mathcal{V}_b} \vec{j}_{\text{el},b}(\vec{r}_b) \times \vec{B}_a(\vec{r}_b) d^3\vec{r}_b.$$

Dabei kann $\vec{B}_a(\vec{r}_b)$ durch das Biot–Savart-Gesetz (VIII.7) ausgedrückt werden:

$$\vec{F}_{a \rightarrow b} = \frac{\mu_0}{4\pi} \int_{\mathcal{V}_b} \vec{j}_{\text{el},b}(\vec{r}_b) \times \left[\int_{\mathcal{V}_a} \frac{\vec{j}_{\text{el},a}(\vec{r}_a) \times (\vec{r}_b - \vec{r}_a)}{|\vec{r}_b - \vec{r}_a|^3} d^3\vec{r}_a \right] d^3\vec{r}_b.$$

Diese Formel lässt sich mit Hilfe der Identität

$$\frac{\vec{r}_b - \vec{r}_a}{|\vec{r}_b - \vec{r}_a|^3} = -\vec{\nabla}_b \left(\frac{1}{|\vec{r}_b - \vec{r}_a|} \right)$$

umschreiben:

$$\vec{F}_{a \rightarrow b} = \frac{\mu_0}{4\pi} \int_{\mathcal{V}_a} \int_{\mathcal{V}_b} \vec{j}_{\text{el},b}(\vec{r}_b) \times \left[\vec{\nabla}_b \left(\frac{1}{|\vec{r}_b - \vec{r}_a|} \right) \times \vec{j}_{\text{el},a}(\vec{r}_a) \right] d^3\vec{r}_b d^3\vec{r}_a.$$

Unter Verwendung des doppelten Kreuzprodukts $\vec{a} \times (\vec{b} \times \vec{c}) = (\vec{a} \cdot \vec{c})\vec{b} - (\vec{a} \cdot \vec{b})\vec{c}$ kommt

$$\begin{aligned} \vec{F}_{a \rightarrow b} &= \frac{\mu_0}{4\pi} \int_{\mathcal{V}_a} \int_{\mathcal{V}_b} [\vec{j}_{\text{el.},a}(\vec{r}_a) \cdot \vec{j}_{\text{el.},b}(\vec{r}_b)] \vec{\nabla}_b \left(\frac{1}{|\vec{r}_b - \vec{r}_a|} \right) d^3\vec{r}_b d^3\vec{r}_a \\ &\quad - \frac{\mu_0}{4\pi} \int_{\mathcal{V}_a} \int_{\mathcal{V}_b} \left[\vec{j}_{\text{el.},b}(\vec{r}_b) \cdot \vec{\nabla}_b \left(\frac{1}{|\vec{r}_b - \vec{r}_a|} \right) \right] \vec{j}_{\text{el.},a}(\vec{r}_a) d^3\vec{r}_b d^3\vec{r}_a. \end{aligned} \quad (\text{VIII.24})$$

Wie wir jetzt beweisen werden, ist der Term in der zweiten Zeile Null, denn das darin enthaltene Integral über \vec{r}_b verschwindet. In der Tat gilt unter Einführung der Koordinaten $\{x_b^k\}$ von \vec{r}_b

$$\begin{aligned} \int_{\mathcal{V}_b} \vec{j}_{\text{el.},b}(\vec{r}) \cdot \vec{\nabla}_{\vec{r}_b} \left(\frac{1}{|\vec{r}_b - \vec{r}_a|} \right) d^3\vec{r}_b &= \int_{\mathcal{V}_b} \sum_{k=1}^3 j_{\text{el.},b}^k(\vec{r}_b) \frac{\partial}{\partial x_b^k} \left(\frac{1}{|\vec{r}_b - \vec{r}_a|} \right) d^3\vec{r}_b \\ &= \int_{\mathcal{V}_b} \sum_{k=1}^3 \frac{\partial}{\partial x_b^k} \left[\frac{j_{\text{el.},b}^k(\vec{r}_b)}{|\vec{r}_b - \vec{r}_a|} \right] d^3\vec{r}_b - \int_{\mathcal{V}_b} \frac{1}{|\vec{r}_b - \vec{r}_a|} \sum_{k=1}^3 \frac{\partial j_{\text{el.},b}^k(\vec{r}_b)}{\partial x_b^k} d^3\vec{r}_b. \end{aligned}$$

Dabei erkennt man die Divergenz zweier Vektorfelder:

$$\int_{\mathcal{V}_b} \vec{j}_{\text{el.},b}(\vec{r}_b) \cdot \vec{\nabla}_{\vec{r}_b} \left(\frac{1}{|\vec{r}_b - \vec{r}_a|} \right) d^3\vec{r}_b = \int_{\mathcal{V}_b} \vec{\nabla}_b \cdot \left[\frac{\vec{j}_{\text{el.},b}(\vec{r}_b)}{|\vec{r}_b - \vec{r}_a|} \right] d^3\vec{r}_b - \int_{\mathcal{V}_b} \frac{1}{|\vec{r}_b - \vec{r}_a|} \vec{\nabla}_b \cdot \vec{j}_{\text{el.},b}(\vec{r}_b) d^3\vec{r}_b.$$

Das zweite Integral auf der rechten Seite verschwindet wegen der Quelfreiheit (VIII.13) der Ladungsstromdichte $\vec{j}_{\text{el.},b}$. Wiederum lässt sich das erste Integral mit dem Satz von Gauß transformieren: daraus kommt das Flächenintegral von $\vec{j}_{\text{el.},b}(\vec{r}_b)/|\vec{r}_b - \vec{r}_a|$ über den Rand $\partial\mathcal{V}_b$, wo $\vec{j}_{\text{el.},b}$ Null ist, so dass auch dieser Beitrag verschwindet.

Insgesamt bleibt somit die Kraft

$$\vec{F}_{a \rightarrow b} = \frac{\mu_0}{4\pi} \int_{\mathcal{V}_a} \int_{\mathcal{V}_b} [\vec{j}_{\text{el.},a}(\vec{r}_a) \cdot \vec{j}_{\text{el.},b}(\vec{r}_b)] \vec{\nabla}_b \left(\frac{1}{|\vec{r}_b - \vec{r}_a|} \right) d^3\vec{r}_b d^3\vec{r}_a. \quad (\text{VIII.25})$$

Da der Gradient von $1/|\vec{r}_b - \vec{r}_a|$ bezüglich \vec{r}_a das Negative des Gradienten bezüglich \vec{r}_b ist, kommt ein globales Minus-Zeichen im Austausch der Rollen von a und b , und zwar $\vec{F}_{b \rightarrow a} = -\vec{F}_{a \rightarrow b}$. Dies ist genau das dritte newtonsche Gesetz (I.19).

VIII.2 Multipolentwicklung

Ähnlich der in Abschn. VII.3 studierten Entwicklung des elektrostatischen Skalarpotentials $\Phi(\vec{r})$ einer Ladungsverteilung $\rho_{\text{el.}}$ als Summe der durch sukzessive Multipolmomente erzeugten Potentiale, wobei das 2^ℓ -te Moment einen Beitrag proportional zu $1/r^{\ell+1}$ liefert, kann das Vektorpotential $\vec{A}(\vec{r})$ einer Ladungsstromverteilung $\vec{j}_{\text{el.}}$ ebenfalls als Summe der Effekte von Multipolmomenten geschrieben werden.

VIII.2.1 Multipolmomente einer Ladungsstromverteilung

Es wird angenommen, dass die Stromverteilung innerhalb eines Volumens \mathcal{V} in der Umgebung des Ursprungs des Bezugssystems lokalisiert ist, wobei $\vec{j}_{\text{el.}}$ in jedem Punkt des Rands $\partial\mathcal{V}$ von \mathcal{V} Null ist. Das magnetische Feld wird in einem weit entfernten Punkt \vec{r} gesucht, d.h. $|\vec{r} - \vec{r}'| \gg R$ für jeden $\vec{r}' \in \mathcal{V}$, wobei R eine typische Länge für das Volumen \mathcal{V} ist.

Gemäß der Gl. (VIII.9) lautet das Vektorpotential (in Coulomb-Eichung)

$$\vec{A}(\vec{r}) = \frac{\mu_0}{4\pi} \int_{\mathcal{V}} \frac{\vec{j}_{\text{el.}}(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3\vec{r}'.$$

Dank den Annahmen ist der Abstand $|\vec{r} - \vec{r}'|$ im Nenner des Integranden ungefähr gleich $r \equiv |\vec{r}|$. Somit kann der Bruch $1/|\vec{r} - \vec{r}'|$ mithilfe einer Taylor-Entwicklung angenähert werden. Zur ersten

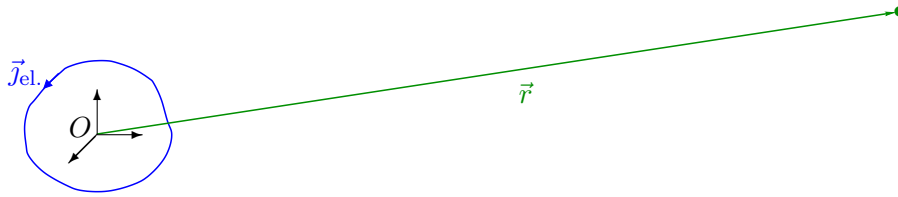


Abbildung VIII.5

Ordnung in \vec{r}' gilt

$$\frac{1}{|\vec{r} - \vec{r}'|} = \frac{1}{r} + \frac{\vec{r} \cdot \vec{r}'}{r^3} + \mathcal{O}\left(\frac{1}{r^3}\right) = \frac{1}{r} + \sum_{k=1}^3 \frac{x^k}{r^3} x'^k + \mathcal{O}\left(\frac{1}{r^3}\right),$$

wobei die $\{x^k\}$ bzw. $\{x'^k\}$ die kartesischen Koordinaten von \vec{r} bzw. \vec{r}' in einem System mit Basisvektoren $\{\vec{e}_k\}$ sind. Das Einsetzen dieser Entwicklung in die Formel für das Vektorpotential gibt

$$\vec{A}(\vec{r}) = \frac{\mu_0}{4\pi r} \int_{\mathcal{V}} \vec{j}_{\text{el.}}(\vec{r}') d^3\vec{r}' + \frac{\mu_0}{4\pi r^3} \sum_{k=1}^3 x^k \int_{\mathcal{V}} x'^k \vec{j}_{\text{el.}}(\vec{r}') d^3\vec{r}' + \dots \quad (\text{VIII.26})$$

Die Integrale in dieser Entwicklung des Vektorpotentials in Potenzen von $1/r$ sind die (kartesischen) *magnetischen Multipolmomente* der Ladungsstromverteilung. Wie wir jetzt sehen werden ist der erste Term in der Tat Null, so dass der führende Beitrag jener des Dipolmoments ist.

Lemma: Sei f bzw. \vec{J} eine kontinuierlich differenzierbare skalare Funktion bzw. ein quellfreies Vektorfeld auf einem Bereich \mathcal{V} , wobei \vec{J} in jedem Punkt des Rands $\partial\mathcal{V}$ von \mathcal{V} verschwindet. Dann gilt

$$\int_{\mathcal{V}} \vec{J}(\vec{r}') \cdot \vec{\nabla} f(\vec{r}') d^3\vec{r}' = 0. \quad (\text{VIII.27})$$

Beweis: Dank der Annahme $\vec{J}(\vec{r}') = \vec{0}$ in jedem Punkt $\vec{r}' \in \partial\mathcal{V}$ gilt

$$\oint_{\partial\mathcal{V}} f(\vec{r}') \vec{J}(\vec{r}') \cdot d^2\vec{S} = 0.$$

Die Transformation des Oberflächenintegrals mit dem Integralsatz von Gauß ergibt unter Verwendung der Produktregel

$$0 = \int_{\mathcal{V}} \vec{\nabla} \cdot [f(\vec{r}') \vec{J}(\vec{r}')] d^3\vec{r}' = \int_{\mathcal{V}} [f(\vec{r}') \vec{\nabla} \cdot \vec{J}(\vec{r}') + \vec{J}(\vec{r}') \cdot \vec{\nabla} f(\vec{r}')] d^3\vec{r}'.$$

Der erste Term in den eckigen Klammern verschwindet dank der Quellfreiheit von \vec{J} , woraus das gesuchte Ergebnis folgt. \square

Wendet man dieses Lemma mit $\vec{J} = \vec{j}_{\text{el.}}$ und $f(\vec{r}') = x'^k$ an, wobei eine triviale Berechnung $\vec{\nabla} f(\vec{r}') = \vec{e}_k$ ergibt, so kommt für jedes $k = 1, 2, 3$

$$\int_{\mathcal{V}} \vec{j}_{\text{el.}}(\vec{r}') \cdot \vec{\nabla} f(\vec{r}') d^3\vec{r}' = \int_{\mathcal{V}} j_{\text{el.}}^k(\vec{r}') d^3\vec{r}' = 0,$$

d.h. insgesamt

$$\int_{\mathcal{V}} \vec{j}_{\text{el.}}(\vec{r}') d^3\vec{r}' = \vec{0}. \quad (\text{VIII.28})$$

Somit ist der erste Beitrag auf der rechten Seite der Multipolentwicklung (VIII.26) Null. Das heißt, es gibt keine „magnetische Monopole“ — entsprechend der Abwesenheit eines Quellterms auf der rechten Seite der Maxwell–Thomson-Gleichung (VIII.1a).

Sei jetzt $f(\vec{r}') = x'^k x'^l$ mit $k, l \in \{1, 2, 3\}$; dann ist $\vec{\nabla} f(\vec{r}') = x'^l \vec{e}_k + x'^k \vec{e}_l$ und das Lemma (VIII.27) gibt

$$\int_{\mathcal{V}} [x'^l j_{\text{el.}}^k(\vec{r}') + x'^k j_{\text{el.}}^l(\vec{r}')] d^3\vec{r}' = 0.$$

Diese Gleichung kann verwendet werden, um das Integral des zweiten Terms in Gl. (VIII.26) umzuschreiben:

$$\begin{aligned} \int_{\mathcal{V}} x'^k \vec{j}_{\text{el.}}(\vec{r}') d^3\vec{r}' &= \sum_{l=1}^3 \vec{e}_l \int_{\mathcal{V}} x'^k j_{\text{el.}}^l(\vec{r}') d^3\vec{r}' \\ &= \sum_{l=1}^3 \vec{e}_l \left\{ \frac{1}{2} \int_{\mathcal{V}} [x'^k j_{\text{el.}}^l(\vec{r}') + x'^l j_{\text{el.}}^k(\vec{r}')] d^3\vec{r}' + \frac{1}{2} \int_{\mathcal{V}} [x'^k j_{\text{el.}}^l(\vec{r}') - x'^l j_{\text{el.}}^k(\vec{r}')] d^3\vec{r}' \right\} \\ &= \sum_{l=1}^3 \frac{\vec{e}_l}{2} \int_{\mathcal{V}} [x'^k j_{\text{el.}}^l(\vec{r}') - x'^l j_{\text{el.}}^k(\vec{r}')] d^3\vec{r}'. \end{aligned} \quad (\text{VIII.29})$$

Als nächstes kann das Integrand in der dritten Zeile als

$$x'^k j_{\text{el.}}^l(\vec{r}') - x'^l j_{\text{el.}}^k(\vec{r}') = \sum_{m,n=1}^3 (\delta^{km} \delta^{ln} - \delta^{lm} \delta^{kn}) x'^m j_{\text{el.}}^n(\vec{r}') = \sum_{i,m,n=1}^3 \epsilon^{ikl} \epsilon^{imn} x'^m j_{\text{el.}}^n(\vec{r}')$$

geschrieben werden. Definiert man das *magnetische Dipolmoment* der elektrischen Ladungsstromverteilung durch

$$\vec{\mu} \equiv \frac{1}{2} \int_{\mathcal{V}} \vec{r}' \times \vec{j}_{\text{el.}}(\vec{r}') d^3\vec{r}' \quad (\text{VIII.30a})$$

mit kartesischen Komponenten

$$\mu^i = \sum_{m,n=1}^3 \frac{\epsilon^{imn}}{2} \int_{\mathcal{V}} x'^m j_{\text{el.}}^n(\vec{r}') d^3\vec{r}', \quad (\text{VIII.30b})$$

so gilt

$$\int_{\mathcal{V}} x'^k \vec{j}_{\text{el.}}(\vec{r}') d^3\vec{r}' = \sum_{i,l=1}^3 \epsilon^{ikl} \mu^i \vec{e}_l.$$

Nach Multiplikation mit x^k und Summe über $k = 1, 2, 3$ ergibt sich

$$\sum_{k=1}^3 x^k \int_{\mathcal{V}} x'^k \vec{j}_{\text{el.}}(\vec{r}') d^3\vec{r}' = \sum_{i,k,l=1}^3 \epsilon^{ikl} \mu^i x^k \vec{e}_l = \vec{\mu} \times \vec{r}.$$

Somit wird die Entwicklung (VIII.26) zu

$$\vec{A}(\vec{r}) = \frac{\mu_0}{4\pi} \frac{\vec{\mu} \times \vec{r}}{r^3} + \dots \quad (\text{VIII.30c})$$

Die zugehörige magnetische Induktion $\vec{B}(\vec{r})$ ergibt sich aus der Rotationsbildung dieses Vektorpotentials. Insbesondere gilt für das magnetische Feld eines reinen magnetischen Dipols $\vec{\mu}$

$$\vec{B}(\vec{r}) = \frac{\mu_0}{4\pi} \frac{3(\vec{r} \cdot \vec{\mu})\vec{r} - r^2 \vec{\mu}}{r^5}. \quad (\text{VIII.31})$$

Wie erwartet nimmt der Betrag des Vektorpotentials (VIII.30c) wie $1/r^2$ ab, und dementsprechend die Stärke des magnetischen Feldes wie $1/r^3$.

Bemerkungen:

* Die SI-Einheit des magnetischen Dipolmoments ist das $\text{A} \cdot \text{m}^2$, entsprechend der physikalischen Dimension $[\mu] = \text{IL}^2$.

* Die Gleichungen (VIII.30c) und (VIII.31) für das Potential und das Feld eines magnetischen Dipols sind analog den Gl. (VII.34a) und (VII.34b) für einen elektrostatischen Dipol.

VIII.2.2 Magnetisches Dipolmoment einer Leiterschleife

Durch eine Leiterschleife \mathcal{C} fließt ein elektrischer Strom I , der einer Ladungsstromdichte \vec{j}_{el} entspricht. Unter Nutzung der Substitution (VIII.18) lautet das zugehörige magnetische Dipolmoment (VIII.31)

$$\vec{\mu} = \frac{1}{2} \int \vec{r} \times \vec{j}_{\text{el}}(\vec{r}) d^3\vec{r} = \frac{1}{2} \oint_{\mathcal{C}} I \vec{r} \times d\vec{\ell},$$

wobei $d\vec{\ell}$ das gerichtete Linienelement entlang der Schleife bezeichnet. Im Integranden ist $\frac{1}{2} \vec{r} \times d\vec{\ell}$ ein Vektor, der senkrecht zur Ebene der Schleife steht, und dessen Betrag gleich dem Flächeninhalt $d\mathcal{S}$ des durch \vec{r} und $d\vec{\ell}$ aufgespannten Dreiecks ist. In der Integration entlang der Schleife ergibt sich die gesamte Fläche \mathcal{S} , die durch den Leiter aufgespannt wird, und somit

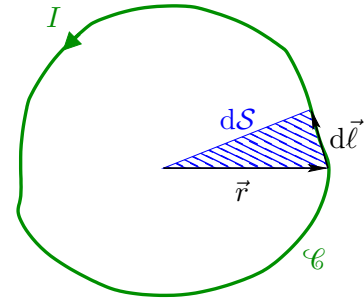


Abbildung VIII.6

$$\vec{\mu} = I \vec{\mathcal{S}} \quad (\text{VIII.32})$$

wobei der orthogonale Flächenvektor $\vec{\mathcal{S}}$ in Richtung des Daumens der rechten Hand zeigt, wenn die anderen Finger die Richtung des Stroms anzeigen.

Bemerkungen:

* Das Ergebnis ist unabhängig von der Wahl des Ursprungspunktes $\vec{r} = \vec{0}$, der auch außer der Schleife gewählt werden kann.

* Betrachtet man die bewegten Ladungsträger in der Leiterschleife, die für die elektrische Stromdichte \vec{j}_{el} verantwortlich sind, so ist das magnetische Dipolmoment $\vec{\mu}$ proportional zu deren Bahndrehimpuls \vec{L} .

VIII.2.3 Magnetischer Dipol in einem äußeren magnetischen Feld

Der Einfluss eines magnetischen Feldes auf eine Stromverteilung kann durch die magnetischen Multipolmomente der letzteren ausgedrückt werden.

VIII.2.3 a Kraft auf einen magnetischen Dipol in einem magnetischen Feld

Sei angenommen, dass sich die Stromverteilung \vec{j}_{el} innerhalb eines Gebiets \mathcal{V} in der Umgebung des Nullpunkts des Koordinatensystems befindet, wo eine nicht-spezifizierte Quelle eine magnetische Induktion \vec{B} erzeugt. Ausgehend von der Lorentz-Kraftdichte (VIII.23) lautet die Kraft auf die Stromverteilung

$$\vec{F} = \int_{\mathcal{V}} \vec{j}_{\text{el}}(\vec{r}) \times \vec{B}(\vec{r}) d^3\vec{r}.$$

Dessen i -te kartesische Komponente lautet

$$F^i = \epsilon^{ikl} \int_{\mathcal{V}} j_{\text{el}}^k(\vec{r}) B^l(\vec{r}) d^3\vec{r}.$$

In diesem Ausdruck kann das magnetische Feld als Taylor-Entwicklung um den Nullpunkt geschrieben werden:

$$B^l(\vec{r}) = B^l(\vec{0}) + \vec{r} \cdot \vec{\nabla} B^l(\vec{0}) + \dots$$

Somit gilt

$$F^i = \sum_{k,l=1}^3 \epsilon^{ikl} \left[\int_{\mathcal{V}} j_{\text{el}}^k(\vec{r}) B^l(\vec{0}) d^3\vec{r} + \int_{\mathcal{V}} j_{\text{el}}^k(\vec{r}) \vec{r} \cdot \vec{\nabla} B^l(\vec{0}) d^3\vec{r} + \dots \right].$$

Im ersten Term auf der rechten Seite kann $B^l(\vec{0})$ aus dem Integral herausgezogen werden: dieser konstante Faktor multipliziert dann das Volumenintegral von j_{el}^k , das gemäß Gl. (VIII.28) Null ist.

Es bleibt

$$F^i = \sum_{k,l=1}^3 \epsilon^{ikl} \int_{\mathcal{V}'} j_{\text{el.}}^k(\vec{r}) \vec{r} \cdot \vec{\nabla} B^l(\vec{0}) d^3\vec{r} + \dots = \sum_{j,k,l=1}^3 \epsilon^{ikl} \int_{\mathcal{V}'} j_{\text{el.}}^k(\vec{r}) x^j \frac{\partial B^l(\vec{0})}{\partial x^j} d^3\vec{r} + \dots$$

übrig. Die Ableitung im Integranden ist eine Konstante, die sich aus dem Integral herausziehen lässt. Dann erkennt man das Integral von $x^j j_{\text{el.}}^k$, das laut der Formel unten Gl. (VIII.30b) als

$$\int_{\mathcal{V}'} x^j j_{\text{el.}}^k(\vec{r}') d^3\vec{r}' = \sum_{m=1}^3 \epsilon^{mjk} \mu^m$$

umgeschrieben werden kann, wobei μ^m die m -te Komponente des magnetischen Dipolmoments der Stromverteilung ist. Dies gibt

$$F^i = \sum_{j,k,l,m=1}^3 \epsilon^{ikl} \epsilon^{mjk} \mu^m \frac{\partial B^l(\vec{0})}{\partial x^j} = \sum_{j,l,m=1}^3 (-\delta^{im} \delta^{jl} + \delta^{ij} \delta^{lm}) \mu^m \frac{\partial B^l(\vec{0})}{\partial x^j} = -\mu^i \vec{\nabla} \cdot \vec{B}(\vec{0}) + \vec{\mu} \cdot \frac{\partial \vec{B}(\vec{0})}{\partial x^i}.$$

Der erste Term im rechten Glied der Gleichung verschwindet dank der Maxwell–Thomson-Gleichung. Dazu ist das magnetische Dipolmoment für eine gegebene Stromverteilung eine Konstante, die in der Ableitung des zweiten Terms mitgenommen werden kann:

$$F^i = \frac{\partial}{\partial x^j} [\vec{\mu} \cdot \vec{B}(\vec{0})].$$

Somit ergibt sich schließlich für die Kraft auf ein magnetisches Dipolmoment, das sich im Punkt \vec{r} befindet ($\vec{r} = \vec{0}$ in der obigen Herleitung)

$$\vec{F} = \vec{\nabla} [\vec{\mu} \cdot \vec{B}(\vec{r})]. \quad (\text{VIII.33})$$

Bemerkung: Laut dieser Formel erfährt ein magnetischer Dipol eine Kraft, wenn er sich in einem inhomogenen Feld befindet.

VIII.2.3b Energie eines magnetischen Dipols in einem magnetischen Feld

Die Kraft (VIII.33) lässt sich trivial als das Negative eines Gradienten schreiben, und zwar

$$\vec{F} = -\vec{\nabla} V_{\text{W}}, \quad (\text{VIII.34a})$$

wobei die Wechselwirkungsenergie zwischen dem Dipol und der magnetischen Induktion, aus der die Kraft abgeleitet wird, durch

$$V_{\text{W}} = -\vec{\mu} \cdot \vec{B}(\vec{r}) \quad (\text{VIII.34b})$$

gegeben ist.

Diese Wechselwirkungsenergie hat genau die gleiche Form wie jene eines elektrischen Dipols in einem elektrostatischen Feld, entsprechend dem zweiten Term in Gl. (VII.31).

Ebenfalls analog zum elektrostatischen Fall ist die Wechselwirkungsenergie zweier magnetischer Dipole $\vec{\mu}_a$ und $\vec{\mu}_b$: unter Verwendung des magnetischen Dipolfeldes (VIII.31) und der Gl. (VIII.34b) ergibt sich

$$V_{\text{W}} = \frac{\mu_0}{4\pi} \frac{\vec{\mu}_a \cdot \vec{\mu}_b - 3(\vec{e}_r \cdot \vec{\mu}_a)(\vec{e}_r \cdot \vec{\mu}_b)}{r^3} \quad (\text{VIII.35})$$

mit dem Abstandsvektor \vec{r} zwischen den beiden (als punktförmig angenommenen) Dipole und dem zugehörigen Einheitsvektor $\vec{e}_r \equiv \vec{r}/|\vec{r}|$. Diese Gleichung ist das genaue Pendant der Gl. (VII.35).

Literatur zum Kapitel VIII

- Fließbach, *Elektrodynamik* [3] Teil III.
- Greiner, *Klassische Elektrodynamik* [8] Teil III.
- Griffiths, *Elektrodynamik* [9] = *Introduction to Electrodynamics* [10] Kap. 5.
- Jackson, *Klassische Elektrodynamik* [11] = *Classical Electrodynamics* [12] Kap. 5.1–5.7.
- Landau & Lifschitz, *Klassische Feldtheorie* [14] = *The classical theory of fields* [28] Kap. 5 § 43–44
- Nolting, *Elektrodynamik* [17] Kap. 3.1–3.3.
- Scheck, *Klassische Feldtheorie* [19] Kap. 1.8.

KAPITEL IX

Zeitabhängige elektromagnetische Felder

IX.1	Grundgesetze	176
IX.1.1	Maxwell-Gleichungen	176
IX.1.2	Bewegungsgleichungen für die elektrischen und magnetischen Felder	178
IX.2	Elektrodynamische Potentiale	179
IX.2.1	Definition	180
IX.2.2	Eichinvarianz	180
IX.2.3	Bewegungsgleichungen für die elektrodynamischen Potentiale	182
IX.3	Energie und Impuls des elektromagnetischen Feldes	183
IX.3.1	Energiedichte und -stromdichte des elektromagnetischen Feldes	183
IX.3.2	Impulsdichte und -stromdichte des elektromagnetischen Feldes	184
IX.4	Elektromagnetische Wellen im Vakuum	187
IX.4.1	Klassische Wellengleichung	187
IX.4.2	Elektromagnetische Wellen	190
IX.5	Klassische Theorie der Strahlung	193
IX.5.1	Greensche Funktion der klassischen Wellengleichung	193
IX.5.2	Retardierte Potentiale	195
IX.5.3	Multipolentwicklung	197
IX.5.4	Potentiale und Felder einer bewegten Punktladung	198

Im Gegensatz zu den stationären Fällen der Elektro- und Magnetostatik sind das elektrische und das magnetische Feld im allgemeineren zeitabhängigen Fall miteinander gekoppelt. Diese Kopplung wird durch zusätzliche Terme in den Grundgleichungen berücksichtigt, die zu neuen Phänomenen führen (Abschn. IX.1).

In Abschn. IX.2 wird gezeigt, dass sich das elektrische und das magnetische Feld aus elektromagnetischen Potentialen ableiten lassen, die nicht eindeutig bestimmt sind. Darüber hinaus werden Bewegungsgleichungen für diese Potentiale hergeleitet. Dem elektromagnetischen Feld können noch eine Energie und ein Impuls zugeordnet werden, oder genauer, entsprechende Dichten und Stromdichten (Abschn. IX.3), die ziemlich anschaulichen Bilanzgleichungen genügen.

Der zweite Teil des Kapitels befasst sich mit Anwendungen der zuvor eingeführten Begriffe und Ergebnisse. Zunächst werden Lösungen zu den Bewegungsgleichungen für die Felder oder Potentiale im Vakuum in Abwesenheit von Quellen untersucht, entsprechend elektromagnetischen Wellen (Abschn. IX.4). Dann wird eine allgemeine Lösung der Maxwell-Gleichungen mit Hilfe sogenannter retardierter Potentiale dargelegt, die das elektromagnetische Feld erzeugt durch beliebige Ladungs- und Stromverteilungen angibt (Abschn. IX.5).

IX.1 Grundgesetze

In diesem Abschnitt werden zuerst die nicht-stationären Maxwell-Gleichungen dargelegt und die Deutung der zeitabhängigen Terme diskutiert (§ IX.1.1). Dann werden Bewegungsgleichungen zweiter Ordnung hergeleitet (§ IX.1.2), deren Lösung oft einfacher ist.

IX.1.1 Maxwell-Gleichungen

Die schon in Kapitel VI eingeführten zeitabhängigen Maxwell-Gleichungen lauten

$$\vec{\nabla} \cdot \vec{E}(t, \vec{r}) = \frac{1}{\epsilon_0} \rho_{\text{el.}}(t, \vec{r}) \quad (\text{IX.1a})$$

$$\vec{\nabla} \cdot \vec{B}(t, \vec{r}) = 0 \quad (\text{IX.1b})$$

$$\vec{\nabla} \times \vec{E}(t, \vec{r}) + \frac{\partial \vec{B}(t, \vec{r})}{\partial t} = \vec{0} \quad (\text{IX.1c})$$

$$\vec{\nabla} \times \vec{B}(t, \vec{r}) - \epsilon_0 \mu_0 \frac{\partial \vec{E}(t, \vec{r})}{\partial t} = \mu_0 \vec{j}_{\text{el.}}(t, \vec{r}). \quad (\text{IX.1d})$$

Relativ zu den stationären Gleichungen (VII.1) und (VIII.1) der Elektro- und Magnetostatik treten zwei neue Terme auf, und zwar die Zeitableitungen in den zwei letzten Gleichungen. Natürlich sind diese Zeitableitungen nötig, um eine mögliche Zeitentwicklung der Felder zu beschreiben.

IX.1.1 a Faraday-Induktionsgesetz

Schreibt man die Maxwell-Faraday-Gleichung (IX.1c) in der Form

$$\vec{\nabla} \times \vec{E}(t, \vec{r}) = -\frac{\partial \vec{B}(t, \vec{r})}{\partial t}, \quad (\text{IX.2})$$

so kann die Änderung der magnetischen Flussdichte auf der rechten Seite als Quellterm für die Rotation des elektrischen Feldes gesehen werden.

Zur Interpretation dieser Gleichung kann man eine feste ruhende orientierte Fläche \mathcal{S} betrachten und Gl. (IX.2) darüber integrieren:

$$\int_{\mathcal{S}} [\vec{\nabla} \times \vec{E}(t, \vec{r})] \cdot d^2 \vec{\mathcal{S}} = - \int_{\mathcal{S}} \frac{\partial \vec{B}(t, \vec{r})}{\partial t} \cdot d^2 \vec{\mathcal{S}}. \quad (\text{IX.3})$$

Da die Fläche zeitunabhängig ist, kann die Zeitableitung auf der rechten Seite aus dem Integral herausgezogen werden:

$$\int_{\mathcal{S}} \frac{\partial \vec{B}(t, \vec{r})}{\partial t} \cdot d^2 \vec{\mathcal{S}} = \frac{d}{dt} \int_{\mathcal{S}} \vec{B}(t, \vec{r}) \cdot d^2 \vec{\mathcal{S}} = \frac{d}{dt} \Phi_B(t),$$

wobei Φ_B der Fluss der magnetischen Induktion durch die Fläche \mathcal{S} ist.

Wiederum kann die linke Seite der Gl. (IX.3) mit dem Integralsatz von Stokes transformiert werden: wenn $\partial \mathcal{S}$ den (orientierten) Rand der Fläche \mathcal{S} bezeichnet, gilt

$$\int_{\mathcal{S}} [\vec{\nabla} \times \vec{E}(t, \vec{r})] \cdot d^2 \vec{\mathcal{S}} = \oint_{\partial \mathcal{S}} \vec{E}(t, \vec{r}) \cdot d\vec{\ell}.$$

Schließlich ergibt sich

$$\oint_{\partial \mathcal{S}} \vec{E}(t, \vec{r}) \cdot d\vec{\ell} = - \frac{d}{dt} \int_{\mathcal{S}} \vec{B}(t, \vec{r}) \cdot d^2 \vec{\mathcal{S}} = - \frac{d\Phi_B(t)}{dt}. \quad (\text{IX.4})$$

Wendet man dieses Ergebnis auf den Fall an, wo die Fläche \mathcal{S} durch einen ruhenden geschlossenen Kreisleiter \mathcal{C} aufspannt wird (Abb. IX.1), so ist das Linienintegral entlang \mathcal{C} die sog. *elektromotorische Kraft* U — die trotz ihrer Bezeichnung keine Kraft ist, sondern eine elektrische Spannung:

$$U \equiv \oint_{\mathcal{C}} \vec{E}(t, \vec{r}) \cdot d\vec{\ell}.$$

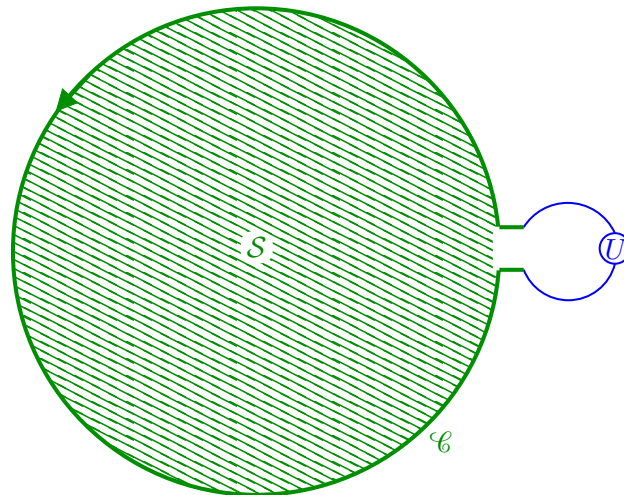


Abbildung IX.1

Dies entspricht der Energie erhalten durch eine Einheitsladung, wenn sie den Kreis einmal umläuft. (Wird der Kreis irgendwo leicht getrennt, so ist U die elektrische Spannung zwischen den Enden, wie in Abb. IX.1 dargestellt wird.) Dann ergibt sich das *Faraday-Induktionsgesetz*

$$U = - \frac{d\Phi_B(t)}{dt}. \quad (\text{IX.5})$$

Bemerkungen:

* Wenn die Fläche \mathcal{S} , durch welche der magnetische Fluss zeitlich variiert, nicht fest ist, müssen zusätzliche Terme in der obigen Argumentation berücksichtigt werden. Physikalisch bleibt das Ergebnis das gleiche.

* In einem geschlossenen Kreisleiter wird das induzierte elektrische Feld die Ladungsträger in Bewegung bringen, d.h. es entsteht ein elektrischer Strom. Wiederum wird dieser Strom über die Maxwell–Ampère-Gleichung ein magnetisches Feld \vec{B}_{ind} erzeugen, das nach der *Lenzschen*^(al) *Regel* der Änderung des magnetischen Flusses durch die Leiterschleife entgegenwirkt.

* Das Prinzip des Induktionsgesetzes wird zu Nutze gemacht, um elektrische Ströme zu erzeugen.

IX.1.1 b Maxwellischer Verschiebungsstrom

Die zweite Modifikation relativ zum stationären Fall ist die Zeitableitung in der Maxwell–Ampère-Gleichung (IX.1d), die sich auch als

$$\vec{\nabla} \times \vec{B}(t, \vec{r}) = \mu_0 \left[\vec{j}_{\text{el.}}(t, \vec{r}) + \epsilon_0 \frac{\partial \vec{E}(t, \vec{r})}{\partial t} \right] \quad (\text{IX.6})$$

schreiben lässt. In dieser Form kann der zweite Term in den eckigen Klammern als eine Ladungsstromdichte interpretiert werden: dieser Term ist der sog. Maxwellische *Verschiebungsstrom*.

Dank diesem Term enthalten die Maxwell-Gleichungen (IX.1) automatisch die *Ladungserhaltung* für die Quellen. Leitet man nämlich die Maxwell–Gauß-Gleichung (IX.1a) nach der Zeit ab, so ergibt sich

$$\frac{\partial \rho_{\text{el.}}(t, \vec{r})}{\partial t} = \epsilon_0 \frac{\partial}{\partial t} [\vec{\nabla} \cdot \vec{E}(t, \vec{r})].$$

^(al)E. LENZ, 1804–1865

Andererseits lautet die Divergenz der Maxwell–Ampère-Gleichung (IX.1d)

$$\vec{\nabla} \cdot \vec{j}_{\text{el.}}(t, \vec{r}) = \frac{1}{\mu_0} \vec{\nabla} \cdot [\vec{\nabla} \times \vec{B}(t, \vec{r})] - \epsilon_0 \vec{\nabla} \cdot \frac{\partial \vec{E}(t, \vec{r})}{\partial t} = -\epsilon_0 \frac{\partial}{\partial t} [\vec{\nabla} \cdot \vec{E}(t, \vec{r})],$$

wobei in der zweiten Gleichung die partiellen Ableitungen von \vec{E} nach der Zeit und nach den Ortskoordinaten ausgetauscht wurden, während $\vec{\nabla} \cdot (\vec{\nabla} \times \vec{B}) = 0$ benutzt wurde. Die Summe der zwei letzten Gleichungen lautet dann

$$\boxed{\frac{\partial \rho_{\text{el.}}(t, \vec{r})}{\partial t} + \vec{\nabla} \cdot \vec{j}_{\text{el.}}(t, \vec{r}) = 0.} \quad (\text{IX.7})$$

Diese Differentialgleichung wird als *Kontinuitätsgleichung* bezeichnet. Sie stellt die lokale Formulierung der Erhaltung der elektrischen Ladung dar.

Um die entsprechende Integralform zu finden kann man ein festes Raumgebiet \mathcal{V} mit Rand $\partial\mathcal{V}$ betrachten und die Gl. (IX.7) über das Volumen \mathcal{V} integrieren:

$$\int_{\mathcal{V}} \frac{\partial \rho_{\text{el.}}(t, \vec{r})}{\partial t} d^3\vec{r} = - \int_{\mathcal{V}} \vec{\nabla} \cdot \vec{j}_{\text{el.}}(t, \vec{r}) d^3\vec{r}.$$

Da das Volumen \mathcal{V} zeitunabhängig ist, kann die Zeitableitung im linken Glied aus dem Integral herausgezogen werden. Währenddessen kann das rechte Glied mit dem Integralsatz von Gauß transformiert werden:

$$\frac{\partial}{\partial t} \int_{\mathcal{V}} \rho_{\text{el.}}(t, \vec{r}) d^3\vec{r} = - \oint_{\partial\mathcal{V}} \vec{j}_{\text{el.}}(t, \vec{r}) \cdot d^2\vec{S}. \quad (\text{IX.8})$$

Der Term auf der linken Seite ist die (totale) Zeitableitung der Gesamtladung $Q_{\text{in}}(t)$ innerhalb des Volumens \mathcal{V} . Das Flächenintegral auf der rechten Seite ist der nach außen gerichtete Fluss der Ladungsstromdichte durch $\partial\mathcal{V}$. Physikalisch bedeutet Gl. (IX.8), dass die Änderung der Gesamtladung Q_{in} pro Zeiteinheit gleich der durch die Oberfläche $\partial\mathcal{V}$ „verlorenen“ Ladung pro Zeiteinheit ist: dies ist eine Bilanzgleichung für eine erhaltene Größe, nämlich hier die elektrische Ladung.

IX.1.2 Bewegungsgleichungen für die elektrischen und magnetischen Felder

Die Maxwell-Gleichungen (IX.1) sind gekoppelte partielle Differentialgleichungen erster Ordnung für $\vec{E}(t, \vec{r})$ und $\vec{B}(t, \vec{r})$. Mit einigen Umrechnungen lassen sich entkoppelte Bewegungsgleichungen erhalten.

Somit ist die Differentialgleichung (IX.9) für $\vec{E}(t, \vec{r})$ unabhängig von $\vec{B}(t, \vec{r})$, und umgekehrt tritt $\vec{E}(t, \vec{r})$ nicht in der Bewegungsgleichung (IX.11) für $\vec{B}(t, \vec{r})$ auf. Die beiden Felder bleiben aber miteinander gekoppelt, beispielsweise über die Maxwell–Faraday-Gleichung (IX.1c). Das heißt, man kann zwar entkoppelte Bewegungsgleichungen finden, die aber nicht unabhängig voneinander sind: physikalisch wird die Kopplung nicht aufgehoben.

IX.1.2a Bewegungsgleichung für das elektrische Feld

Die Rotationsbildung der Maxwell–Faraday-Gleichung (IX.1c) ergibt

$$\vec{\nabla} \times [\vec{\nabla} \times \vec{E}(t, \vec{r})] + \vec{\nabla} \times \frac{\partial \vec{B}(t, \vec{r})}{\partial t} = \vec{0}.$$

Der erste Term kann mit der Identität $\vec{\nabla} \times (\vec{\nabla} \times \vec{V}) = \vec{\nabla}(\vec{\nabla} \cdot \vec{V}) - \Delta \vec{V}$ transformiert werden. Im zweiten Term können partielle Zeit- und Raumableitung ausgetauscht werden, was zu

$$\vec{\nabla} [\vec{\nabla} \cdot \vec{E}(t, \vec{r})] - \Delta \vec{E}(t, \vec{r}) + \frac{\partial}{\partial t} [\vec{\nabla} \times \vec{B}(t, \vec{r})] = \vec{0}$$

führt. Dabei kann $\vec{\nabla} \cdot \vec{E}$ unter Verwendung der Maxwell–Gauß-Gleichung (IX.1a) durch die Ladungsdichte ausgedrückt werden. Währenddessen lässt sich $\vec{\nabla} \times \vec{B}$ mithilfe der Maxwell–Ampère-Gleichung (IX.1d) umschreiben. Dies ergibt

$$\frac{1}{\epsilon_0} \vec{\nabla} \rho_{\text{el.}}(t, \vec{r}) - \Delta \vec{E}(t, \vec{r}) + \frac{\partial}{\partial t} \left[\mu_0 \vec{j}_{\text{el.}}(t, \vec{r}) + \epsilon_0 \mu_0 \frac{\partial \vec{E}(t, \vec{r})}{\partial t} \right] = \vec{0}.$$

Äquivalent gilt

$$\square \vec{E}(t, \vec{r}) \equiv -\epsilon_0 \mu_0 \frac{\partial^2 \vec{E}(t, \vec{r})}{\partial t^2} + \Delta \vec{E}(t, \vec{r}) = \mu_0 \frac{\partial \vec{j}_{\text{el.}}(t, \vec{r})}{\partial t} + \frac{1}{\epsilon_0} \vec{\nabla} \rho_{\text{el.}}(t, \vec{r}). \quad (\text{IX.9})$$

Dabei ist der *d'Alembert^(am)-Operator* \square gemäß

$$\square \equiv -\epsilon_0 \mu_0 \frac{\partial^2}{\partial t^2} + \Delta \quad (\text{IX.10})$$

definiert.⁽⁷⁵⁾

IX.1.2b Bewegungsgleichung für das magnetische Feld

Betrachtet man als Anfangspunkt die Maxwell–Ampère-Gleichung (IX.1d) und bildet man deren Rotation, so ergibt sich

$$\vec{\nabla} \times [\vec{\nabla} \times \vec{B}(t, \vec{r})] - \epsilon_0 \mu_0 \vec{\nabla} \times \frac{\partial \vec{E}(t, \vec{r})}{\partial t} = \mu_0 \vec{\nabla} \times \vec{j}_{\text{el.}}(t, \vec{r}).$$

Dabei ist $\vec{\nabla} \times (\vec{\nabla} \times \vec{B})$ wegen der Maxwell–Thomson-Gleichung $\vec{\nabla} \cdot \vec{B} = 0$ einfach gleich $-\Delta \vec{B}$. Dann können Ableitungen bezüglich der Zeit- und Ortsvariablen im zweiten Term ausgetauscht werden, und die resultierende Rotation $\vec{\nabla} \times \vec{E}$ mithilfe der Maxwell–Faraday-Gleichung (IX.1c) umgeschrieben werden. Insgesamt ergibt sich

$$-\Delta \vec{B}(t, \vec{r}) + \epsilon_0 \mu_0 \frac{\partial^2 \vec{B}(t, \vec{r})}{\partial t^2} = \mu_0 \vec{\nabla} \times \vec{j}_{\text{el.}}(t, \vec{r}),$$

d.h.

$$\square \vec{B}(t, \vec{r}) \equiv -\epsilon_0 \mu_0 \frac{\partial^2 \vec{B}(t, \vec{r})}{\partial t^2} + \Delta \vec{B}(t, \vec{r}) = \mu_0 \vec{\nabla} \times \vec{j}_{\text{el.}}(t, \vec{r}). \quad (\text{IX.11})$$

Die Gleichungen (IX.9) für das elektrische Feld \vec{E} und (IX.11) für die magnetische Induktion \vec{B} sind anscheinend entkoppelte (inhomogene) lineare partielle Differentialgleichungen zweiter Ordnung. Die Felder sollen aber noch der Maxwell–Faraday-Gleichung (IX.1c) genügen, so dass sie nicht unabhängig voneinander sind: die Bewegungsgleichungen (IX.9) und (IX.11) sind also eigentlich redundant.

IX.2 Elektrodynamische Potentiale

Wie im stationären Fall können ein Skalar- und ein Vektorpotential eingeführt werden, aus denen sich das elektrische und das magnetische Feld ableiten lassen (§ IX.2.1). Im Gegensatz zu den stationären Potentialen hängen aber jetzt beide Felder \vec{E} und \vec{B} vom Vektorpotential ab. Dementsprechend sind die Eichtransformationen der Potentiale, welche die elektromagnetischen Felder invariant lassen,

⁽⁷⁵⁾Oft wird der d'Alembert-Operator als das Negative des hier betrachteten Differentialoperators definiert.

^(am)J. LE ROND D'ALEMBERT, 1717–1783

nicht unabhängig voneinander, sondern beide Potentiale müssen gleichzeitig transformiert werden (§ IX.2.2). Schließlich werden in § IX.2.3 Bewegungsgleichungen für die Potentiale dargelegt, die aus den Maxwell-Gleichungen folgen.

IX.2.1 Definition

IX.2.1 a Vektorpotential

Die Maxwell–Thomson-Gleichung (IX.1b) hat die gleiche Form, wie in der Magnetostatik. Daher kann man wie in § VIII.1.1 die Existenz eines differenzierbaren Vektorfeldes $\vec{A}(\vec{r})$ folgern, mit dem die Beziehung

$$\vec{B}(t, \vec{r}) = \vec{\nabla} \times \vec{A}(t, \vec{r}) \quad (\text{IX.12})$$

in jedem Punkt \vec{r} und zu jeder Zeit t erfüllt wird: $\vec{A}(t, \vec{r})$ ist das (elektrodynamische) *Vektorpotential*.

Ähnlich wie in der Magnetostatik existiert eine (Eich)Freiheit in der Auswahl von $\vec{A}(t, \vec{r})$. Ist ein erstes Vektorpotential \vec{A} für eine bestimmte magnetische Induktion \vec{B} geeignet, dann so ist auch das Vektorfeld $\vec{A}'(t, \vec{r}) \equiv \vec{A}(t, \vec{r}) + \vec{\nabla}\chi(t, \vec{r})$, wobei $\chi(t, \vec{r})$ eine beliebige skalare Funktion von Zeit und Ort ist.

IX.2.1 b Skalarpotential

Setzt man die Beziehung (IX.12) in die Maxwell–Faraday-Gleichung (IX.1c) ein, so kommt

$$\vec{\nabla} \times \vec{E}(t, \vec{r}) + \frac{\partial}{\partial t} [\vec{\nabla} \times \vec{A}(t, \vec{r})] = \vec{0}.$$

Nach Austauschen der Zeit- und Ortsableitungen im zweiten Term kann man einfacher

$$\vec{\nabla} \times \left[\vec{E}(t, \vec{r}) + \frac{\partial \vec{A}(t, \vec{r})}{\partial t} \right] = \vec{0}.$$

schreiben. Daraus folgt, dass es ein skalares Feld $\Phi(t, \vec{r})$ existiert, das (elektrodynamische) *Skalarpotential*, das die Beziehung

$$\vec{E}(t, \vec{r}) = -\vec{\nabla}\Phi(t, \vec{r}) - \frac{\partial \vec{A}(t, \vec{r})}{\partial t} \quad (\text{IX.13})$$

zu jeder Zeit t und in jedem Punkt \vec{r} erfüllt.

Im stationären Fall findet man die Beziehung (VII.3) zwischen elektrostatischem Potential und Feld wieder.

IX.2.2 Eichinvarianz

IX.2.2 a Eichtransformationen

Zwei Vektorpotentiale \vec{A} und $\vec{A}' \equiv \vec{A} + \vec{\nabla}\chi$ führen zwar über die Beziehung (IX.12) zum gleichen magnetischen Feld \vec{B} ; laut der Gl. (IX.13) würden sie zu unterschiedlichen elektrischen Feldern \vec{E} , \vec{E}' führen.

Führt man aber eine gleichzeitige Transformation des Vektor- und des Skalarpotentials, und zwar die *Eichtransformation*

$$\Phi(t, \vec{r}) \rightarrow \Phi'(t, \vec{r}) \equiv \Phi(t, \vec{r}) - \frac{\partial \chi(t, \vec{r})}{\partial t}, \quad (\text{IX.14a})$$

$$\vec{A}(t, \vec{r}) \rightarrow \vec{A}'(t, \vec{r}) \equiv \vec{A}(t, \vec{r}) + \vec{\nabla}\chi(t, \vec{r}), \quad (\text{IX.14b})$$

wobei die skalare Funktion $\chi(t, \vec{r})$ beliebig ist, so bleiben die zwei abgeleiteten Felder $\vec{E}(t, \vec{r})$ und $\vec{B}(t, \vec{r})$ unverändert.

Beispiel: Dem elektromagnetischen Feld bestehend aus einem stationären und gleichförmigen elektrischen Feld $\vec{E}(t, \vec{r}) = \vec{E}_0$ und einer verschwindenden magnetischen Induktion $\vec{B}(t, \vec{r}) = \vec{0}$ kann man z.B. die einfachen Potentiale

$$\Phi(t, \vec{r}) = -\vec{r} \cdot \vec{E}_0, \quad \vec{A}(t, \vec{r}) = \vec{0},$$

oder

$$\Phi'(t, \vec{r}) = 0, \quad \vec{A}'(t, \vec{r}) = -\vec{E}_0 t$$

zuordnen. Eine Funktion χ , die von den ersten zu den zweiten führt, ist $\chi(t, \vec{r}) = -\vec{r} \cdot \vec{E}_0 t$ (dazu darf man natürlich eine additive Konstante hinzufügen).

IX.2.2b Spezielle Eichungen

Die Eichfreiheit kann benutzt werden, um eine Eichung zu benutzen, in welcher einige Gleichungen eine einfachere Form annehmen.

Eine erste oft auftretende Wahl ist die *Coulomb-Eichung*, die durch die Bedingung

$$\vec{\nabla} \cdot \vec{A}(t, \vec{r}) = 0. \quad (\text{IX.15})$$

[vgl. Gl. (VIII.4)] definiert ist. Bei festen \vec{E} - und \vec{B} -Feldern ist es immer möglich, Potentiale zu finden, die der Coulomb-Eichbedingung (IX.15) genügen.

Seien Φ, \vec{A} Potentiale für ein gegebenes elektromagnetisches Feld. Wenn \vec{A} die Bedingung (IX.15) erfüllt, ist der Fall erledigt. Sonst sucht man eine Funktion $\chi(t, \vec{r})$ derart, dass die über Gl. (IX.14) eichtransformierten Potentiale Φ', \vec{A}' die Coulomb-Eichbedingung erfüllen. Die Divergenzbildung der Beziehung (IX.14b) zwischen \vec{A}' und \vec{A} lautet

$$\vec{\nabla} \cdot \vec{A}'(t, \vec{r}) = \vec{\nabla} \cdot \vec{A}(t, \vec{r}) + \vec{\nabla} \cdot \vec{\nabla} \chi(t, \vec{r}) = \vec{\nabla} \cdot \vec{A}(t, \vec{r}) + \Delta \chi(t, \vec{r}).$$

Das Ziel ist, dass $\vec{\nabla} \cdot \vec{A}'(t, \vec{r}) = 0$ gilt. Dafür soll man eine Funktion χ finden, die die inhomogene Poisson-Gleichung

$$\Delta \chi(t, \vec{r}) = -\vec{\nabla} \cdot \vec{A}(t, \vec{r})$$

löst, wobei die rechte Seite eine vorgegebene Funktion ist. Laut den Ergebnissen des Abschn. VII.2 existiert eine solche Lösung immer. \square

Eine andere günstige Eichung ist die *Lorenz^(an)-Eichung*, entsprechend der Bedingung

$$\epsilon_0 \mu_0 \frac{\partial \Phi(t, \vec{r})}{\partial t} + \vec{\nabla} \cdot \vec{A}(t, \vec{r}) = 0. \quad (\text{IX.16})$$

Wir werden in § IX.2.3 c sehen, dass sich die Bewegungsgleichungen für die Potentiale in der Lorenz-Eichung stark vereinfachen. Ein zweiter Vorteil der Eichung wird auch in § XII.1.2 diskutiert.

Die Coulomb- und Lorenz-Eichbedingungen legen die elektromagnetischen Potentiale Φ und \vec{A} nicht vollständig fest, sondern man kann in beiden Fällen unterschiedliche Paare (Φ, \vec{A}) finden, welche die zugehörige Bedingung erfüllen.

Dies ist trivial der Fall der Coulomb-Eichung, deren Bedingung (IX.15) das Skalarpotential frei lässt. Dass auch die Lorenz-Eichung „unvollständig“ ist, lässt sich einfach beweisen. Seien (Φ, \vec{A}) und (Φ', \vec{A}') Paare von Potentialen, die der Gl. (IX.16) genügen, wobei die gestrichelten Potentiale mit den nicht-gestrichelten über die Beziehungen (IX.14) zusammenhängen. Dann gilt

$$\epsilon_0 \mu_0 \frac{\partial \Phi'(t, \vec{r})}{\partial t} + \vec{\nabla} \cdot \vec{A}'(t, \vec{r}) - \left[\epsilon_0 \mu_0 \frac{\partial \Phi(t, \vec{r})}{\partial t} + \vec{\nabla} \cdot \vec{A}(t, \vec{r}) \right] = 0,$$

d.h. nach einer einfachen Berechnung

$$-\epsilon_0 \mu_0 \frac{\partial^2 \chi(t, \vec{r})}{\partial t^2} + \Delta \chi(t, \vec{r}) = \square \chi(t, \vec{r}) = 0. \quad (\text{IX.17})$$

^(an)L. LORENZ, 1829–1891

Diese Differentialgleichung hat nicht-triviale Lösungen — die in § IX.4.1 unten diskutiert werden —, so dass die gestrichenen und nicht-gestrichenen Potentiale ungleich sind.

IX.2.3 Bewegungsgleichungen für die elektrodynamischen Potentiale

IX.2.3a Bewegungsgleichung für das Vektorpotential

Das Einsetzen der Beziehungen (IX.12) und (IX.13) in die Maxwell–Ampère-Gleichung ergibt unter Berücksichtigung der Formel für die Rotation einer Rotation

$$\begin{aligned}\mu_0 \vec{j}_{\text{el.}}(t, \vec{r}) &= \vec{\nabla} \times [\vec{\nabla} \times \vec{A}(t, \vec{r})] - \epsilon_0 \mu_0 \frac{\partial}{\partial t} \left[-\vec{\nabla} \Phi(t, \vec{r}) - \frac{\partial \vec{A}(t, \vec{r})}{\partial t} \right] \\ &= \vec{\nabla} [\vec{\nabla} \cdot \vec{A}(t, \vec{r})] - \Delta \vec{A}(t, \vec{r}) + \epsilon_0 \mu_0 \vec{\nabla} \left[\frac{\partial \Phi(t, \vec{r})}{\partial t} \right] + \epsilon_0 \mu_0 \frac{\partial^2 \vec{A}(t, \vec{r})}{\partial t^2}.\end{aligned}$$

Nach trivialer Umschreibung lautet dies

$$\square \vec{A}(t, \vec{r}) \equiv -\epsilon_0 \mu_0 \frac{\partial^2 \vec{A}(t, \vec{r})}{\partial t^2} + \Delta \vec{A}(t, \vec{r}) = -\mu_0 \vec{j}_{\text{el.}}(t, \vec{r}) + \vec{\nabla} \left[\epsilon_0 \mu_0 \frac{\partial \Phi(t, \vec{r})}{\partial t} + \vec{\nabla} \cdot \vec{A}(t, \vec{r}) \right]. \quad (\text{IX.18})$$

Diese lineare partielle Differentialgleichung zweiter Ordnung stellt die Bewegungsgleichung für das Vektorpotential $\vec{A}(t, \vec{r})$ dar.

Bemerkung: Im zeitunabhängigen Fall vereinfacht sich Gl. (IX.18) zur Gleichung (VIII.8).

IX.2.3b Bewegungsgleichung für das Skalarpotential

Setzt man in die Maxwell–Gauß-Gleichung (IX.1a) die Beziehung (IX.13) ein, so kommt

$$\frac{\rho_{\text{el.}}(t, \vec{r})}{\epsilon_0} = \vec{\nabla} \cdot \left[-\vec{\nabla} \Phi(t, \vec{r}) - \frac{\partial \vec{A}(t, \vec{r})}{\partial t} \right] = -\Delta \Phi(t, \vec{r}) - \frac{\partial}{\partial t} [\vec{\nabla} \cdot \vec{A}(t, \vec{r})]$$

oder äquivalent

$$\Delta \Phi(t, \vec{r}) = -\frac{\rho_{\text{el.}}(t, \vec{r})}{\epsilon_0} - \frac{\partial}{\partial t} [\vec{\nabla} \cdot \vec{A}(t, \vec{r})]. \quad (\text{IX.19})$$

Man kann noch $\epsilon_0 \mu_0 \partial^2 \Phi(t, \vec{r}) / \partial t^2$ von beiden Seiten dieser Gleichung abziehen, um die Bewegungsgleichung in eine Form ähnlich der Gl. (IX.18) zu bringen:

$$\square \Phi(t, \vec{r}) \equiv -\epsilon_0 \mu_0 \frac{\partial^2 \Phi(t, \vec{r})}{\partial t^2} + \Delta \Phi(t, \vec{r}) = -\frac{\rho_{\text{el.}}(t, \vec{r})}{\epsilon_0} - \frac{\partial}{\partial t} \left[\epsilon_0 \mu_0 \frac{\partial \Phi(t, \vec{r})}{\partial t} + \vec{\nabla} \cdot \vec{A}(t, \vec{r}) \right]. \quad (\text{IX.20})$$

Bemerkung: Im stationären Fall vereinfacht sich Gl. (IX.20), oder noch einfacher Gl. (IX.19), zur Poisson-Gleichung (VII.4) der Elektrostatik.

IX.2.3c Bewegungsgleichungen in speziellen Eichungen

Arbeitet man in der durch die Bedingung (IX.16) definierten Lorenz-Eichung, so vereinfachen sich die Bewegungsgleichungen (IX.18) und (IX.20) erheblich, und zwar zu

$$\text{in Lorenz-Eichung} \quad \begin{cases} \square \Phi(t, \vec{r}) = -\frac{\rho_{\text{el.}}(t, \vec{r})}{\epsilon_0} & (\text{IX.21a}) \\ \square \vec{A}(t, \vec{r}) = -\mu_0 \vec{j}_{\text{el.}}(t, \vec{r}). & (\text{IX.21b}) \end{cases}$$

Somit nehmen die Bewegungsgleichungen für die Potentiale die gleiche Form wie solche (IX.9), (IX.11) für die daraus abgeleiteten \vec{E} - und \vec{B} -Felder an.

IX.3 Energie und Impuls des elektromagnetischen Feldes

Genau wie mechanische Systeme trägt das elektromagnetische Feld Energie (§ IX.3.1) und Impuls (§ IX.3.2). Lokal lassen sich diese Größen durch zugehörige Dichten und Stromdichten charakterisieren, die dann Bilanzgleichungen genügen.

IX.3.1 Energiedichte und -stromdichte des elektromagnetischen Feldes

IX.3.1 a Bilanzgleichung für die Energie

Bildet man das Skalarprodukt aus der Maxwell–Faraday-Gleichung (IX.1c) und der magnetischen Induktion $\vec{B}(t, \vec{r})$, so kommt

$$\vec{B}(t, \vec{r}) \cdot [\vec{\nabla} \times \vec{E}(t, \vec{r})] + \vec{B}(t, \vec{r}) \cdot \frac{\partial \vec{B}(t, \vec{r})}{\partial t} = 0.$$

Wiederum lautet das Skalarprodukt aus Maxwell–Ampère-Gleichung (IX.1d) und elektrischem Feld $\vec{E}(t, \vec{r})$

$$\vec{E}(t, \vec{r}) \cdot [\vec{\nabla} \times \vec{B}(t, \vec{r})] - \epsilon_0 \mu_0 \vec{E}(t, \vec{r}) \cdot \frac{\partial \vec{E}(t, \vec{r})}{\partial t} = \mu_0 \vec{j}_{\text{el.}}(t, \vec{r}) \cdot \vec{E}(t, \vec{r}).$$

Die letztere Gleichung kann dann von der ersteren abgezogen werden. Unter Verwendung der Identität $\vec{B} \cdot (\vec{\nabla} \times \vec{E}) - \vec{E} \cdot (\vec{\nabla} \times \vec{B}) = \vec{\nabla} \cdot (\vec{E} \times \vec{B})$ ergibt sich

$$\vec{\nabla} \cdot [\vec{E}(t, \vec{r}) \times \vec{B}(t, \vec{r})] + \mu_0 \frac{\partial}{\partial t} \left[\frac{\epsilon_0 \vec{E}(t, \vec{r})^2}{2} + \frac{\vec{B}(t, \vec{r})^2}{2\mu_0} \right] = -\mu_0 \vec{j}_{\text{el.}}(t, \vec{r}) \cdot \vec{E}(t, \vec{r}). \quad (\text{IX.22})$$

Die Struktur des Terms auf der linken Seite dieser Gleichung, mit der Summe aus der Zeitableitung eines skalaren Feldes und dem Gradienten eines Vektorfeldes, ist ähnlich jener der Kontinuitätsgleichung (IX.7), d.h. allgemeiner einer lokalen Bilanzgleichung. Dies weist darauf hin, dass

$$e_{\text{e.m.}}(t, \vec{r}) \equiv \frac{\epsilon_0 \vec{E}(t, \vec{r})^2}{2} + \frac{\vec{B}(t, \vec{r})^2}{2\mu_0} \quad (\text{IX.23a})$$

die Dichte einer Größe darstellt, während der *Poynting*^(ao)-Vektor

$$\vec{S}_{\text{e.m.}}(t, \vec{r}) \equiv \frac{\vec{E}(t, \vec{r}) \times \vec{B}(t, \vec{r})}{\mu_0} \quad (\text{IX.23b})$$

die assoziierte Stromdichte ist. Da die Dichte $e_{\text{e.m.}}(t, \vec{r})$ im stationären Fall ohne magnetisches Feld mit der elektrostatischen Energiedichte [vgl. Gl. (VII.17a)] übereinstimmt, möchte man sie gerne als *Energiedichte* des elektromagnetischen Feldes interpretieren. Dementsprechend wäre $\vec{S}_{\text{e.m.}}(t, \vec{r})$ die zugehörige *Energiestromdichte* und die aus Gl. (IX.22) folgende Gleichung

$$\frac{\partial e_{\text{e.m.}}(t, \vec{r})}{\partial t} + \vec{\nabla} \cdot \vec{S}_{\text{e.m.}}(t, \vec{r}) = -\vec{j}_{\text{el.}}(t, \vec{r}) \cdot \vec{E}(t, \vec{r}). \quad (\text{IX.23c})$$

sollte eine (lokale) Bilanzgleichung für die Energie darstellen.

Bemerkungen:

* Die Gleichung (IX.23c) wird auch (differentielle Form von dem) *Satz von Poynting* genannt — wobei die Integralform die Gl. (IX.24) unten ist.

* Der zweite Beitrag zur Energiedichte, $\vec{B}(t, \vec{r})^2/2\mu_0$, kann als die im magnetischen Feld gespeicherte Energiedichte betrachtet werden.

^(ao)J. H. POYNTING, 1852–1914

IX.3.1b Interpretation der Energiebilanzgleichung

Sei \mathcal{V} ein Raumbereich mit Oberfläche $\partial\mathcal{V}$. Die Integration der Bilanzgleichung (IX.23c) über diesen Bereich ergibt

$$\frac{d}{dt} \int_{\mathcal{V}} e_{\text{e.m.}}(t, \vec{r}) d^3\vec{r} + \oint_{\partial\mathcal{V}} \vec{S}_{\text{e.m.}}(t, \vec{r}) \cdot d^2\vec{S} = - \int_{\mathcal{V}} \vec{j}_{\text{el.}}(t, \vec{r}) \cdot \vec{E}(t, \vec{r}) d^3\vec{r}, \quad (\text{IX.24})$$

wobei der Integralsatz von Gauß benutzt wurde, um den zweiten Term umzuschreiben.

Zur Deutung dieser Gleichung, insbesondere des Terms auf deren rechten Seite, kann man einen elektrischen Strom bestehend aus bewegten Punktladungen betrachten, entsprechend der Ladungsstromdichte $\sum_a q_a \vec{v}_a(t) \delta^{(3)}(\vec{r} - \vec{r}_a)$. Mit dieser Stromdichte gilt

$$- \int_{\mathcal{V}} \vec{j}_{\text{el.}}(t, \vec{r}) \cdot \vec{E}(t, \vec{r}) d^3\vec{r} = - \sum'_a q_a \vec{v}_a(t) \cdot \vec{E}(t, \vec{r}_a),$$

wobei der Strich neben dem Summenzeichen bedeutet, dass die Summe nur über die Punktladungen läuft, die sich im Bereich \mathcal{V} befinden.

Die gesamte kinetische Energie dieser Punktladungen zur Zeit t ist

$$E_{\text{mat.}}(t) = \sum'_a \frac{m_a \vec{v}_a(t)^2}{2}.$$

Daraus folgt

$$\frac{dE_{\text{mat.}}(t)}{dt} = \sum'_a \vec{v}_a(t) \cdot \left[m_a \frac{d\vec{v}_a(t)}{dt} \right].$$

Laut dem zweiten Newtonschen Gesetz (I.14) ist der Term in eckigen Klammern genau gleich der Kraft auf die Punktladung a . Bei der letzteren handelt es sich um die durch das elektromagnetische Feld verursachte Lorentz-Kraft, d.h.

$$\frac{dE_{\text{mat.}}(t)}{dt} = \sum'_a \vec{v}_a(t) \cdot \left\{ q_a \left[\vec{E}(t, \vec{r}_a) + \vec{v}_a(t) \times \vec{B}(t, \vec{r}_a) \right] \right\}.$$

Das Spatprodukt $\vec{v}_a \cdot (\vec{v}_a \times \vec{B})$ auf der rechten Seite verschwindet, so dass

$$\frac{dE_{\text{mat.}}(t)}{dt} = \sum'_a q_a \vec{v}_a(t) \cdot \vec{E}(t, \vec{r}_a)$$

übrig bleibt. Definiert man jetzt

$$E_{\text{e.m.}}(t) \equiv \int_{\mathcal{V}} e_{\text{e.m.}}(t, \vec{r}) d^3\vec{r}, \quad (\text{IX.25a})$$

die sich als Energie des elektromagnetischen Feldes in \mathcal{V} zur Zeit t interpretieren lässt, im Einklang mit der Deutung von $e_{\text{e.m.}}$, so wird die Gleichung (IX.24) zu

$$\frac{dE_{\text{e.m.}}(t)}{dt} = - \oint_{\partial\mathcal{V}} \vec{S}_{\text{e.m.}}(t, \vec{r}) \cdot d^2\vec{S} - \frac{dE_{\text{mat.}}(t)}{dt}. \quad (\text{IX.25b})$$

Die Interpretation dieser Gleichung ist ziemlich klar: die (Rate der) Änderung der Feldenergie im Bereich \mathcal{V} — Term auf der linken Seite — besteht aus der Energie, die durch die Oberfläche $\partial\mathcal{V}$ nach außen fließt — erster Term im rechten Glied —, und der Energie, die auf die Ladungen übertragen wird.

Somit stellt wirklich Gl. (IX.23c) eine lokale Bilanzgleichung für die Energie dar, in der die in Gl. (IX.23a) und (IX.23b) definierten $e_{\text{e.m.}}$ und $\vec{S}_{\text{e.m.}}$ die Energiedichte und -stromdichte des elektromagnetischen Feldes sind.

IX.3.2 Impulsdichte und -stromdichte des elektromagnetischen Feldes

Ausgehend aus den Maxwell-Gleichungen lässt sich eine weitere Bilanzgleichung herleiten, wobei die „erhaltene“ Größe jetzt vektoriell ist, und kann als Impuls des elektromagnetischen Feldes interpretiert werden.

Man definiert das Vektorfeld

$$\vec{g}_{\text{e.m.}}(t, \vec{r}) \equiv \epsilon_0 \vec{E}(t, \vec{r}) \times \vec{B}(t, \vec{r}), \quad (\text{IX.26a})$$

das hiernach als *Impulsdichte* interpretiert wird, und den *Maxwellschen Spannungstensor* $\sigma_{\text{e.m.}}$ mit kartesischen Komponenten⁽⁷⁶⁾

$$\sigma_{\text{e.m.}}^{ij}(t, \vec{r}) \equiv \epsilon_0 \left[E^i(t, \vec{r}) E^j(t, \vec{r}) + c^2 B^i(t, \vec{r}) B^j(t, \vec{r}) - \frac{1}{2} \delta^{ij} \left[\vec{E}(t, \vec{r})^2 + c^2 \vec{B}(t, \vec{r})^2 \right] \right]. \quad (\text{IX.26b})$$

Dazu kann man noch den Tensor $\mathbf{T}_{\text{e.m.}} \equiv -\sigma_{\text{e.m.}}$ mit Komponenten

$$T_{\text{e.m.}}^{ij}(t, \vec{r}) \equiv -\sigma_{\text{e.m.}}^{ij}(t, \vec{r}), \quad (\text{IX.26c})$$

eingeführen: die Komponente ij dieses Tensors wird die Dichte des Stroms in Richtung i von der j -ten Komponente des Impulses vom elektromagnetischen Feld modellieren.

Dann gilt die Bilanzgleichung

$$\frac{\partial g_{\text{e.m.}}^j(t, \vec{r})}{\partial t} + \sum_{i=1}^3 \frac{\partial T_{\text{e.m.}}^{ij}(t, \vec{r})}{\partial x^i} = -f_{\text{L}}^j(t, \vec{r}) \quad \text{für } j = 1, 2, 3, \quad (\text{IX.26d})$$

mit der j -ten Komponente $f_{\text{L}}^j(t, \vec{r})$ der Lorentz-Kraftdichte $\vec{f}_{\text{L}} = \rho_{\text{el.}} \vec{E} + \vec{j}_{\text{el.}} \times \vec{B}$ auf die Ladungen und Ströme, die das elektromagnetische Feld verursachen. Da diese Kraftdichte die Dichte des Impulses, der pro Zeiteinheit auf die Ladungen übertragen wird, darstellt — laut dem zweiten newtonschen Gesetz ist nämlich die Kraft gleich der Rate der Impulsänderung —, hat die Gleichung (IX.26d) genau die gleiche Form wie die Energiebilanzgleichung (IX.23c): auf der linken Seite steht die Summe aus der Zeitableitung von Dichte und der Divergenz der Stromdichte, die nur das elektromagnetische Feld betreffen; auf der rechten Seite steht ein Term, der die Wechselwirkung zwischen dem Feld und den Quellen berücksichtigt.

Herleitung der Gl. (IX.26d): der Kürze halber werden die Variablen (t, \vec{r}) durchaus weggelassen. Unter Verwendung der Maxwell–Gauß und der Maxwell–Ampère-Gleichungen (IX.1a), (IX.1d) lässt sich die Lorentz-Kraftdichte als

$$\vec{f}_{\text{L}} = \rho_{\text{el.}} \vec{E} + \vec{j}_{\text{el.}} \times \vec{B} = \epsilon_0 (\vec{\nabla} \cdot \vec{E}) \vec{E} + \left(\frac{1}{\mu_0} \vec{\nabla} \times \vec{B} - \epsilon_0 \frac{\partial \vec{E}}{\partial t} \right) \times \vec{B}$$

umschreiben. Dabei kann der letzte Term gemäß

$$\frac{\partial \vec{E}}{\partial t} \times \vec{B} = \frac{\partial}{\partial t} (\vec{E} \times \vec{B}) - \vec{E} \times \frac{\partial \vec{B}}{\partial t} = \frac{\partial}{\partial t} (\vec{E} \times \vec{B}) + \vec{E} \times (\vec{\nabla} \times \vec{E})$$

transformiert werden, wobei die letzte Gleichung aus der Maxwell–Faraday (IX.1c) folgt. Somit gilt

$$\vec{f}_{\text{L}} = \epsilon_0 \left[(\vec{\nabla} \cdot \vec{E}) \vec{E} - \vec{E} \times (\vec{\nabla} \times \vec{E}) \right] - \frac{1}{\mu_0} \vec{B} \times (\vec{\nabla} \times \vec{B}) - \frac{\partial \vec{g}_{\text{e.m.}}}{\partial t},$$

wobei Definition (IX.26a) benutzt wurde. Mithilfe der Maxwell–Thomson-Gleichung (IX.1b)

⁽⁷⁶⁾ Der Maxwellsche Spannungstensor wird manchmal mit der umgekehrten Zeichenkonvention, also wie der Tensor $\mathbf{T}_{\text{e.m.}}$, definiert, z.B. in den frühen Auflagen (vor etwa 1985) von Landau & Lifschitz [14, 28]. Die hier verwendete Konvention ist also die der späteren Auflagen von Refs. [14, 28] sowie von Jackson [11, 12, Gl. (6.120)].

kann man die Symmetrie der Terme mit den \vec{E} und \vec{B} Feldern erhöhen:

$$\vec{f}_L = \epsilon_0 \left[(\vec{\nabla} \cdot \vec{E}) \vec{E} - \vec{E} \times (\vec{\nabla} \times \vec{E}) \right] + \frac{1}{\mu_0} \left[(\vec{\nabla} \cdot \vec{B}) \vec{B} - \vec{B} \times (\vec{\nabla} \times \vec{B}) \right] - \frac{\partial \vec{g}_{\text{e.m.}}}{\partial t}.$$

Betrachte man die i -te Komponente dieser Gleichung. Für die zwei Terme der Form $\vec{V} \times (\vec{\nabla} \times \vec{V})$ gilt⁽⁷⁷⁾

$$\begin{aligned} \left[\vec{V} \times (\vec{\nabla} \times \vec{V}) \right]^i &= \sum_{j,k=1}^3 \epsilon^{ijk} V^j \sum_{l,m=1}^3 \epsilon^{klm} \frac{\partial V^m}{\partial x^l} = \sum_{j,l,m=1}^3 (\delta^{il} \delta^{jm} - \delta^{im} \delta^{jl}) V^j \frac{\partial V^m}{\partial x^l} \\ &= \sum_{j=1}^3 V^j \left(\frac{\partial V^j}{\partial x^i} - \frac{\partial V^i}{\partial x^j} \right) = \frac{\partial}{\partial x^i} \left(\frac{\vec{V}^2}{2} \right) - \sum_{j=1}^3 V^j \frac{\partial V^j}{\partial x^i}. \end{aligned}$$

Nach Abziehen der i -ten Komponente von $(\vec{\nabla} \cdot \vec{V}) \vec{V}$ kommt

$$\begin{aligned} \left[\vec{V} \times (\vec{\nabla} \times \vec{V}) - (\vec{\nabla} \cdot \vec{V}) \vec{V} \right]^i &= \frac{\partial}{\partial x^i} \left(\frac{\vec{V}^2}{2} \right) - \sum_{j=1}^3 V^j \frac{\partial V^j}{\partial x^i} - \sum_{j=1}^3 V^j \frac{\partial V^j}{\partial x^i} \\ &= \frac{\partial}{\partial x^i} \left(\frac{\vec{V}^2}{2} \right) - \sum_{j=1}^3 \frac{\partial}{\partial x^j} (V^i V^j). \end{aligned}$$

Daher lautet die i -te Komponente der Lorentz-Kraftdichte

$$\begin{aligned} f_L^i &= \epsilon_0 \left[\sum_{j=1}^3 \frac{\partial}{\partial x^j} (E^i E^j) - \frac{\partial}{\partial x^i} \left(\frac{\vec{E}^2}{2} \right) \right] + \frac{1}{\mu_0} \left[\sum_{j=1}^3 \frac{\partial}{\partial x^j} (B^i B^j) - \frac{\partial}{\partial x^i} \left(\frac{\vec{B}^2}{2} \right) \right] - \frac{\partial g_{\text{e.m.}}}{\partial t} \\ &= \epsilon_0 \sum_{j=1}^3 \frac{\partial}{\partial x^j} \left[E^i E^j - \delta^{ij} \frac{\vec{E}^2}{2} \right] + \frac{1}{\mu_0} \sum_{j=1}^3 \frac{\partial}{\partial x^j} \left[B^i B^j - \delta^{ij} \frac{\vec{B}^2}{2} \right] - \frac{\partial g_{\text{e.m.}}}{\partial t}. \end{aligned}$$

Unter Berücksichtigung der Identität $\epsilon_0 \mu_0 c^2 = 1$ erkennt man den Maxwell'schen Spannungstensor

$$f_L^i = \sum_{j=1}^3 \frac{\partial}{\partial x^j} \sigma_{\text{e.m.}}^{ij} - \frac{\partial g_{\text{e.m.}}}{\partial t} = - \sum_{j=1}^3 \frac{\partial T_{\text{e.m.}}^{ij}}{\partial x^j} - \frac{\partial g_{\text{e.m.}}}{\partial t}.$$

Wegen der Symmetrie von $T_{\text{e.m.}}^{ij}$ unter dem Austausch der beiden Indizes ist diese Gleichung, nach Umbenennung der Indizes $i \leftrightarrow j$, genau die Bilanzgleichung (IX.26d). \square

Bemerkungen:

* Offensichtlich hängt die Impulsdichte (IX.26b) mit dem Poynting-Vektor (IX.23b) einfach über

$$\vec{g}_{\text{e.m.}}(t, \vec{r}) = \frac{1}{c^2} \vec{S}_{\text{e.m.}}(t, \vec{r}). \quad (\text{IX.27})$$

zusammen.

* Die Spur des Maxwell'schen Spannungstensor bzw. des Tensors $\mathbf{T}_{\text{e.m.}}$ und die Energiedichte (IX.23a) genügen der einfachen Beziehung

$$\text{Tr } \mathbf{T}_{\text{e.m.}}(t, \vec{r}) = T_{\text{e.m.}}^{11}(t, \vec{r}) + T_{\text{e.m.}}^{22}(t, \vec{r}) + T_{\text{e.m.}}^{33}(t, \vec{r}) = e_{\text{e.m.}}(t, \vec{r}). \quad (\text{IX.28})$$

⁽⁷⁷⁾In vektorieller Form lautet diese Gleichung $\vec{V} \times (\vec{\nabla} \times \vec{V}) = \vec{\nabla} \left(\frac{\vec{V}^2}{2} \right) - (\vec{V} \cdot \vec{\nabla}) \vec{V}$.

IX.4 Elektromagnetische Wellen im Vakuum

In Abwesenheit von Quellen, $\rho_{\text{el.}} = 0$ und $\vec{j}_{\text{el.}} = \vec{0}$, nehmen die Bewegungsgleichungen (IX.9) und (IX.11) für die elektromagnetischen Felder oder (IX.21) für die Potentiale in Lorenz-Eichung die gleiche Form an, und zwar

$$\text{im Vakuum} \quad \begin{cases} \square \vec{E}(t, \vec{r}) = \vec{0} & \text{und} & \square \vec{B}(t, \vec{r}) = \vec{0}; \\ \square \Phi(t, \vec{r}) = 0 & \text{und} & \square \vec{A}(t, \vec{r}) = \vec{0} \quad (\text{in Lorenz-Eichung}). \end{cases} \quad \begin{array}{l} \text{(IX.29a)} \\ \text{(IX.29b)} \end{array}$$

Zur Umschreibung einiger Gleichungen führt man eine positive Zahl c gemäß

$$c^2 \equiv \frac{1}{\epsilon_0 \mu_0} \quad \text{(IX.29c)}$$

ein. Dann wird der D'Alembert-Operator (IX.10) zu

$$\square \equiv -\frac{1}{c^2} \frac{\partial^2}{\partial t^2} + \Delta. \quad \text{(IX.29d)}$$

Die Bewegungsgleichungen (IX.29) sind alle der gleichen Form $\square f(t, \vec{r}) = 0$, mit einer skalaren oder vektoriellen Funktion f . Die Lösungen dieser Differentialgleichung — die auch in anderen Bereichen der Physik auftritt — werden zunächst in § IX.4.1 vorgestellt. Dann werden die Ergebnisse auf den Fall des elektromagnetischen Feldes im Vakuum angewandt, wobei die Eigenschaften des Feldes einer Welle präzisiert werden (§ IX.4.2).

IX.4.1 Klassische Wellengleichung

Dieses Paragraph befasst sich mit der Herleitung der (genügend regulären) reellen Lösungen der (*klassischen*) *Wellengleichung*⁽⁷⁸⁾ oder *D'Alembert-Gleichung*

$$\square f(t, \vec{r}) = 0 \quad \text{(IX.30)}$$

für skalare Funktionen f , wobei der D'Alembert-Operator durch Gl. (IX.29d) definiert ist.

IX.4.1 a Ebene Wellen

Als einfaches aber wichtiges Beispiel kann man zunächst eine Lösung betrachten, die von einer einzigen (kartesischen) Raumkoordinate $z = x^3$ abhängt: $f(t, z)$. Diese ist somit unabhängig von den dazu orthogonalen Koordinaten x^1, x^2 , d.h. sie ist konstant in der sog. *transversalen* (x^1, x^2) Ebene, weshalb sie als *ebene Welle* bezeichnet wird.

Unter dieser Annahme wird die klassische Wellengleichung (IX.30) zur (1 + 1)-dimensionalen Differentialgleichung

$$-\frac{1}{c^2} \frac{\partial^2 f(t, z)}{\partial t^2} + \frac{\partial^2 f(t, z)}{\partial z^2} = 0. \quad \text{(IX.31)}$$

Der Differentialoperator dieser Gleichung lässt sich als Produkt von Operatoren erster Ordnung schreiben:

$$\left(\frac{\partial}{\partial z} - \frac{1}{c} \frac{\partial}{\partial t} \right) \left(\frac{\partial}{\partial z} + \frac{1}{c} \frac{\partial}{\partial t} \right) f(t, z) = 0. \quad \text{(IX.32)}$$

⁽⁷⁸⁾Gleichung (IX.30) wird oft als *die* Wellengleichung bezeichnet. Man findet aber auch andere partielle Differentialgleichungen mit zeit- und ortsabhängigen Lösungen, die als Wellen bezeichnet werden — z.B. für Schwerewellen in Flüssigkeiten, Stoßwellen in Fluiden, oder Wellenfunktionen in der Quantenmechanik. Deshalb wird hier die Bezeichnung „klassische Wellengleichung“ verwendet.

Zur Lösung dieser partiellen Differentialgleichung lohnt es sich, die Variablenänderung

$$x^+ \equiv z + ct \quad , \quad x^- \equiv z - ct \quad (\text{IX.33a})$$

durchzuführen. Die entsprechende Rücktransformation zu den ursprünglichen Variablen lautet

$$t = \frac{x^+ - x^-}{2c} \quad , \quad x = \frac{x^+ + x^-}{2}. \quad (\text{IX.33b})$$

Die partiellen Ableitungen nach den neuen Variablen lassen sich mit Hilfe der Kettenregel durch die Ableitungen nach den alten Variablen ausdrücken, und zwar

$$\frac{\partial}{\partial x^+} = \frac{\partial t}{\partial x^+} \frac{\partial}{\partial t} + \frac{\partial z}{\partial x^+} \frac{\partial}{\partial z} = \frac{1}{2} \left(\frac{1}{c} \frac{\partial}{\partial t} + \frac{\partial}{\partial z} \right) \quad (\text{IX.34a})$$

und

$$\frac{\partial}{\partial x^-} = \frac{\partial t}{\partial x^-} \frac{\partial}{\partial t} + \frac{\partial z}{\partial x^-} \frac{\partial}{\partial z} = \frac{1}{2} \left(-\frac{1}{c} \frac{\partial}{\partial t} + \frac{\partial}{\partial z} \right). \quad (\text{IX.34b})$$

Dank diesen Ergebnissen lautet die Differentialgleichung (IX.32) in den neuen Variablen

$$\frac{1}{4} \frac{\partial}{\partial x^-} \frac{\partial}{\partial x^+} f(x^+, x^-) = 0,$$

d.h. noch

$$\frac{\partial^2 f(x^+, x^-)}{\partial x^- \partial x^+} = 0. \quad (\text{IX.35})$$

Die allgemeine Lösung dieser partiellen Differentialgleichung zweiter Ordnung ist der Form

$$f(x^+, x^-) = f_-(x^+) + f_+(x^-)$$

wobei f_- und f_+ zwei beliebige Funktionen einer einzigen reellen Variablen sind. Kommt man zurück zu den ursprünglichen Variablen t, z , so lautet die allgemeine Lösung der Gl. (IX.31)

$$f(t, z) = f_-(z + ct) + f_+(z - ct). \quad (\text{IX.36})$$

Dieses mathematische Ergebnis lässt sich einfach physikalisch interpretieren. Betrachte man beispielsweise die Funktion $f_+(z - ct)$: sie nimmt den gleichen Wert für alle Raumzeitpunkte (t, z) an, für welche $z - ct$ einen konstanten Wert hat. Somit bleibt das Profil entlang der z -Achse des durch f modellierten Signals zur Zeit $t = t_0$ global unverändert zu einem späteren Zeitpunkt t_1 . Es wird nur in positive z -Richtung um $\Delta z \equiv c(t_1 - t_0)$ verschoben, entsprechend einer Ausbreitung des Signals mit Geschwindigkeit $+c$.

Wiederum modelliert der Term f_- ein Signal, das sich in negative z -Richtung mit Ausbreitungsgeschwindigkeit c — will man die Richtung auch in der Geschwindigkeit berücksichtigen: $-c$ — fortpflanzt.

Solche sich ausbreitende Signale werden als *fortschreitende Wellen* bezeichnet.⁽⁷⁸⁾ Dann ist die allgemeine Lösung (IX.36) eine Superposition von rechts- und linkslaufenden ebenen Wellen.

Bemerkungen:

* Da die D'Alembert-Gleichung in einer räumlichen Dimension (IX.31) sowohl rechts- als linkslaufende Wellen beschreiben kann, wird sie manchmal *Zweiweg-Wellengleichung* genannt. Im Gegensatz dazu haben die Differentialgleichungen

$$\left(\frac{\partial}{\partial z} - \frac{1}{c} \frac{\partial}{\partial t} \right) f(t, z) = 0 \quad \text{bzw.} \quad \left(\frac{\partial}{\partial z} + \frac{1}{c} \frac{\partial}{\partial t} \right) f(t, z) = 0 \quad (\text{IX.37})$$

die allgemeinen Lösungen $f(t, z) = f_-(z + ct)$ bzw. $f(t, z) = f_+(z - ct)$, d.h. fortschreitende Wellen, die nur nach links bzw. rechts propagieren. Dementsprechend werden die Gleichungen (IX.37) *Einweg-Wellengleichungen* (in einer Dimension) genannt.

* Im Fall mit $f_- = f_+$ in der Lösung (IX.36) breitet sich die resultierende Welle f nicht mehr aus: es handelt sich um eine *stehende Welle*.

IX.4.1 b Allgemeine Lösung

Um die allgemeine Lösung der klassischen Wellengleichung (IX.30) herzuleiten, ist es günstig, die Fourier-Darstellung der räumlichen Abhängigkeit von $f(t, \vec{r})$ einzuführen. Somit schreibt man

$$f(t, \vec{r}) = \int_{\mathbb{R}^3} \tilde{f}(t, \vec{k}) e^{i\vec{k}\cdot\vec{r}} \frac{d^3\vec{k}}{(2\pi)^3}, \quad (\text{IX.38a})$$

wobei die räumlich Fourier-transformierte Funktion $\tilde{f}(t, \vec{k})$ durch

$$\tilde{f}(t, \vec{k}) = \int_{\mathbb{R}^3} f(t, \vec{r}) e^{-i\vec{k}\cdot\vec{r}} d^3\vec{r} \quad (\text{IX.38b})$$

gegeben ist.

Unter Verwendung der Identität $\vec{\nabla}(e^{i\vec{k}\cdot\vec{r}}) = i\vec{k} e^{i\vec{k}\cdot\vec{r}}$ wird die klassische Wellengleichung (IX.30) angewandt auf die Darstellung (IX.38a) zu

$$\int \left[-\frac{1}{c^2} \frac{\partial^2 \tilde{f}(t, \vec{k})}{\partial t^2} - \vec{k}^2 \tilde{f}(t, \vec{k}) \right] e^{i\vec{k}\cdot\vec{r}} \frac{d^3\vec{k}}{(2\pi)^3} = 0,$$

wobei Integration über \vec{k} und Ableitung nach der Zeit t oder nach den Ortskoordinaten \vec{r} ausgetauscht wurden. Nach inverser Fourier-Transformation soll der Term in den eckigen Klammern verschwinden. Definiert man

$$\omega_{\vec{k}} \equiv c|\vec{k}|, \quad (\text{IX.39})$$

so ergibt sich für jeden *Wellenvektor* \vec{k} die Differentialgleichung

$$\frac{\partial^2 \tilde{f}(t, \vec{k})}{\partial t^2} + \omega_{\vec{k}}^2 \tilde{f}(t, \vec{k}) = 0. \quad (\text{IX.40a})$$

Man erkennt die Bewegungsgleichung eines harmonischen Oszillators mit Kreisfrequenz $\omega_{\vec{k}}$. Deren allgemeine Lösung ist der Form

$$\tilde{f}(t, \vec{k}) = \tilde{f}_+(\vec{k}) e^{-i\omega_{\vec{k}}t} + \tilde{f}_-(\vec{k}) e^{i\omega_{\vec{k}}t} \quad (\text{IX.40b})$$

mit zwei beliebigen Funktionen $\tilde{f}_+(\vec{k})$ und $\tilde{f}_-(\vec{k})$.

Wenn die gesuchte Lösung $f(t, \vec{r})$ reellwertig sein soll, stehen $\tilde{f}_+(\vec{k})$ und $\tilde{f}_-(\vec{k})$ in Zusammenhang miteinander. Die komplexe Konjugation der Beziehung (IX.38a) lautet

$$f(t, \vec{r})^* = \int_{\mathbb{R}^3} [\tilde{f}(t, \vec{k})]^* e^{-i\vec{k}\cdot\vec{r}} \frac{d^3\vec{k}}{(2\pi)^3} = \int_{\mathbb{R}^3} [\tilde{f}(t, -\vec{k})]^* e^{i\vec{k}\cdot\vec{r}} \frac{d^3\vec{k}}{(2\pi)^3},$$

wobei die zweite Gleichung aus der Substitution $\vec{k} \rightarrow -\vec{k}$ folgt. $f(t, \vec{r})$ nimmt reelle Werte an, wenn $f(t, \vec{r}) = f(t, \vec{r})^*$, d.h. nach Vergleich des letzten Terms in der obigen Gleichung mit dem Term auf der rechten Seite von Gl. (IX.38a), wenn

$$\tilde{f}(t, \vec{k}) = \tilde{f}(t, -\vec{k})^*$$

für alle t und \vec{k} . Angewandt auf die Lösung (IX.40b) lautet diese Bedingung

$$\tilde{f}_+(\vec{k}) e^{-i\omega_{\vec{k}}t} + \tilde{f}_-(\vec{k}) e^{i\omega_{\vec{k}}t} = [\tilde{f}_+(-\vec{k})]^* e^{i\omega_{\vec{k}}t} + [\tilde{f}_-(-\vec{k})]^* e^{-i\omega_{\vec{k}}t},$$

d.h. nach Identifizierung der Terme in $e^{i\omega_{\vec{k}}t}$

$$\tilde{f}_-(\vec{k}) = [\tilde{f}_+(-\vec{k})]^* \quad \forall \vec{k}. \quad (\text{IX.41})$$

Die Terme in $e^{-i\omega_{\vec{k}}t}$ führen zur gleichen Bedingung.

Setzt man schließlich die Lösung (IX.40b) der Gl. (IX.40a) in die Fourier-Darstellung (IX.38a) ein, so lautet die letztere

$$f(t, \vec{r}) = \int_{\mathbb{R}^3} [\tilde{f}_+(\vec{k}) e^{-i\omega_{\vec{k}}t} + \tilde{f}_-(\vec{k}) e^{i\omega_{\vec{k}}t}] e^{i\vec{k}\cdot\vec{r}} \frac{d^3\vec{k}}{(2\pi)^3} = \int_{\mathbb{R}^3} [\tilde{f}_+(\vec{k}) e^{-i(\omega_{\vec{k}}t - \vec{k}\cdot\vec{r})} + \tilde{f}_-(-\vec{k}) e^{i(\omega_{\vec{k}}t - \vec{k}\cdot\vec{r})}] \frac{d^3\vec{k}}{(2\pi)^3}.$$

Dabei wurde im zweiten Summanden des Integranden des letzten Integrals die Substitution $\vec{k} \rightarrow -\vec{k}$ unter Berücksichtigung von $\omega_{-\vec{k}} = \omega_{\vec{k}}$ durchgeführt. Ersetzt man in jenem Term $\tilde{f}_-(-\vec{k})$ durch $[\tilde{f}_+(\vec{k})]^*$, wie aus Bedingung (IX.41) folgt, so sieht man, dass der Term genau komplex konjugiert zum ersten Summanden ist. Unter Einführung der Notation $a(\vec{k}) \equiv 2\tilde{f}_+(\vec{k})$ ergibt sich somit für die allgemeine reelle Lösung der klassischen Wellengleichung (IX.30)

$$f(t, \vec{r}) = \operatorname{Re} \left[\int_{\mathbb{R}^3} a(\vec{k}) e^{-i(\omega_{\vec{k}}t - \vec{k}\cdot\vec{r})} \frac{d^3\vec{k}}{(2\pi)^3} \right]. \quad (\text{IX.42})$$

Somit lässt sich die allgemeine Lösung als Superposition von (unendlich vielen) ebenen Wellen $a(\vec{k}) e^{-i(\omega_{\vec{k}}t - \vec{k}\cdot\vec{r})}$ schreiben. Die Beziehung (IX.39) zwischen Kreisfrequenz $\omega_{\vec{k}}$ und Wellenvektor \vec{k} dieser ebenen Wellen heißt *Dispersionsrelation*.

Das Verhältnis $c_\varphi(\vec{k}) \equiv \omega_{\vec{k}}/|\vec{k}|$ ist die *Phasengeschwindigkeit* der Welle mit Wellenvektor \vec{k} — wobei im Fall der klassischen Wellengleichung (IX.30) $c_\varphi(\vec{k}) = c$ für alle \vec{k} .

Bemerkung: Die komplexe Amplitude $a(\vec{k}) \in \mathbb{C}$ lässt sich prinzipiell durch die Angabe von Anfangsbedingungen $f(t=0, \vec{r})$ und $\partial f(t=0, \vec{r})/\partial t$ festlegen.

IX.4.2 Elektromagnetische Wellen

IX.4.2a Elektromagnetische Potentiale

Gemäß den Ergebnissen des vorigen § IX.4.1 lauten die allgemeinen Lösungen der Bewegungsgleichungen (IX.29b) für die elektromagnetischen Potentiale im Vakuum

$$\Phi(t, \vec{r}) = \operatorname{Re} \left[\int a(\vec{k}) e^{-i(\omega_{\vec{k}}t - \vec{k}\cdot\vec{r})} \frac{d^3\vec{k}}{(2\pi)^3} \right] \quad (\text{IX.43a})$$

und

$$\vec{A}(t, \vec{r}) = \operatorname{Re} \left[\int \vec{b}(\vec{k}) e^{-i(\omega_{\vec{k}}t - \vec{k}\cdot\vec{r})} \frac{d^3\vec{k}}{(2\pi)^3} \right] \quad (\text{IX.43b})$$

mit beliebigen $a(\vec{k}) \in \mathbb{C}$ und $\vec{b}(\vec{k}) \in \mathbb{C}^3$ und der Dispersionsrelation

$$\omega_{\vec{k}} \equiv c|\vec{k}|. \quad (\text{IX.43c})$$

Dabei sollen Φ und \vec{A} die Lorenz-Eichbedingung (IX.16) erfüllen:

$$\frac{1}{c^2} \frac{\partial \Phi(t, \vec{r})}{\partial t} + \vec{\nabla} \cdot \vec{A}(t, \vec{r}) = \operatorname{Re} \left(\int \left[-\frac{i\omega_{\vec{k}}}{c^2} a(\vec{k}) + i\vec{k} \cdot \vec{b}(\vec{k}) \right] e^{-i(\omega_{\vec{k}}t - \vec{k}\cdot\vec{r})} \frac{d^3\vec{k}}{(2\pi)^3} \right) = 0.$$

Die elektromagnetischen Potentiale Φ und \vec{A} können noch über die Eichtransformation (IX.14) durch äquivalente Potentiale Φ' und \vec{A}' ersetzt werden, die ebenfalls der Lorenz-Bedingung genügen, solange die Funktion χ der Transformation die Gleichung (IX.17) — d.h. die klassische Wellengleichung — erfüllt. Mit

$$\chi(t, \vec{r}) = \operatorname{Re} \left[\int d(\vec{k}) e^{-i(\omega_{\vec{k}}t - \vec{k}\cdot\vec{r})} \frac{d^3\vec{k}}{(2\pi)^3} \right]$$

gelten

$$\Phi'(t, \vec{r}) = \Phi(t, \vec{r}) - \frac{\partial \chi(t, \vec{r})}{\partial t} = \operatorname{Re} \left(\int \left[a(\vec{k}) + i\omega_{\vec{k}} d(\vec{k}) \right] e^{-i(\omega_{\vec{k}}t - \vec{k}\cdot\vec{r})} \frac{d^3\vec{k}}{(2\pi)^3} \right)$$

und

$$\vec{A}'(t, \vec{r}) = \vec{A}(t, \vec{r}) + \vec{\nabla} \chi(t, \vec{r}) = \operatorname{Re} \left(\int \left[\vec{b}(\vec{k}) + i\vec{k} d(\vec{k}) \right] e^{-i(\omega_{\vec{k}}t - \vec{k}\cdot\vec{r})} \frac{d^3\vec{k}}{(2\pi)^3} \right).$$

Eine besonders geeignete Wahl ist $d(\vec{k}) = ia(\vec{k})/\omega_{\vec{k}}$ für jeden \vec{k} , die zum einfachen Skalarpotential

$$\Phi'(t, \vec{r}) = 0 \quad (\text{IX.44a})$$

führt. Wiederum lautet das entsprechende Vektorpotential

$$\vec{A}'(t, \vec{r}) = \text{Re} \left[\int \vec{\varepsilon}(\vec{k}) e^{-i(\omega_{\vec{k}}t - \vec{k} \cdot \vec{r})} \frac{d^3\vec{k}}{(2\pi)^3} \right], \quad (\text{IX.44b})$$

wobei der *Polarisationsvektor* $\vec{\varepsilon}(\vec{k})$ durch

$$\vec{\varepsilon}(\vec{k}) \equiv \vec{b}(\vec{k}) + i\vec{k}d(\vec{k}) = \vec{b}(\vec{k}) - \frac{a(\vec{k})}{\omega_{\vec{k}}} \vec{k}$$

gegeben ist. Dabei kann noch $\vec{\varepsilon}(\vec{k})$ reell oder komplex sein.

Eine Einschränkung über den Polarisationsvektor folgt aus der Eichbedingung. Mit $\Phi'(t, \vec{r}) = 0$ wird die Lorenz-Eichbedingung automatisch zur Coulomb-Eichbedingung (IX.15)

$$\vec{\nabla} \cdot \vec{A}'(t, \vec{r}) = \text{Re} \left[\int i \vec{\varepsilon}(\vec{k}) \cdot \vec{k} e^{-i(\omega_{\vec{k}}t - \vec{k} \cdot \vec{r})} \frac{d^3\vec{k}}{(2\pi)^3} \right] = 0.$$

Diese Bedingung ist nur dann erfüllt, wenn

$$\vec{\varepsilon}(\vec{k}) \cdot \vec{k} = 0 \quad (\text{IX.45})$$

für jeden Wellenvektor \vec{k} , d.h. wenn der Polarisationsvektor senkrecht auf die Ausbreitungsrichtung der entsprechenden ebenen Welle ist. Somit sind elektromagnetische Wellen im Vakuum *transversal polarisiert*.

IX.4.2b Elektrisches und magnetisches Feld

Setzt man die elektromagnetischen Potentiale (IX.44) in die Beziehungen (IX.12) und (IX.13) ein, so lauten die zugehörigen elektrischen und magnetischen Felder

$$\vec{E}(t, \vec{r}) = -\frac{\partial \vec{A}'(t, \vec{r})}{\partial t} = \text{Re} \left[\int i\omega_{\vec{k}} \vec{\varepsilon}(\vec{k}) e^{-i(\omega_{\vec{k}}t - \vec{k} \cdot \vec{r})} \frac{d^3\vec{k}}{(2\pi)^3} \right] \quad (\text{IX.46a})$$

$$\vec{B}(t, \vec{r}) = \vec{\nabla} \times \vec{A}'(t, \vec{r}) = \text{Re} \left[\int i\vec{k} \times \vec{\varepsilon}(\vec{k}) e^{-i(\omega_{\vec{k}}t - \vec{k} \cdot \vec{r})} \frac{d^3\vec{k}}{(2\pi)^3} \right]. \quad (\text{IX.46b})$$

Betrachte man eine *monochromatische* ebene Welle, d.h. eine Lösung mit nur einem einzigen Wellenvektor: $\vec{\varepsilon}(\vec{k}) = \mathcal{N} \vec{\varepsilon}(\vec{k}_0) (2\pi)^3 \delta^{(3)}(\vec{k} - \vec{k}_0)$, wobei \mathcal{N} eine unwesentliche Normierungskonstante bezeichnet. Dann gelten

$$\vec{E}(t, \vec{r}) = \text{Re} \left[i\omega_{\vec{k}_0} \vec{\varepsilon}(\vec{k}_0) e^{-i(\omega_{\vec{k}_0}t - \vec{k}_0 \cdot \vec{r})} \right] \quad \text{und} \quad \vec{B}(t, \vec{r}) = \text{Re} \left[i\vec{k}_0 \times \vec{\varepsilon}(\vec{k}_0) e^{-i(\omega_{\vec{k}_0}t - \vec{k}_0 \cdot \vec{r})} \right].$$

Auf diesen Feldern erkennt man die folgenden Eigenschaften. Erstens sind $\vec{E}(t, \vec{r})$ und $\vec{B}(t, \vec{r})$ automatisch senkrecht aufeinander

$$\vec{E}(t, \vec{r}) \cdot \vec{B}(t, \vec{r}) = 0. \quad (\text{IX.47})$$

Dann sind $\vec{E}(t, \vec{r})$ und $\vec{B}(t, \vec{r})$ beide senkrecht zur Bewegungsrichtung \vec{k}_0

$$\vec{k}_0 \cdot \vec{E}(t, \vec{r}) = \vec{k}_0 \cdot \vec{B}(t, \vec{r}) = 0. \quad (\text{IX.48})$$

Schließlich gilt dank $|\vec{k}_0| = \omega_{\vec{k}_0}/c$

$$|\vec{B}(t, \vec{r})| = \frac{|\vec{E}(t, \vec{r})|}{c}. \quad (\text{IX.49})$$

IX.4.2c Energie des elektromagnetischen Feldes

Betrachte man der Einfachheit halber die Felder (IX.46) für den Fall einer monochromatischen ebenen Welle mit Polarisationsvektor $\vec{\varepsilon}(\vec{k}) = (2\pi)^3 \delta^{(3)}(\vec{k} - \vec{k}_0) \vec{E}_0 / (i\omega_0)$ mit $\vec{E}_0 \in \mathbb{R}^3$ und $\omega_0 \equiv \omega_{\vec{k}_0}$. Da die Richtung des elektrischen Feldes — und daher auch der magnetischen Induktion — konstant bleibt, handelt es sich um eine *linear polarisierte* Welle. Es gelten dann

$$\vec{E}(t, \vec{r}) = \vec{E}_0 \cos(\omega_0 t - \vec{k}_0 \cdot \vec{r}) \quad (\text{IX.50a})$$

und

$$\vec{B}(t, \vec{r}) = \frac{\vec{\varepsilon}_{\vec{k}_0} \times \vec{E}_0}{c} \cos(\omega_0 t - \vec{k}_0 \cdot \vec{r}) = \frac{\vec{k}_0 \times \vec{E}_0}{\omega_0} \cos(\omega_0 t - \vec{k}_0 \cdot \vec{r}). \quad (\text{IX.50b})$$

Aus der Orthogonalität von \vec{k}_0 und \vec{E}_0 folgt $|\vec{\varepsilon}_{\vec{k}_0} \times \vec{E}_0| = |\vec{\varepsilon}_{\vec{k}_0}| |\vec{E}_0| = |\vec{E}_0|$, womit sich die Gleichung (IX.49) wiederfinden lässt.

Setzt man diese Felder in die Energiedichte (IX.23a) ein, so ergibt sich

$$e_{\text{e.m.}}(t, \vec{r}) = \left(\frac{\epsilon_0 \vec{E}_0^2}{2} + \frac{\vec{E}_0^2}{2\mu_0 c^2} \right) \cos^2(\omega_0 t - \vec{k}_0 \cdot \vec{r})$$

d.h. unter Verwendung der Beziehung $\epsilon_0 \mu_0 c^2 = 1$

$$e_{\text{e.m.}}(t, \vec{r}) = \epsilon_0 \vec{E}_0^2 \cos^2(\omega_0 t - \vec{k}_0 \cdot \vec{r}). \quad (\text{IX.51})$$

Interessant ist auch der über die Zeit gemittelte Wert — definiert für eine beliebige periodische Funktion f durch

$$\langle f \rangle \equiv \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} f(t) dt \quad (\text{IX.52})$$

mit der Periode \mathcal{T} der Funktion — von der Energiedichte. Im Fall der letzteren ist die Periodendauer $\mathcal{T} = 2\pi/\omega_0$: da der Mittelwert von \cos^2 über eine Periode $\frac{1}{2}$ ist, gilt

$$\langle e_{\text{e.m.}}(\vec{r}) \rangle = \frac{\epsilon_0 \vec{E}_0^2}{2}. \quad (\text{IX.53})$$

Diese zeitgemittelte Energiedichte ist auch unabhängig vom Ort.

Mit den Feldern (IX.50) wird der Poynting-Vektor (IX.23b) zu

$$\vec{S}_{\text{e.m.}}(t, \vec{r}) = \frac{\vec{E}_0 \times (\vec{\varepsilon}_{\vec{k}_0} \times \vec{E}_0)}{\mu_0 c} \cos^2(\omega_0 t - \vec{k}_0 \cdot \vec{r}).$$

Im doppelten Kreuzprodukt $\vec{E}_0 \times (\vec{\varepsilon}_{\vec{k}_0} \times \vec{E}_0) = (\vec{E}_0)^2 \vec{\varepsilon}_{\vec{k}_0} - (\vec{E}_0 \cdot \vec{\varepsilon}_{\vec{k}_0}) \vec{E}_0$ verschwindet der zweite Term wegen der Transversalität der elektromagnetischen Welle im Vakuum. Somit gilt

$$\vec{S}_{\text{e.m.}}(t, \vec{r}) = \frac{\vec{E}_0^2}{\mu_0 c} \cos^2(\omega_0 t - \vec{k}_0 \cdot \vec{r}) \vec{\varepsilon}_{\vec{k}_0} \quad (\text{IX.54})$$

gerichtet entlang der Ausbreitungsrichtung $\vec{\varepsilon}_{\vec{k}_0}$, wie es intuitiv der Fall sein soll.

Gemittelt über die Zeit ergibt sich

$$\langle \vec{S}_{\text{e.m.}}(\vec{r}) \rangle = \frac{\vec{E}_0^2}{2\mu_0 c} \vec{\varepsilon}_{\vec{k}_0} = \frac{\epsilon_0 \vec{E}_0^2}{2} c \vec{\varepsilon}_{\vec{k}_0},$$

d.h. nach Vergleich mit der zeitgemittelten Energiedichte (IX.53)

$$\langle \vec{S}_{\text{e.m.}}(\vec{r}) \rangle = \langle e_{\text{e.m.}}(\vec{r}) \rangle c \vec{\varepsilon}_{\vec{k}_0}. \quad (\text{IX.55})$$

Dies entspricht der Stromdichte assoziiert mit einer Dichte $\langle e_{\text{e.m.}}(\vec{r}) \rangle$, die sich mit der (gerichteten) Geschwindigkeit $c \vec{\varepsilon}_{\vec{k}_0}$ ausbreitet.

IX.5 Klassische Theorie der Strahlung

In diesem Abschnitt werden die Maxwell-Gleichungen in Anwesenheit gegebener äußerer Quellen anhand von retardierten Potentialen gelöst. Dafür wird zunächst die retardierte Greensche Funktion der D'Alembert-Gleichung eingeführt (§ IX.5.1). Mit deren Hilfe können die elektromagnetischen Potentiale für jede beliebige Ladungs- und Stromverteilung ausgedrückt werden (§ IX.5.2). Dann befasst sich § IX.5.3 mit der Herleitung einer Näherung der Potentiale, die in großer Entfernung von den Quellen gilt. Schließlich wird das durch ein bewegtes geladenes Punktteilchen erzeugte elektromagnetische Feld bestimmt (§ IX.5.4). Dabei wird insbesondere die durch eine beschleunigte Punktladung abgestrahlte Leistung berechnet.

IX.5.1 Greensche Funktion der klassischen Wellengleichung

Die Bewegungsgleichungen (IX.21) für die elektromagnetischen Potentiale in der Lorenz-Eichung sind lineare partielle Differentialgleichungen der Form

$$\square f(t, \vec{r}) = J(t, \vec{r}) \quad (\text{IX.56})$$

mit vorgegebenem rechtem Glied J , d.h. einer inhomogenen klassischen Wellengleichung (inhomogenen D'Alembert-Gleichung).

Die Lösung ist gleich der Summe aus der allgemeinen Lösung der assoziierten homogenen Differentialgleichung — die in § IX.4.1 b berechnet wurde — und aus einer speziellen Lösung, die hiernach bestimmt wird.

Sei $G(t, \vec{r}; t', \vec{r}')$ eine Greensche Funktion der Gleichung (IX.56), d.h. eine Lösung der Differentialgleichung

$$\square G(t, \vec{r}; t', \vec{r}') = \delta(t - t') \delta^{(3)}(\vec{r} - \vec{r}') \quad (\text{IX.57a})$$

genügen. Dann ist die Faltung von G und J eine spezielle Lösung der Gl. (IX.56):

$$f(t, \vec{r}) = \int G(t, \vec{r}; t', \vec{r}') J(t', \vec{r}') dt' d^3\vec{r}'. \quad (\text{IX.57b})$$

Wenn \square den d'Alembert-Operator bezüglich der nicht-gestrichenen Variablen t, \vec{r} bezeichnet, gilt tatsächlich

$$\begin{aligned} \square f(t, \vec{r}) &= \square \left[\int G(t, \vec{r}; t', \vec{r}') J(t', \vec{r}') dt' d^3\vec{r}' \right] = \int \square G(t, \vec{r}; t', \vec{r}') J(t', \vec{r}') dt' d^3\vec{r}' \\ &= \int \delta(t - t') \delta^{(3)}(\vec{r} - \vec{r}') J(t', \vec{r}') dt' d^3\vec{r}' = J(t, \vec{r}). \end{aligned}$$

Hiernach wird gezeigt, dass die Differentialgleichung (IX.57a) zwei unabhängige Lösungen hat, die im Unendlichen verschwinden, und zwar

$$G_{\text{ret.}}(t, \vec{r}; t', \vec{r}') \equiv \frac{-1}{4\pi|\vec{r} - \vec{r}'|} \delta\left(\frac{|\vec{r} - \vec{r}'|}{c} - (t - t')\right), \quad (\text{IX.58a})$$

$$G_{\text{adv.}}(t, \vec{r}; t', \vec{r}') \equiv \frac{-1}{4\pi|\vec{r} - \vec{r}'|} \delta\left(\frac{|\vec{r} - \vec{r}'|}{c} + (t - t')\right). \quad (\text{IX.58b})$$

- $G_{\text{ret.}}(t, \vec{r}; t', \vec{r}')$ ist die *retardierte Greensche Funktion*: liegt der „Beobachtungspunkt“ — wo die Lösung f von Gl. (IX.56) ausgewertet bzw. gemessen wird — im Raumzeitpunkt $P = (t, \vec{r})$, dann ist der Träger von $G_{\text{ret.}}$ lokalisiert auf der Menge der Punkte (t', \vec{r}') mit $|\vec{r} - \vec{r}'| = c(t - t')$ und somit $t' \leq t$: diese Punkte bilden den *Rückwärtslichtkegel* (oder *Vergangenheits-Lichtkegel*) des Punkts P .
- $G_{\text{adv.}}(t, \vec{r}; t', \vec{r}')$ ist die *avancierte Greensche Funktion*: $G_{\text{adv.}}$ ist lokalisiert auf dem Vorwärtslichtkegel $|\vec{r} - \vec{r}'| = -c(t - t')$, woraus $t' \geq t$ folgt, des Beobachtungspunktes (t, \vec{r}) .

Allgemeiner ist jede Funktion der Form

$$G(t, \vec{r}; t', \vec{r}') = \alpha G_{\text{ret.}}(t, \vec{r}; t', \vec{r}') + \beta G_{\text{adv.}}(t, \vec{r}; t', \vec{r}') \quad \text{mit } \alpha + \beta = 1$$

ebenfalls eine im Unendlichen verschwindende Greensche Funktion der inhomogenen klassischen Wellengleichung (IX.56).

Bemerkung: Bezeichnet man die retardierte bzw. avancierte Greensche Funktion $G_{\text{ret.}}$ bzw. $G_{\text{adv.}}$ mit G_- bzw. G_+ , so gilt

$$G_{\mp}(t, \vec{r}; t', \vec{r}') = \frac{-c}{2\pi} \Theta(\pm(t-t')) \delta(-c^2(t-t')^2 + |\vec{r}-\vec{r}'|^2),$$

wobei die einzige Rolle der Heaviside-Funktion Θ darin besteht, den Rückwärts- bzw. Vorwärtslichtkegel des Punkts (t, \vec{r}) auszuwählen.

Sei $\tau \equiv t - t'$ und $\varrho \equiv |\vec{r} - \vec{r}'|$. Aus $\delta(f(x)) = \sum_i \delta(x - x_i)/|f'(x_i)|$, wobei die Summe über die Nullstellen x_i der Funktion f läuft [Gl. (C.22)], folgt [vgl. auch Gl. (C.23)]

$$\frac{-c}{2\pi} \Theta(\pm\tau) \delta(-c^2\tau^2 + \varrho^2) = \frac{-c}{2\pi} \Theta(\pm\tau) \left[\frac{\delta(\varrho - c\tau)}{|2c\tau|} + \frac{\delta(\varrho + c\tau)}{|2c\tau|} \right] = \frac{-c}{4\pi\varrho} \delta(\varrho \mp c\tau) = \frac{-1}{4\pi\varrho} \delta\left(\frac{\varrho}{c} \mp \tau\right),$$

wobei $\varrho \geq 0$ und die Skalierungseigenschaft (C.17a) der δ -Distribution benutzt wurden. \square

Bestimmung der Greenschen Funktionen zur inhomogenen D'Alembert-Gleichung

Um eine Lösung der Gleichung (IX.57) zu finden, ist es günstig, die Fourier-Darstellungen

$$\delta(t - t') = \int e^{-i\omega(t-t')} \frac{d\omega}{2\pi}, \quad \delta^{(3)}(\vec{r} - \vec{r}') = \int e^{i\vec{k}\cdot(\vec{r}-\vec{r}')} \frac{d^3\vec{k}}{(2\pi)^3}$$

und

$$G(t, \vec{r}; t', \vec{r}') = \int \tilde{G}(\omega, \vec{k}) e^{-i[\omega(t-t') - \vec{k}\cdot(\vec{r}-\vec{r}')] } \frac{d\omega}{2\pi} \frac{d^3\vec{k}}{(2\pi)^3} \quad (\text{IX.59})$$

der δ -Distributionen und der gesuchten Greenschen Funktion einzuführen. Somit gilt

$$\begin{aligned} \square G(t, \vec{r}; t', \vec{r}') &= \int \tilde{G}(\omega, \vec{k}) \left[\left(-\frac{1}{c^2} \frac{\partial^2}{\partial t^2} + \Delta \right) e^{-i[\omega(t-t') - \vec{k}\cdot(\vec{r}-\vec{r}')] } \right] \frac{d\omega}{2\pi} \frac{d^3\vec{k}}{(2\pi)^3} \\ &= \int \left(\frac{\omega^2}{c^2} - \vec{k}^2 \right) \tilde{G}(\omega, \vec{k}) e^{-i[\omega(t-t') - \vec{k}\cdot(\vec{r}-\vec{r}')] } \frac{d\omega}{2\pi} \frac{d^3\vec{k}}{(2\pi)^3}. \end{aligned}$$

Dies muss gleich

$$\delta(t - t') \delta^{(3)}(\vec{r} - \vec{r}') = \int e^{-i[\omega(t-t') - \vec{k}\cdot(\vec{r}-\vec{r}')] } \frac{d\omega}{2\pi} \frac{d^3\vec{k}}{(2\pi)^3}$$

sein, woraus

$$\left(\frac{\omega^2}{c^2} - \vec{k}^2 \right) \tilde{G}(\omega, \vec{k}) = 1 \quad \text{bzw.} \quad \tilde{G}(\omega, \vec{k}) = \frac{-1}{\vec{k}^2 - \omega^2/c^2}$$

folgt. Nach Einsetzen in die Fourier-Darstellung (IX.59) ergibt sich dann

$$G(t, \vec{r}; t', \vec{r}') = - \int \frac{e^{-i[\omega(t-t') - \vec{k}\cdot(\vec{r}-\vec{r}')] } d\omega}{\vec{k}^2 - \omega^2/c^2} \frac{d^3\vec{k}}{2\pi (2\pi)^3}. \quad (\text{IX.60})$$

Dieser Ausdruck ist aber mehrdeutig, weil der Nenner des Integranden verschwindet.

Sei $g(k_0, \vec{R}) \equiv \int \frac{e^{i\vec{k}\cdot\vec{R}}}{\vec{k}^2 - k_0^2} d^3\vec{k}$. Unter Einführung des Winkels θ zwischen \vec{k} und \vec{R} ergibt sich

$$g(k_0, \vec{R}) = 2\pi \int_0^\infty k^2 \left[\int_{-1}^1 \frac{e^{ikR \cos \theta}}{k^2 - k_0^2} d(\cos \theta) \right] dk = \frac{2\pi}{iR} \int_0^\infty \frac{k}{k^2 - k_0^2} (e^{ikR} - e^{-ikR}) dk,$$

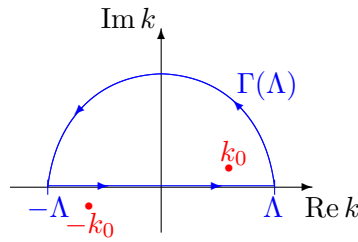


Abbildung IX.2 – Integrationskontour für die Berechnung von $\int_{-\infty}^{\infty} \frac{k e^{ikR}}{k^2 - k_0^2} dk$ mit $R > 0$.

wobei $R = |\vec{R}|$ und $k = |\vec{k}|$. Dies lautet noch

$$g(k_0, \vec{R}) = \frac{2\pi}{iR} \left(\int_0^{\infty} \frac{k e^{ikR}}{k^2 - k_0^2} dk - \int_0^{\infty} \frac{k e^{-ikR}}{k^2 - k_0^2} dk \right) = \frac{2\pi}{iR} \int_{-\infty}^{\infty} \frac{k e^{ikR}}{k^2 - k_0^2} dk,$$

wobei die letzte Gleichung aus der Substitution $k \rightarrow -k$ im zweiten Integral folgt.

Für $k_0 \in \mathbb{C}$ mit $\text{Im } k_0 \neq 0$ ist die Funktion $g(k_0, \vec{r})$ wohldefiniert. Wenn $\Gamma(\Lambda)$ die in Abb. IX.2 dargestellte Integrationskontour bezeichnet, dann ist für $R > 0$

$$\int_{-\infty}^{\infty} \frac{k e^{ikR}}{k^2 - k_0^2} dk = \lim_{\Lambda \rightarrow \infty} \int_{\Gamma(\Lambda)} \frac{k e^{ikR}}{k^2 - k_0^2} dk$$

laut dem Residuensatz gleich $2\pi i$ multipliziert mit dem Residuum von $k e^{ikR}/(k^2 - k_0^2)$ am Pol in der Halbebene $\text{Im } k > 0$. Für $\text{Im } k_0 > 0$ liegt dieser Pol bei $k = k_0$ und das Residuum ist $e^{ik_0 R}/2$. Für $\text{Im } k_0 < 0$ liegt der Pol bei $k = -k_0$, mit Residuum $e^{-ik_0 R}/2$. Somit gilt

$$g(k_0, \vec{R}) = \begin{cases} \frac{2\pi^2}{|\vec{R}|} e^{ik_0 |\vec{R}|} & \text{wenn } \text{Im } k_0 > 0, \\ \frac{2\pi^2}{|\vec{R}|} e^{-ik_0 |\vec{R}|} & \text{wenn } \text{Im } k_0 < 0. \end{cases}$$

Um den Fall $k_0 \in \mathbb{R}$ zu behandeln, werden die Funktionen

$$g_{\pm}(k_0, \vec{R}) \equiv \lim_{\epsilon \rightarrow 0^+} g(k_0 \pm i\epsilon, \vec{R}) = \frac{2\pi^2}{|\vec{R}|} e^{\pm ik_0 |\vec{R}|}$$

definiert. Mit deren Hilfe lassen sich aus Gl. (IX.60) zwei Greensche Funktionen herleiten, und zwar

$$G_{\pm}(t, \vec{r}; t', \vec{r}') = -\frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} g_{\pm}\left(\frac{\omega}{c}, \vec{r} - \vec{r}'\right) e^{-i\omega(t-t')} \frac{d\omega}{2\pi} = \frac{-1}{4\pi|\vec{r} - \vec{r}'|} \int_{-\infty}^{\infty} e^{-i\omega(t-t' \mp |\vec{r} - \vec{r}'|/c)} \frac{d\omega}{2\pi},$$

d.h.

$$G_{\pm}(t, \vec{r}; t', \vec{r}') = \frac{-1}{4\pi|\vec{r} - \vec{r}'|} \delta\left(\frac{|\vec{r} - \vec{r}'|}{c} \mp (t - t')\right), \quad (\text{IX.61})$$

entsprechend der retardierten (G_+) und avancierten (G_-) Greenschen Funktionen (IX.58).

IX.5.2 Retardierte Potentiale

Mit Hilfe der gerade eingeführten Greenschen Funktionen können nun spezielle Lösungen der Bewegungsgleichungen (IX.21) für die elektromagnetischen Potentiale geschrieben werden. Somit liefert das Einsetzen der retardierten Greenschen Funktion (IX.58a) in die allgemeine Formel (IX.57b) mit Quellterm $-\rho_{\text{el.}}(t', \vec{r}')/\epsilon_0$ bzw. $-\mu_0 \vec{j}_{\text{el.}}(t', \vec{r}')$

$$\Phi_{\text{ret.}}(t, \vec{r}) = -\frac{1}{\epsilon_0} \int G_{\text{ret.}}(t, \vec{r}; t', \vec{r}') \rho_{\text{el.}}(t', \vec{r}') dt' d^3\vec{r}' \quad (\text{IX.62a})$$

bzw.

$$\vec{A}_{\text{ret.}}(t, \vec{r}) = -\mu_0 \int G_{\text{ret.}}(t, \vec{r}; t', \vec{r}') \vec{j}_{\text{el.}}(t', \vec{r}') dt' d^3\vec{r}'. \quad (\text{IX.62b})$$

Diese Ausdrücke sehen offensichtlich sehr ähnlich aus. Daher werden detaillierte Berechnungen hier-nach nur am Beispiel des Vektorpotentials $\vec{A}_{\text{ret.}}$ durchgeführt.

Mit dem expliziten Ausdruck (IX.58a) der retardierten Greenschen Funktion gelten

$$\begin{aligned} \vec{A}_{\text{ret.}}(t, \vec{r}) &= -\mu_0 \int \frac{-1}{4\pi|\vec{r}-\vec{r}'|} \delta\left(\frac{|\vec{r}-\vec{r}'|}{c} - (t-t')\right) \vec{j}_{\text{el.}}(t', \vec{r}') dt' d^3\vec{r}' \\ &= \frac{\mu_0}{4\pi} \int \frac{1}{|\vec{r}-\vec{r}'|} \left[\int \delta\left(t-t' - \frac{|\vec{r}-\vec{r}'|}{c}\right) \vec{j}_{\text{el.}}(t', \vec{r}') dt' \right] d^3\vec{r}' \end{aligned} \quad (\text{IX.63a})$$

und für das Skalarpotential

$$\Phi_{\text{ret.}}(t, \vec{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{1}{|\vec{r}-\vec{r}'|} \left[\int \delta\left(t-t' - \frac{|\vec{r}-\vec{r}'|}{c}\right) \rho_{\text{el.}}(t', \vec{r}') dt' \right] d^3\vec{r}'. \quad (\text{IX.63b})$$

Nach Durchführen des Integrals über die Zeitvariable t' kommen

$$\Phi_{\text{ret.}}(t, \vec{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{1}{|\vec{r}-\vec{r}'|} \rho_{\text{el.}}\left(t - \frac{|\vec{r}-\vec{r}'|}{c}, \vec{r}'\right) d^3\vec{r}' \quad (\text{IX.64a})$$

$$\vec{A}_{\text{ret.}}(t, \vec{r}) = \frac{\mu_0}{4\pi} \int \frac{1}{|\vec{r}-\vec{r}'|} \vec{j}_{\text{el.}}\left(t - \frac{|\vec{r}-\vec{r}'|}{c}, \vec{r}'\right) d^3\vec{r}'. \quad (\text{IX.64b})$$

Φ_{ret} und $\vec{A}_{\text{ret.}}$ heißen *retardierte Potentiale*. Dabei hängt das Skalar bzw. Vektorpotential am Ort \vec{r} zur Zeit t von der Ladungsdichte $\rho_{\text{el.}}$ bzw. Stromdichte $\vec{j}_{\text{el.}}$ am Ort \vec{r}' zur früheren, *retardierten* Zeit $t - |\vec{r} - \vec{r}'|/c$ ab. Die Verzögerung entspricht genau der Reisezeit des Lichts von \vec{r}' bis zu \vec{r} .

Bemerkungen:

- * Die retardierten Potentiale genügen automatisch der Lorenz-Eichbedingung (IX.16).

Das Einsetzen der Gl. (IX.62) in die letztere gibt nämlich

$$\begin{aligned} \epsilon_0\mu_0 \frac{\partial\Phi_{\text{ret.}}(t, \vec{r})}{\partial t} + \vec{\nabla} \cdot \vec{A}_{\text{ret.}}(t, \vec{r}) = \\ -\mu_0 \int \left[\frac{\partial G_{\text{ret.}}(t, \vec{r}; t', \vec{r}')}{\partial t} \rho_{\text{el.}}(t', \vec{r}') + \vec{\nabla} G_{\text{ret.}}(t, \vec{r}; t', \vec{r}') \cdot \vec{j}_{\text{el.}}(t', \vec{r}') \right] dt' d^3\vec{r}'. \end{aligned}$$

Die Ableitung der retardierten Greenschen Funktion nach t bzw. nach den Komponenten von \vec{r} kann durch das Negative der Ableitung nach t' bzw. nach den Komponenten von \vec{r}' ersetzt werden:

$$\begin{aligned} \epsilon_0\mu_0 \frac{\partial\Phi_{\text{ret.}}(t, \vec{r})}{\partial t} + \vec{\nabla} \cdot \vec{A}_{\text{ret.}}(t, \vec{r}) = \\ \mu_0 \int \left[\frac{\partial G_{\text{ret.}}(t, \vec{r}; t', \vec{r}')}{\partial t'} \rho_{\text{el.}}(t', \vec{r}') + \vec{\nabla}_{\vec{r}'} G_{\text{ret.}}(t, \vec{r}; t', \vec{r}') \cdot \vec{j}_{\text{el.}}(t', \vec{r}') \right] dt' d^3\vec{r}'. \end{aligned}$$

Jetzt können partielle Integrationen über die Zeit t' (für den ersten Summanden im Integral) oder die Ortskoordinaten \vec{r}' (für den zweiten Summanden) durchgeführt werden. Unter der Annahme, dass die integrierten Terme verschwinden — entsprechend der Abwesenheit von Ladungen und Strömen im Unendlichen —, kommt

$$\epsilon_0\mu_0 \frac{\partial\Phi_{\text{ret.}}(t, \vec{r})}{\partial t} + \vec{\nabla} \cdot \vec{A}_{\text{ret.}}(t, \vec{r}) = -\mu_0 \int G_{\text{ret.}}(t, \vec{r}; t', \vec{r}') \left[\frac{\partial\rho_{\text{el.}}(t', \vec{r}')}{\partial t'} + \vec{\nabla}_{\vec{r}'} \cdot \vec{j}_{\text{el.}}(t', \vec{r}') \right] dt' d^3\vec{r}'.$$

Der Term in eckigen Klammern stellt genau die linke Seite der Kontinuitätsgleichung (IX.7) dar, d.h. er verschwindet. \square

* Wenn die Quellen $\rho_{\text{el.}}(t', \vec{r}')$, $\vec{j}_{\text{el.}}(t', \vec{r}')$ des elektromagnetischen Feldes vor einem Zeitpunkt t_0 identisch verschwinden, dann sind die Potentiale $\Phi_{\text{ret.}}(t, \vec{r})$, $\vec{A}_{\text{ret.}}(t, \vec{r})$ für $t \leq t_0$ ebenfalls Null.

* Mithilfe der avancierten Greenschen Funktion $G_{\text{adv.}}$ lassen sich auch *avancierte Potentiale* bestimmen. Physikalisch werden solche Potentiale üblicherweise im Namen vom Prinzip der *Kausalität* nicht angenommen: das Effekt — Potential am Punkt (t, \vec{r}) — könne nicht vor der Ursache — Quelle am Punkt $(t + |\vec{r} - \vec{r}'|/c, \vec{r}')$ — kommen. Diese *Wahl* der Lösung mit retardiertem Potential schließt also einen Unterschied zwischen Vergangenheit und Zukunft („elektromagnetischer Zeitpfeil“) ein, der in den Maxwell-Gleichungen nicht existiert.

Referenz [30] legt eine „Begründung“ dieser Wahl dar, die aber nicht zwangsläufig ist: beispielsweise haben Wheeler und Feynman [31] lineare Kombinationen von retardierten und avancierten Potentialen benutzt, um Probleme der klassischen Elektrodynamik von Punktladungen — insbesondere deren Selbstwechselwirkung — zu lösen.

Stationäre Quellen

Falls die Quellen — Ladungs- und Stromverteilungen — des elektromagnetischen Feldes stationär sind, lauten die retardierten Potentiale (IX.64)

$$\Phi_{\text{ret.}}(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{\rho_{\text{el.}}(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3\vec{r}', \quad (\text{IX.65a})$$

$$\vec{A}_{\text{ret.}}(\vec{r}) = \frac{\mu_0}{4\pi} \int \frac{\vec{j}_{\text{el.}}(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3\vec{r}'. \quad (\text{IX.65b})$$

Man findet also genau die Lösungen (VII.22) bzw. (VIII.10) der Poisson-Gleichungen der Elektrostatik (VII.4) bzw. der Magnetostatik (VIII.9) wieder.

IX.5.3 Multipolentwicklung

Die Ergebnisse des vorigen Paragraphen können zuerst auf den Fall von periodisch oszillierenden Quellen angewandt werden zwar

$$\rho_{\text{el.}}(t, \vec{r}) = \text{Re} \left[\rho_0(\vec{r}) e^{-i\omega t} \right], \quad \vec{j}_{\text{el.}}(t, \vec{r}) = \text{Re} \left[\vec{j}_0(\vec{r}) e^{-i\omega t} \right] \quad (\text{IX.66})$$

mit gegebenen Amplituden ρ_0 , \vec{j}_0 . Dann gilt

$$\frac{\partial \rho_{\text{el.}}(t, \vec{r})}{\partial t} + \vec{\nabla} \cdot \vec{j}_{\text{el.}}(t, \vec{r}) = \text{Re} \left(\left[-i\omega \rho_0(\vec{r}) + \vec{\nabla} \cdot \vec{j}_0(\vec{r}) \right] e^{-i\omega t} \right).$$

Laut der Kontinuitätsgleichung (IX.7) ist die linke Seite dieser Gleichung gleich Null, was nur dann möglich ist, wenn der Term in rechteckigen Klammern verschwindet, d.h.

$$-i\omega \rho_0(\vec{r}) + \vec{\nabla} \cdot \vec{j}_0(\vec{r}) = 0. \quad (\text{IX.67})$$

Mit der elektrischen Ladungsstromdichte $\vec{j}_{\text{el.}}$ der Gl. (IX.66) wird das retardierte Vektorpotential (IX.64b) zu

$$\vec{A}_{\text{ret.}}(t, \vec{r}) = \text{Re} \left[\frac{\mu_0 e^{-i\omega t}}{4\pi} \int \vec{j}_0(\vec{r}') \frac{e^{i\omega|\vec{r} - \vec{r}'|/c}}{|\vec{r} - \vec{r}'|} d^3\vec{r}' \right]. \quad (\text{IX.68})$$

Sei angenommen, dass die Ladungsstromdichte $\vec{j}_{\text{el.}}$ innerhalb eines Gebiets \mathcal{V} in der Umgebung des Nullpunkts $\vec{r}' = \vec{0}$ lokalisiert ist, und dass sie am Rand $\partial\mathcal{V}$ des Gebiets verschwindet.

Das elektromagnetische Feld wird in einem „weit entfernten“ Punkt \vec{r} untersucht, d.h. $|\vec{r}| \gg |\vec{r}'|$ gilt für jeden Punkt $\vec{r}' \in \mathcal{V}$. Dementsprechend darf man in erster Näherung $|\vec{r} - \vec{r}'|$ durch $r \equiv |\vec{r}|$ in Gl. (IX.68) ersetzen; daraus folgt

$$\vec{A}_{\text{ret.}}(t, \vec{r}) \simeq \text{Re} \left[\frac{\mu_0 e^{-i\omega(t-r/c)}}{4\pi r} \int_{\mathcal{V}} \vec{j}_0(\vec{r}') d^3\vec{r}' \right]. \quad (\text{IX.69})$$

Zur Berechnung des Terms in eckigen Klammern kann man das Integral

$$\int_{\mathcal{V}} \sum_{k=1}^3 \frac{\partial}{\partial x'^k} [x'^l j_0^k(\vec{r}')] d^3\vec{r}'$$

betrachten. Es lässt sich einerseits direkt integrieren, denn der Integrand eine Ableitung ist: da $j_0(\vec{r}') = 0$ für $\vec{r}' \in \partial\mathcal{V}$ ist das Integral gleich Null. Andererseits kann man die Produktregel verwenden, woraus sich

$$\int_{\mathcal{V}} \sum_{k=1}^3 \frac{\partial}{\partial x'^k} [x'^l j_0^k(\vec{r}')] d^3\vec{r}' = \int_{\mathcal{V}} \sum_{k=1}^3 \left[\delta^{kl} j_0^k(\vec{r}') + x'^l \frac{\partial j_0^k(\vec{r}')}{\partial x'^k} \right] d^3\vec{r}' = \int_{\mathcal{V}} [j_0^l(\vec{r}') + x'^l \vec{\nabla}_{\vec{r}'} \cdot \vec{j}_0(\vec{r}')] d^3\vec{r}'$$

ergibt, mit dem Gradient $\vec{\nabla}_{\vec{r}'}$ bezüglich der Komponenten von \vec{r}' . Daher gilt

$$\int_{\mathcal{V}} \vec{j}_0(\vec{r}') d^3\vec{r}' = - \int_{\mathcal{V}} \vec{r}' [\vec{\nabla}_{\vec{r}'} \cdot \vec{j}_0(\vec{r}')] d^3\vec{r}'.$$

Aus der Beziehung (IX.67) folgt dann

$$\int_{\mathcal{V}} \vec{j}_0(\vec{r}') d^3\vec{r}' = -i\omega \int_{\mathcal{V}} \vec{r}' \rho_0(\vec{r}') d^3\vec{r}'.$$

Somit wird das retardierte Potential (IX.69) zu

$$\vec{A}_{\text{ret.}}(t, \vec{r}) \simeq \text{Re} \left[-i\omega \frac{\mu_0 e^{-i\omega(t-r/c)}}{4\pi r} \int_{\mathcal{V}} \vec{r}' \rho_0(\vec{r}') d^3\vec{r}' \right].$$

Dabei ist das Integral genau das in Gl. (VII.27c) definierte elektrische Dipolmoment \vec{P}_0 der Ladungsverteilung, die das Potential verursacht. Insgesamt lautet das retardierte Vektorpotential

$$\vec{A}_{\text{ret.}}(t, \vec{r}) \simeq -\frac{\mu_0}{4\pi r} \omega \vec{P}_0 \sin \omega \left(t - \frac{r}{c} \right). \quad (\text{IX.70})$$

Dabei handelt es sich um *Dipolstrahlung*

Aus diesem Vektorpotential leitet man für $r \gg c/\omega$ die elektromagnetischen Felder

$$\vec{B}(t, \vec{r}) \simeq \frac{\mu_0}{4\pi cr} \omega^2 \vec{e}_r \times \vec{P}_0 \cos \omega \left(t - \frac{r}{c} \right) \quad (\text{IX.71a})$$

und

$$\vec{E}(t, \vec{r}) \simeq \frac{\mu_0}{4\pi r} \omega^2 \vec{e}_r \times (\vec{e}_r \times \vec{P}_0) \cos \omega \left(t - \frac{r}{c} \right) \quad (\text{IX.71b})$$

ab, wobei \vec{e}_r den Einheitsvektor entlang \vec{r} bezeichnet.

IX.5.4 Potentiale und Felder einer bewegten Punktladung

Dieser Paragraph beginnt mit der Berechnung der Potentiale und Felder, die durch eine bewegte Punktladung mit der Bahnkurve $\vec{x}(t)$ erzeugt werden, entsprechend einer Ladungsdichte bzw. Stromdichte [vgl. Gl. (VI.6)]

$$\rho_{\text{el.}}(t, \vec{r}) = q \delta^{(3)}(\vec{r} - \vec{x}(t)) \quad \text{bzw.} \quad \vec{j}_{\text{el.}}(t, \vec{r}) = q \vec{v}(t) \delta^{(3)}(\vec{r} - \vec{x}(t)).$$

Danach wird die Leistung, die durch die abgestrahlten Felder transportiert wird, bestimmt.

IX.5.4a Liénard–Wiechert-Potentiale

Mit dieser Form für die Quellterme lauten die retardierten Potentiale (IX.63)

$$\begin{aligned} \vec{A}_{\text{ret.}}(t, \vec{r}) &= -\frac{\mu_0 q}{4\pi} \int \frac{1}{|\vec{r} - \vec{r}'|} \delta \left(t - t' - \frac{|\vec{r} - \vec{r}'|}{c} \right) \vec{v}(t') \delta^{(3)}(\vec{r}' - \vec{x}(t')) dt' d^3\vec{r}' \\ &= \frac{\mu_0 q}{4\pi} \int_{-\infty}^{\infty} \vec{v}(t') \frac{1}{|\vec{r} - \vec{x}(t')|} \delta \left(t - t' - \frac{|\vec{r} - \vec{x}(t')|}{c} \right) dt', \end{aligned} \quad (\text{IX.72a})$$

wobei das Integral über \vec{r}' durchgeführt wurde, und ähnlich

$$\Phi_{\text{ret.}}(t, \vec{r}) = \frac{q}{4\pi\epsilon_0} \int_{-\infty}^{\infty} \frac{1}{|\vec{r} - \vec{x}(t')|} \delta\left(t - t' - \frac{|\vec{r} - \vec{x}(t')|}{c}\right) dt' \quad (\text{IX.72b})$$

Wie im § IX.5.2 werden die Herleitungen hiernach nur im Fall des Vektorpotentials detailliert.

Das Argument $f(t') \equiv t - t' - |\vec{r} - \vec{x}(t')|/c$ der δ -Distribution im Integranden verschwindet für einen einzigen Wert von t' , der als $t_{\text{ret.}}$ bezeichnet wird und *retardierte Zeit* heißt. $t_{\text{ret.}}$ ist die Zeit, zu der die Raumzeitlinie — d.h. die Trajektorie in der Raumzeit — $(t', \vec{x}(t'))$ der Punktladung den Rückwärtslichtkegel des Punkts (t, \vec{r}) durchschneidet (Abb. IX.3).

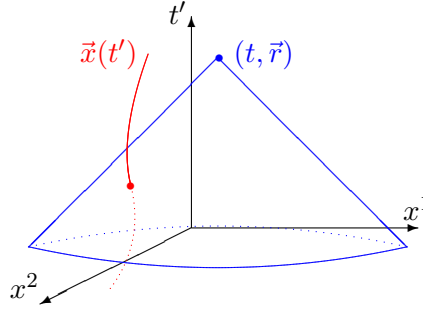


Abbildung IX.3 – Retardierte Zeit.

Somit genügt die retardierte Zeit der impliziten Gleichung

$$t_{\text{ret.}} = t - \frac{|\vec{r} - \vec{x}(t_{\text{ret.}})|}{c}. \quad (\text{IX.73})$$

Die Integration nach t' in Gl. (IX.72a) folgt aus der Formel $\delta(f(t')) = \delta(t' - t_{\text{ret.}})/|f'(t_{\text{ret.}})|$, wobei hier

$$f'(t') = -1 + \frac{1}{c} \frac{\vec{r} - \vec{x}(t')}{|\vec{r} - \vec{x}(t')|} \cdot \vec{v}(t').$$

Daher gilt

$$f'(t_{\text{ret.}}) = -1 + \vec{e}_{\text{ret.}} \cdot \vec{\beta}_{\text{ret.}},$$

mit dem Einheitsvektor

$$\vec{e}_{\text{ret.}} \equiv \frac{\vec{r} - \vec{x}(t_{\text{ret.}})}{|\vec{r} - \vec{x}(t_{\text{ret.}})|} \equiv \frac{\vec{X}_{\text{ret.}}}{|\vec{X}_{\text{ret.}}|} \quad (\text{IX.74a})$$

in Richtung von $\vec{X}_{\text{ret.}} \equiv \vec{r} - \vec{x}(t_{\text{ret.}})$, d.h. von der Quelle des Potentials bis zu dessen Beobachtungspunkt, während

$$\vec{\beta}_{\text{ret.}} = \frac{\vec{v}_{\text{ret.}}}{c} \equiv \frac{\vec{v}(t_{\text{ret.}})}{c} \quad (\text{IX.74b})$$

die Geschwindigkeit der Punktladung zur retardierte Zeit bezeichnet. Da der Betrag von $\vec{\beta}_{\text{ret.}}$ streng kleiner als 1 ist, bleibt $f'(t_{\text{ret.}})$ immer negativ.

Somit ergeben sich die *Liénard^(ap)–Wiechert^(aq)-Potentiale*

$$\vec{A}(t, \vec{r}) = \frac{\mu_0}{4\pi} \frac{q \vec{v}_{\text{ret.}}}{|\vec{r} - \vec{x}(t_{\text{ret.}})| - [\vec{r} - \vec{x}(t_{\text{ret.}})] \cdot \frac{\vec{v}_{\text{ret.}}}{c}}, \quad (\text{IX.74c})$$

und

$$\Phi(t, \vec{r}) = \frac{1}{4\pi\epsilon_0} \frac{q}{|\vec{r} - \vec{x}(t_{\text{ret.}})| - [\vec{r} - \vec{x}(t_{\text{ret.}})] \cdot \frac{\vec{v}_{\text{ret.}}}{c}}. \quad (\text{IX.74d})$$

^(ap)A.-M. LIÉNARD, 1869–1958 ^(aq)E. WIECHERT, 1861–1928

Dabei hängt die Geschwindigkeit $\vec{v}_{\text{ret.}}$ der Punktladung zur retardierten Zeit von $t_{\text{ret.}}$ und somit implizit von t und \vec{r} ab.

Bemerkung: Der Nenner dieser Potentiale ist immer positiv.

IX.5.4 b Elektrisches und magnetisches Feld

Das elektrische und das magnetische Feld können aus den Gl. (IX.12) und (IX.13) hergeleitet werden. Die Liénard–Wiechert-Potentiale (IX.74) hängen aber nicht nur explizit von der Raumzeitvariablen t und \vec{r} ab, sondern auch implizit über die retardierte Zeit (IX.73). Dementsprechend ist die Berechnung der Felder etwa mühsam. Letztendlich findet man die retardierten Feldern:

$$\vec{E}(t, \vec{r}) = \frac{q}{4\pi\epsilon_0(1 - \vec{e}_{\text{ret.}} \cdot \vec{\beta}_{\text{ret.}})^3} \left(\frac{\vec{e}_{\text{ret.}} - \vec{\beta}_{\text{ret.}}}{\gamma_{\text{ret.}}^2 |\vec{X}_{\text{ret.}}|^2} + \frac{\vec{e}_{\text{ret.}} \times [(\vec{e}_{\text{ret.}} - \vec{\beta}_{\text{ret.}}) \times \vec{a}_{\text{ret.}}]}{c^2 |\vec{X}_{\text{ret.}}|} \right), \quad (\text{IX.75a})$$

$$\vec{B}(t, \vec{r}) = \frac{1}{c} \vec{e}_{\text{ret.}} \times \vec{E}(t, \vec{r}), \quad (\text{IX.75b})$$

mit $\vec{\beta}_{\text{ret.}}$ bzw. $\vec{a}_{\text{ret.}}$ der auf c normierten Geschwindigkeit bzw. der Beschleunigung der Punktladung zur retardierten Zeit und $\gamma_{\text{ret.}} \equiv (1 - \vec{\beta}_{\text{ret.}}^2)^{-1/2}$ dem entsprechenden Lorentz-Faktor.

Beweis: Um die retardierten Felder herzuleiten, lohnt es sich, die Potentiale (IX.74c)–(IX.74d) umzuschreiben [vgl. Gl. (IX.72)]

$$\Phi(t, \vec{r}) = \frac{q}{4\pi\epsilon_0} \int_{-\infty}^{\infty} \frac{1}{|\vec{r} - \vec{x}(t')|} \delta\left(t - t' - \frac{|\vec{r} - \vec{x}(t')|}{c}\right) dt' \quad (\text{IX.76a})$$

$$\vec{A}(t, \vec{r}) = \frac{\mu_0 q}{4\pi} \int_{-\infty}^{\infty} \frac{\vec{v}(t')}{|\vec{r} - \vec{x}(t')|} \delta\left(t - t' - \frac{|\vec{r} - \vec{x}(t')|}{c}\right) dt'. \quad (\text{IX.76b})$$

Sei jetzt $\vec{X}(t') \equiv \vec{r} - \vec{x}(t')$ mit $\vec{e}(t') \equiv \vec{X}(t')/|\vec{X}(t')|$ dem entsprechenden Einheitsvektor. Aus diesen Gleichungen folgen

$$\vec{\nabla} \Phi(t, \vec{r}) = \frac{q}{4\pi\epsilon_0} \int_{-\infty}^{\infty} \left[-\frac{1}{|\vec{X}(t')|^2} \delta\left(t - t' - \frac{|\vec{X}(t')|}{c}\right) - \frac{1}{c|\vec{X}(t')|} \delta'\left(t - t' - \frac{|\vec{X}(t')|}{c}\right) \right] \vec{e}(t') dt',$$

$$\frac{\partial \vec{A}(t, \vec{r})}{\partial t} = \frac{\mu_0 q}{4\pi} \int_{-\infty}^{\infty} \frac{\vec{v}(t')}{|\vec{X}(t')|} \delta'\left(t - t' - \frac{|\vec{X}(t')|}{c}\right) dt'.$$

Die Substitution $s = f(t') = t - t' - |\vec{X}(t')|/c$ gibt $ds/dt' = -1 + \vec{e} \cdot \vec{\beta}$ und dadurch

$$\vec{E}(t, \vec{r}) = \frac{q}{4\pi\epsilon_0} \int_{-\infty}^{\infty} \frac{1}{1 - \vec{e} \cdot \vec{\beta}} \left[\frac{\vec{e}}{|\vec{X}|^2} \delta(s) + \frac{\vec{e} - \vec{\beta}}{c|\vec{X}|} \delta'(s) \right] ds, \quad (\text{IX.77})$$

wobei \vec{e} , \vec{X} , \vec{v} und $\vec{\beta} \equiv \vec{v}/c$ zur Zeit $t' = f^{-1}(s)$ zu betrachten sind. Der Term mit der Ableitung der δ -Distribution kann mithilfe partieller Integration berechnet werden:

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{\vec{e} - \vec{\beta}}{c(1 - \vec{e} \cdot \vec{\beta})|\vec{X}|} \delta'(s) ds &= - \int_{-\infty}^{\infty} \delta(s) \frac{d}{ds} \left[\frac{\vec{e} - \vec{\beta}}{c(1 - \vec{e} \cdot \vec{\beta})|\vec{X}|} \right] ds \\ &= \int_{-\infty}^{\infty} \delta(s) \frac{1}{1 - \vec{e} \cdot \vec{\beta}} \frac{d}{dt'} \left[\frac{\vec{e} - \vec{\beta}}{c(1 - \vec{e} \cdot \vec{\beta})|\vec{X}|} \right] ds. \end{aligned}$$

Die Ableitung nach t' im Integranden folgt aus den Ableitungen

$$\frac{d}{dt'} \frac{1}{|\vec{X}|} = \frac{\vec{e} \cdot \vec{v}}{|\vec{X}|^2}, \quad \frac{d\vec{e}}{dt'} = \frac{(\vec{e} \cdot \vec{v})\vec{e} - \vec{v}}{|\vec{X}|}, \quad \frac{d}{dt'} \frac{1}{1 - \vec{e} \cdot \vec{\beta}} = \frac{1}{(1 - \vec{e} \cdot \vec{\beta})^2} \left[\frac{\vec{e} \cdot \vec{a}}{c} + \frac{(\vec{e} \cdot \vec{v})^2 - v^2}{c|\vec{X}|} \right],$$

mit der Beschleunigung $\vec{a} = d\vec{v}/dt'$ der Punktladung. Dies führt zu

$$\frac{d}{dt'} \left[\frac{\vec{e} - \vec{\beta}}{c(1 - \vec{e} \cdot \vec{\beta})|\vec{X}|} \right] = \frac{1}{|\vec{X}|(1 - \vec{e} \cdot \vec{\beta})^2} \left[\frac{(\vec{e} \cdot \vec{a})\vec{e} - \vec{a} - (\vec{e} \cdot \vec{a})\vec{\beta} + (\vec{e} \cdot \vec{\beta})\vec{a}}{c^2} + \frac{2(\vec{e} \cdot \vec{\beta}) - (\vec{e} \cdot \vec{\beta})^2 - \beta^2}{|\vec{X}|} \vec{e} - \frac{1 - \beta^2}{|\vec{X}|} \vec{\beta} \right],$$

so dass Gl. (IX.77) insgesamt lautet

$$\vec{E}(t, \vec{r}) = \frac{q}{4\pi\epsilon_0} \int_{-\infty}^{\infty} \frac{\delta(s)}{|\vec{X}|(1 - \vec{e} \cdot \vec{\beta})^3} \left(\frac{(\vec{e} \cdot \vec{a})(\vec{e} - \vec{\beta}) - [\vec{e} \cdot (\vec{e} - \vec{\beta})]\vec{a}}{c^2} + \frac{(1 - \beta^2)(\vec{e} - \vec{\beta})}{|\vec{X}|} \right) ds.$$

Dann ist der Integrand gleich dem Produkt eines Terms $\delta(s)$ mit einer Funktion von s : das Integral ist einfach der Wert der Letzteren für $s = 0$, d.h. $t' = t_{\text{ret.}}$. Somit findet man Gl. (IX.75a).

Das magnetische Feld $\vec{B}(t, \vec{r})$ lässt sich ähnlich berechnen, ausgehend aus

$$\vec{\nabla} \times \vec{A}(t, \vec{r}) = \frac{\mu_0 q}{4\pi} \int_{-\infty}^{\infty} \vec{e}(t') \times \vec{v}(t') \left[\frac{-1}{|\vec{X}(t')|^2} \delta\left(t - t' - \frac{|\vec{X}(t')|}{c}\right) - \frac{1}{c|\vec{X}(t')|} \delta'\left(t - t' - \frac{|\vec{X}(t')|}{c}\right) \right] dt'.$$

Die retardierten elektromagnetischen Felder (IX.75) bestehen aus einem Beitrag proportional zu $1/|\vec{X}_{\text{ret.}}|^2$ und einem Term proportional zu $1/|\vec{X}_{\text{ret.}}|$. In einem Bezugssystem, das sich mit der Punktladung zur retardierten Zeit mitbewegt, d.h. wo $\vec{\beta}_{\text{ret.}} = \vec{0}$, lauten sie

$$\vec{E} = \frac{q}{4\pi\epsilon_0} \left[\frac{\vec{e}_{\text{ret.}}}{|\vec{X}_{\text{ret.}}|^2} + \frac{\vec{e}_{\text{ret.}} \times (\vec{e}_{\text{ret.}} \times \vec{a}_{\text{ret.}})}{c^2 |\vec{X}_{\text{ret.}}|} \right], \quad \vec{B} = \frac{1}{c} \vec{e}_{\text{ret.}} \times \vec{E}.$$

Das elektrische Feld ist die Summe aus dem Coulomb-Feld in $1/|\vec{X}_{\text{ret.}}|^2$ einer ruhenden Punktladung

$$\vec{E}_{\text{Coul.}} = \frac{q}{4\pi\epsilon_0} \frac{\vec{e}_{\text{ret.}}}{|\vec{X}_{\text{ret.}}|^2}, \quad \vec{B}_{\text{Coul.}} = \vec{0},$$

und, wenn die retardierte Beschleunigung $\vec{a}_{\text{ret.}}$ der Punktladung nicht Null ist, aus dem *Strahlungsfeld* in $1/|\vec{X}_{\text{ret.}}|$

$$\vec{E}_{\text{Strahl.}} = \frac{q}{4\pi\epsilon_0 c^2 |\vec{X}_{\text{ret.}}|} \vec{e}_{\text{ret.}} \times (\vec{e}_{\text{ret.}} \times \vec{a}_{\text{ret.}}), \quad \vec{B}_{\text{Strahl.}} = \frac{1}{c} \vec{e}_{\text{ret.}} \times \vec{E}_{\text{Strahl.}}$$

Diese Strahlungsfelder sind senkrecht zu $\vec{e}_{\text{ret.}}$.

Ähnlicherweise entspricht der Beitrag in $1/|\vec{X}_{\text{ret.}}|^2$ in Gl. (IX.75) dem Coulomb-Feld, und der Term in $1/|\vec{X}_{\text{ret.}}|$, dem Strahlungsfeld.

Bemerkung: Wenn die Punktladung beschleunigt wird, muss sie einer Kraft unterliegen, d.h. sie muss in einem elektromagnetischen Feld sein: dieses „äußere“ Feld wird hier nicht präzisiert. Dagegen lässt sich die Kraft $\vec{F}(t)$, die für die Beschleunigung zur Zeit t verantwortlich ist, dank dem Grundgesetz der Mechanik (XI.14) einfach durch die kinetischen Größen der Punktladung ausdrücken:

$$\vec{F}(t) = \frac{d}{dt} [m\gamma(t)\vec{v}(t)] = \gamma(t)m \left(\vec{a}(t) + \gamma^2(t) [\vec{\beta}(t) \cdot \vec{a}(t)] \vec{\beta}(t) \right) \quad (\text{IX.78})$$

mit m der Masse der Punktladung.

Beispiel: Punktladung in gleichförmiger geradliniger Bewegung

Wenn $\vec{a}_{\text{ret.}} = \vec{0}$ zu jeder retardierten Zeit vereinfacht sich das elektrische Feld (IX.75) zu

$$\vec{E}(t, \vec{r}) = \frac{q}{4\pi\epsilon_0 (1 - \vec{e}_{\text{ret.}} \cdot \vec{\beta})^3} \frac{\vec{e}_{\text{ret.}} - \vec{\beta}}{\gamma^2 |\vec{X}_{\text{ret.}}|^2} = \frac{q}{4\pi\epsilon_0} \frac{\vec{X}_{\text{ret.}} - |\vec{X}_{\text{ret.}}| \vec{\beta}}{\gamma^2 (|\vec{X}_{\text{ret.}}| - \vec{X}_{\text{ret.}} \cdot \vec{\beta})^3}, \quad (\text{IX.79})$$

mit konstanter Geschwindigkeit $\vec{\beta}$.

Aus $|\vec{X}_{\text{ret.}}| = |\vec{r} - \vec{x}(t_{\text{ret.}})| = c(t - t_{\text{ret.}})$ folgt für die gleichförmige Bewegung der Punktladung

$$\vec{x}(t) = \vec{x}(t_{\text{ret.}}) + c\vec{\beta}(t - t_{\text{ret.}}) = \vec{x}(t_{\text{ret.}}) + |\vec{X}_{\text{ret.}}|\vec{\beta},$$

d.h. für den Vektor im Zähler der Gl. (IX.79) $\vec{X}_{\text{ret.}} - |\vec{X}_{\text{ret.}}|\vec{\beta} = \vec{r} - \vec{x}(t)$: das elektrische Feld zur Zeit t ist in der *momentanen* Richtung von der Ladung nach \vec{r} , nicht in der „retardierten“ Richtung gerichtet.

IX.5.4c Abgestrahlte Leistung

Sei d^2S ein Flächenelement in einem Punkt \vec{r} im Abstand X einer Punktladung, von der aus d^2S unter einem Raumwinkelelement $d^2\Omega$ gesehen wird. Der Einheitsvektor in Richtung von der Punktladung nach \vec{r} wird mit $\vec{e}_{\text{ret.}}$ bezeichnet.

Ein Beobachter \mathcal{B} in \vec{r} misst die elektromagnetischen Felder (IX.75): die Energie, die pro Einheit der Eigenzeit t von \mathcal{B} durch d^2S strömt, ist die von \mathcal{B} empfangene Leistung

$$d^2\mathcal{P} = \vec{S}_{\text{e.m.}} \cdot \vec{e}_{\text{ret.}} X^2 d^2\Omega,$$

mit $\vec{S}_{\text{e.m.}} = \vec{E} \times \vec{B}/\mu_0$ dem Poynting-Vektor in \vec{r} [Gl. (IX.23b)]. Für große Abstände $X \rightarrow \infty$ trägt der Coulomb-Anteil der retardierten Felder nicht bei, und man darf das Strahlungsfeld alleine betrachten. Dann gilt

$$\frac{d^2\mathcal{P}}{d^2\Omega} = \frac{1}{\mu_0 c} X^2 \left[\vec{E}_{\text{Strahl.}} \times (\vec{e}_{\text{ret.}} \times \vec{E}_{\text{Strahl.}}) \right] \cdot \vec{e}_{\text{ret.}}.$$

Das doppelte Kreuzprodukt lautet $\vec{E}_{\text{Strahl.}}^2 \vec{e}_{\text{ret.}} - (\vec{e}_{\text{ret.}} \cdot \vec{E}_{\text{Strahl.}}) \vec{E}_{\text{Strahl.}}$, wobei der zweite Term wegen der Orthogonalität von $\vec{E}_{\text{Strahl.}}$ und $\vec{e}_{\text{ret.}}$ Null ist. Damit ergibt sich unter Berücksichtigung des Ausdrucks des Strahlungsfeldes in Gl. (IX.75a)

$$\frac{d^2\mathcal{P}}{d^2\Omega} = \frac{\mu_0 q^2}{(4\pi)^2 c (1 - \vec{e}_{\text{ret.}} \cdot \vec{\beta}_{\text{ret.}})^6} \left| \vec{e}_{\text{ret.}} \times [(\vec{e}_{\text{ret.}} - \vec{\beta}_{\text{ret.}}) \times \vec{a}_{\text{ret.}}] \right|^2,$$

wobei ein Faktor $1/(\mu_0 \epsilon_0^2 c^5)$ in μ_0/c transformiert wurde.

Wegen Energieerhaltung ist die elektromagnetische Energie, welche die Punktladung in Richtung von d^2S abstrahlt, gleich der Energie, die (später) im Beobachtungspunkt \vec{r} durch d^2S strömt. Die Leistung, die durch die Punktladung abgestrahlt wird, ist genau diese Energie, bezogen auf die Einheit der Eigenzeit der Punktladung. Diese Eigenzeit ist gerade die retardierte Zeit $t_{\text{ret.}}$, so dass die pro Raumwinkelelement abgestrahlte Leistung durch

$$\frac{d^2\mathcal{P}_0}{d^2\Omega} = \frac{d^2\mathcal{P}}{d^2\Omega} \frac{\partial t}{\partial t_{\text{ret.}}} = \frac{\mu_0 q^2}{(4\pi)^2 c (1 - \vec{e}_{\text{ret.}} \cdot \vec{\beta}_{\text{ret.}})^5} \left| \vec{e}_{\text{ret.}} \times [(\vec{e}_{\text{ret.}} - \vec{\beta}_{\text{ret.}}) \times \vec{a}_{\text{ret.}}] \right|^2 \quad (\text{IX.80})$$

gegeben wird, wobei $\partial t/\partial t_{\text{ret.}} = 1 - \vec{e}_{\text{ret.}} \cdot \vec{\beta}_{\text{ret.}}$ benutzt wurde.

Um die letztere partielle Ableitung zu erhalten, kann man Gl. (IX.73) differenzieren, was zu

$$dt_{\text{ret.}} - dt + \frac{1}{c} \frac{\vec{X}_{\text{ret.}} \cdot d\vec{X}_{\text{ret.}}}{|\vec{X}_{\text{ret.}}|} = dt_{\text{ret.}} - dt + \frac{1}{c} \vec{e}_{\text{ret.}} \cdot (d\vec{r} - \vec{v}_{\text{ret.}} dt_{\text{ret.}}) = 0$$

führt, d.h. $(1 - \vec{e}_{\text{ret.}} \cdot \vec{\beta}_{\text{ret.}}) dt_{\text{ret.}} = dt - \vec{e}_{\text{ret.}} \cdot d\vec{r}/c$. Dann gelten

$$\frac{\partial t_{\text{ret.}}}{\partial t} = \frac{1}{1 - \vec{e}_{\text{ret.}} \cdot \vec{\beta}_{\text{ret.}}}, \quad \vec{\nabla} t_{\text{ret.}} = -\frac{\vec{e}_{\text{ret.}}}{c(1 - \vec{e}_{\text{ret.}} \cdot \vec{\beta}_{\text{ret.}})}.$$

Bemerkungen:

- * Eine nicht-beschleunigte Punktladung strahlt keine Energie ab!
- * Aus Sicht der Punktladung entspricht die als elektromagnetisches Feld abgestrahlte Energie bzw. Leistung einem *Strahlungsverlust*.

Dieser Verlust liegt einem wichtigen Problem der klassischen Physik zugrunde, und zwar der Frage der Stabilität von Atomen: die an den Atomkern gebundenen Elektronen unterliegen einer anziehenden Kraft, und werden daher beschleunigt. Wenn sie teil ihrer Energie als Strahlung verlieren, sollte ihre kinetische Energie allmählich kleiner werden, d.h. sie sollten irgendwann auf den Atomkern fallen. Dies ist aber nicht der Fall, was im Rahmen der klassischen Physik nicht erklärbar ist, und stellt damit eine wichtige Begründung für die Quantenmechanik — oder zumindest für etwas jenseits der klassischen Physik — dar.

Nicht-relativistischer Grenzfall

Im Fall $|\vec{\beta}_{\text{ret.}}| \ll |\vec{e}_{\text{ret.}}| = 1$ vereinfacht sich Gl. (IX.80) zu

$$\frac{d^2\mathcal{P}_0}{d^2\Omega} = \frac{\mu_0 q^2}{(4\pi)^2 c} |\vec{e}_{\text{ret.}} \times (\vec{e}_{\text{ret.}} \times \vec{a}_{\text{ret.}})|^2.$$

Sei θ der Winkel zwischen der retardierten Beschleunigung $\vec{a}_{\text{ret.}}$ und der Beobachtungsrichtung $\vec{e}_{\text{ret.}}$; dann gilt

$$\frac{d^2\mathcal{P}_0}{d^2\Omega} = \frac{\mu_0 q^2}{(4\pi)^2 c} |\vec{a}_{\text{ret.}}|^2 \sin^2 \theta = \frac{q^2 |\vec{a}_{\text{ret.}}|^2 \sin^2 \theta}{4\pi \epsilon_0 c^3 \cdot 4\pi}.$$

Die Strahlungsleistung ist maximal für $\theta = \frac{\pi}{2}$, d.h. senkrecht zur Richtung der Beschleunigung.

Die gesamte abgestrahlte Leistung ergibt sich durch Integration dieser Formel über den Raumwinkel: mit $d^2\Omega = \sin \theta d\theta d\varphi$ erhält man die *Larmor^(ar)-Formel*

$$\mathcal{P}_0 = \int \frac{d^2\mathcal{P}_0}{d^2\Omega} d^2\Omega = \frac{2}{3} \frac{q^2 |\vec{a}_{\text{ret.}}|^2}{4\pi \epsilon_0 c^3}. \quad (\text{IX.81})$$

Allgemeiner Fall

Die Larmor-Formel (IX.81) gilt in einem mitbewegten Inertialsystem \mathcal{B}_0 , relativ zu welchem die retardierte Geschwindigkeit der Punktladung momentan verschwindet. In einem Bezugssystem $\mathcal{B}_{\vec{v}}$, wo die retardierte Geschwindigkeit der Punktladung den Wert \vec{v} annimmt, gilt

$$\mathcal{P}_0 = \frac{2}{3} \frac{q^2}{4\pi \epsilon_0 c^3} a_{\text{ret.}}^\mu a_{\text{ret.}\mu}, \quad (\text{IX.82})$$

mit $a^\mu \equiv du^\mu/d\tau \equiv d^2x^\mu/d\tau^2$ den kontravarianten Komponenten der Viererbeschleunigung (XI.15) der Punktladung, die hier zur retardierten Zeit betrachtet werden soll.

Diese Formel lässt sich noch umschreiben als

$$\mathcal{P}_0 = \frac{2}{3} \frac{q^2}{4\pi \epsilon_0 c^3} \gamma_{\text{ret.}}^6 \left[(\vec{a}_{\text{ret.}})^2 - (\vec{\beta}_{\text{ret.}} \times \vec{a}_{\text{ret.}})^2 \right]. \quad (\text{IX.83})$$

Beweis der relativistischen Formeln (IX.82)–(IX.83):

Der erste Schritt besteht in der Beobachtung, dass die abgestrahlte Leistung Lorentz-invariant ist. Es seien $d\mathcal{E}_0$ bzw. $d\mathcal{E}$ das Differential der Energie im Bezugssystem \mathcal{B}_0 bzw. $\mathcal{B}_{\vec{v}}$, und dt_0 bzw. dt die entsprechenden Differentiale der Zeit. Dann gelten $d\mathcal{E} = \gamma d\mathcal{E}_0$ und $dt = \gamma dt_0$, mit $\gamma = 1/\sqrt{1 - \vec{v}^2/c^2}$ dem Lorentz-Faktor. Daraus folgt

$$\frac{d\mathcal{E}}{dt} = \frac{d\mathcal{E}_0}{dt_0} = \mathcal{P}_0.$$

Berechnet man dann das Lorentz-Skalar $a^\mu a_\mu$ der Viererbeschleunigung, so kommt erstens

$$a^\mu = \frac{du^\mu}{d\tau} = \gamma \frac{du^\mu}{dt} = (\gamma^4 \vec{\beta} \cdot \vec{a}, \gamma^2 \vec{a} + \gamma^4 (\vec{\beta} \cdot \vec{a}) \vec{\beta})$$

mit der Beschleunigung $\vec{a} = d\vec{v}/dt$ der Punktladung, und daher

$$a^\mu a_\mu = \gamma^4 [\vec{a}^2 + \gamma^2 (\vec{\beta} \cdot \vec{a})^2]. \quad (\text{IX.84})$$

^(ar)J. LARMOR, 1857–1942

Im Bezugssystem \mathcal{B}_0 , wo $\vec{\beta} = \vec{0}$ bzw. $\gamma = 1$, stimmen die Formeln (IX.81) und (IX.82) miteinander überein. Da der Ausdruck auf der rechten Seite von Gl. (IX.82) ein Lorentz-Skalar ist, erhält er den gleichen Wert in allen Inertialsystemen, d.h. die Formel gilt in allen solchen Bezugssystemen.

Sei ϑ der Winkel zwischen $\vec{\beta}$ und \vec{a} . Aus

$$\vec{a}^2 - (\vec{\beta} \times \vec{a})^2 = \vec{a}^2 - \vec{\beta}^2 \vec{a}^2 (1 - \cos^2 \vartheta) = (1 - \vec{\beta}^2) \vec{a}^2 + (\vec{\beta} \cdot \vec{a})^2 = \frac{1}{\gamma^2} \vec{a}^2 + (\vec{\beta} \cdot \vec{a})^2$$

und dem Viererquadrat (IX.84) folgt dann die Identität der rechten Gliedern der Gl. (IX.82) und (IX.83). \square

Beschleunigung parallel zur Geschwindigkeit

Wenn Beschleunigung und Geschwindigkeit parallel zueinander sind, gelten $\vec{a}_{\text{ret.}} \times \vec{\beta}_{\text{ret.}} = \vec{0}$ und $\vec{a}_{\text{ret.}} \cdot \vec{\beta}_{\text{ret.}} = |\vec{\beta}_{\text{ret.}}| \cos \theta$. Dann lautet Gl. (IX.80)

$$\frac{d^2 \mathcal{P}_0}{d^2 \Omega} = \frac{\mu_0 q^2 |\vec{a}_{\text{ret.}}|^2}{(4\pi)^2 c} \frac{\sin^2 \theta}{(1 - |\vec{\beta}_{\text{ret.}}| \cos \theta)^5}. \quad (\text{IX.85})$$

Für $|\vec{\beta}_{\text{ret.}}| \rightarrow 1$ wird also $d^2 \mathcal{P}_0 / d^2 \Omega$ groß in der Vorwärtsrichtung $\cos \theta = 1$.

Genauer kann man prüfen, dass das Maximum von $d^2 \mathcal{P}_0 / d^2 \Omega$ im Fall $1 - |\vec{\beta}_{\text{ret.}}| \ll 1$ für $\theta \simeq 1/2\gamma$ erreicht wird.

Die Integration der Leistung (IX.85) über den ganzen Raumwinkel führt zu

$$\mathcal{P}_0 = \frac{2}{3} \frac{\mu_0 q^2 |\vec{a}_{\text{ret.}}|^2}{4\pi c} \gamma^6 = \frac{2}{3} \frac{q^2 |\vec{a}_{\text{ret.}}|^2}{4\pi \epsilon_0 c^3} \gamma^6, \quad (\text{IX.86})$$

mit dem Lorentz-Faktor γ assoziiert mit der Geschwindigkeit der Punktladung. Dieses Resultat folgt auch direkt aus der relativistischen Formel (IX.83).

Die Leistung (IX.86) wird also unendlich groß, wenn die Geschwindigkeit der Punktladung gegen c strebt. Dieses Resultat zeigt, dass es nicht möglich ist, eine Punktladung bis zur Lichtgeschwindigkeit zu beschleunigen: dies würde eine unendliche Energie erfordern, um die abgestrahlte Leistung zu kompensieren.

Wenn Beschleunigung und Geschwindigkeit parallel zueinander sind, nimmt die Kraft (IX.78), die für die Beschleunigung der Punktladung verantwortlich ist, die Form

$$\vec{F}(t) = \gamma(t)^3 m \vec{a}(t)$$

an. Somit lässt sich die Leistung (IX.86) als

$$\mathcal{P}_0 = \frac{2}{3} \frac{\mu_0 q^2}{4\pi m^2 c} |\vec{F}(t_{\text{ret.}})|^2 = \frac{2}{3} \frac{q^2}{4\pi \epsilon_0 m^2 c^3} |\vec{F}(t_{\text{ret.}})|^2$$

schreiben.

Beschleunigung senkrecht zur Richtung der Geschwindigkeit

Wenn Beschleunigung und Geschwindigkeit orthogonal sind, entsprechend einer Kreisbewegung, lautet die abgestrahlte Leistung (IX.83)

$$\mathcal{P}_0 = \frac{2}{3} \frac{q^2}{4\pi \epsilon_0 c^3} \gamma_{\text{ret.}}^6 (1 - \vec{\beta}_{\text{ret.}}^2) |\vec{a}_{\text{ret.}}|^2 = \frac{2}{3} \frac{q^2}{4\pi \epsilon_0 c^3} \gamma_{\text{ret.}}^4 |\vec{a}_{\text{ret.}}|^2. \quad (\text{IX.87})$$

Dazu lautet die Kraft (IX.78) auf die Punktladung $\vec{F}(t) = \gamma(t) m \vec{a}(t)$, so dass diese Leistung dadurch noch als

$$\mathcal{P}_0 = \frac{2}{3} \frac{q^2}{4\pi \epsilon_0 m^2 c^3} \gamma_{\text{ret.}}^2 |\vec{F}(t_{\text{ret.}})|^2.$$

ausgedrückt werden kann. Die abgestrahlte Leistung ist also größer um einen Faktor γ^2 relativ zum Fall einer Linearbewegung mit der gleichen Kraft. Somit wird mehr Energie als sog. *Synchrotron-*

strahlung in einem kreisförmigen Teilchenbeschleuniger (einem „Speicherring“, wie z.B. das LHC am CERN) verloren,⁽⁷⁹⁾ als in einem Linearbeschleuniger.

Bemerkung: Wenn ein geladenes Teilchen auf Materie stößt und dort von Wechselwirkungen mit den Atomen abgebremst wird, strahlt es ab. Die entsprechende Strahlung wird *Bremsstrahlung* (auch auf Englisch!) genannt.

Literatur zum Kapitel IX

- Feynman, *Vorlesungen über Physik. Band 2* [22] = *Lectures on Physics. Vol. II* [23] Kap. 18, 20–21, 27 & 29;
Vorlesungen über Physik. Band 1 [24] = *Lectures on Physics. Vol. I* [25] Kap. 34.
- Fließbach, *Elektrodynamik* [3] Teil IV Kap. 16–17 & Teil V Kap. 20, 23–24.
- Greiner, *Klassische Elektrodynamik* [8] Teil IV Kap. 12–13, 15, 20–21.
- Griffiths, *Elektrodynamik* [9] = *Introduction to Electrodynamics* [10] Kap. 7–11.
- Jackson, *Klassische Elektrodynamik* [11] = *Classical Electrodynamics* [12] Kap. 6.1–6.5, 6.7, 7.1–7.2, 9.1–9.2 & 14.1–14.4.
- Landau & Lifschitz, *Klassische Feldtheorie* [14] = *The classical theory of fields* [28] Kap. 4 § 26, 29–31, 33, Kap. 6 § 46–48, Kap. 8 § 62–64 & Kap. 9 § 66–67 & 73–76.
- Nolting, *Elektrodynamik* [17] Kap. 4.1, 4.3 & 4.5.
- Scheck, *Klassische Feldtheorie* [19] Kap. 1.3–1.6, 2.3, 3.5–3.6 & 4.1–4.2.

⁽⁷⁹⁾Oder genauer, abgestrahlt: die emittierte Strahlung kann benutzt werden, um Materie zu untersuchen.

Dritter Teil

Grundlagen der

Speziellen Relativitätstheorie

KAPITEL X

Mathematischer Apparat der Speziellen Relativitätstheorie

Einleitung

X.1 Einstein'sche Postulate

X.1.1 Motivation

Trotz der vielen Erfolge der klassischen Newton'schen Mechanik tritt ein wichtiges Problem auf, wenn man Effekte aus dem Elektromagnetismus betrachtet. Aus den Maxwell^(as)-Gleichungen folgert man leicht, dass die Ausbreitungsgeschwindigkeit im Vakuum von elektromagnetischen Wellen — insbesondere von Licht, das laut Arbeiten von Hertz^(at) ein elektromagnetisches Phänomen ist — denselben Wert (c) in allen Inertialsystemen annimmt. In der Tat zeigten die Interferenz-Experimente von Michelson^(au) und Morley^(av), dass die Lichtgeschwindigkeit parallel oder senkrecht zur (instantanen) Bewegungsrichtung der Erde um die Sonne gleich bleibt.

Die theoretische Vorhersage sowie ihre experimentelle Bestätigung⁽⁸⁰⁾ lassen sich aber offensichtlich nicht mit dem Additionstheorem für Geschwindigkeiten in der Newton'schen Raumzeit vereinbaren.

X.1.2 Einstein'sche Postulate

Um die theoretisch und experimentell gefundene Konstanz der Vakuumlichtgeschwindigkeit zu berücksichtigen hat Einstein seine Postulate eingeführt [?], welche die durch Newton postulierten Gesetze ersetzen müssen.

X.1.2a Erstes Postulat

Das erste dieser Postulate ist das (Einstein'sche) *Relativitätsprinzip*

Die Gesetze der Physik nehmen in allen Inertialsystemen die gleiche Form an. (X.1)

Dabei handelt es sich um eine Erweiterung des sog. Galilei-Relativitätsprinzip — das eigentlich durch Newton formuliert wurde —, das aber nur für mechanische System galt. Mit seiner Verallgemeinerung meinte insbesondere Einstein, dass auch elektromagnetische Phänomene diesem Prinzip genügen sollen.

X.1.2b Zweites Postulat

⁽⁸⁰⁾... vorausgesetzt, die Michelson–Morley-Experimente wirklich das Fortpflanzen von Licht im Vakuum messen, nicht in irgendeinem hypothetischen Medium — dem Äther —, wie damals vorgeschlagen wurde, um die Newton-Mechanik zu retten.

^(as)J. C. MAXWELL, 1831–1879 ^(at)H. HERTZ, 1857–1894 ^(au)A. MICHELSON, 1852–1931 ^(av)E. MORLEY, 1838–1923

Mit seinem zweiten Postulat entschied sich Einstein gegen die Newton'schen Gesetze mit ihren Invarianz unter Galilei-Transformationen, und legte mehr Wert auf die Maxwell-Gleichungen und

deren experimentellen Bestätigungen, indem er die Konstanz der Lichtgeschwindigkeit als grundsätzlich erklärte:

Die Lichtgeschwindigkeit im Vakuum ist in allen Inertialsystemen gleich. (X.2)

Demzufolge kann das klassische Additionsgesetz für Geschwindigkeiten nicht mehr gelten. Dies bedeutet wiederum, dass Galilei-Transformationen, insbesondere Galilei-Booster, nicht die richtigen Transformationen zwischen Inertialsystemen sind. Wegen dieses Prinzip muss also die ganze Kinematik neu geschrieben werden.

X.2 Lorentz-Transformationen

In diesem Abschnitt sind \mathcal{B} und \mathcal{B}' zwei Inertialsysteme, wobei das letztere sich relativ zum ersteren mit der Geschwindigkeit \vec{u} bewegt. t und \vec{r} bzw. t' und \vec{r}' bezeichnen die Zeit und den Ortsvektor gemessen durch einen in \mathcal{B} bzw. \mathcal{B}' ruhenden Beobachter.

X.2.1 Linienelement

Seien P_1 und P_2 zwei Punkte, zwischen denen Vakuum herrscht. Zu einem ersten Zeitpunkt wird Licht in P_1 emittiert. Für den Beobachter in \mathcal{B} geschieht dies zur Zeit t_1 bei der Position \vec{r}_1 ; für den Beobachter in \mathcal{B}' , zur Zeit t'_1 bei der Position \vec{r}'_1 . Das Licht wird zu einem späteren Zeitpunkt in P_2 detektiert, und zwar für den Beobachter in \mathcal{B} bzw. \mathcal{B}' zur Zeit t_2 bzw. t'_2 bei der Position \vec{r}_2 bzw. \vec{r}'_2 .

Laut dem zweiten Einstein'schen Postulat muss die Vakuumlichtgeschwindigkeit gleich c für beide Beobachter sein, d.h.

$$\frac{|\vec{r}_2 - \vec{r}_1|}{t_2 - t_1} \equiv \frac{|\Delta\vec{r}|}{\Delta t} = c \quad \text{und} \quad \frac{|\vec{r}'_2 - \vec{r}'_1|}{t'_2 - t'_1} \equiv \frac{|\Delta\vec{r}'|}{\Delta t'} = c$$

oder äquivalent $|\Delta\vec{r}| = c\Delta t$ und $|\Delta\vec{r}'| = c\Delta t'$, oder noch

$$(\Delta\vec{r})^2 = c^2\Delta t^2 \quad \text{und} \quad (\Delta\vec{r}')^2 = c^2\Delta t'^2.$$

Betrachtet man infinitesimal benachbarte Punkte P_1 und P_2 , und dementsprechend ein infinitesimales Zeitintervall für die Lichtpropagation, so kann man die Änderungen $\Delta\vec{r}$, Δt , usw. durch kleine Variationen $d\vec{r}$, dt , usw. ersetzen. Dann lässt sich die obige Beziehung als

$$d\vec{r}^2 - c^2 dt^2 = d\vec{r}'^2 - c^2 dt'^2 = 0.$$

umschreiben. Sei

$$ds^2 \equiv -c^2 dt^2 + d\vec{r}^2. \quad (\text{X.3a})$$

Gemäß der obigen Diskussion muss dieses infinitesimale *Linienelement* dasselbe — und hier gleich Null — für beide Beobachter sein, damit beide dieselbe Lichtgeschwindigkeit im Vakuum messen.

Seien jetzt P_1 , P_2 zwei beliebige Punkte derart, dass in beiden Orten zu gegebenen Zeitpunkten etwas passiert. Für den Beobachter in \mathcal{B} bzw. \mathcal{B}' findet das „erste“ Ereignis zur Zeit t_1 bzw. t'_1 statt, als sich der Punkt P_1 bei der Position \vec{r}_1 bzw. \vec{r}'_1 befindet. Wiederum sind die jeweiligen Zeit- und Ortskoordinaten des „zweiten“ Ereignisses $t_2 \equiv t_1 + dt$ und $\vec{r}_2 \equiv \vec{r}_1 + d\vec{r}$ bzw. $t'_2 \equiv t'_1 + dt'$ und $\vec{r}'_2 \equiv \vec{r}'_1 + d\vec{r}'$. Dabei wird nichts über die Vorzeichen von dt und dt' angenommen.

Das Linienelement ds^2 bzw. ds'^2 zwischen den beiden Ereignissen für den Beobachter in \mathcal{B} bzw. \mathcal{B}' kann wieder über Gl. (X.3a) definiert werden. Wenn \mathcal{B} und \mathcal{B}' beide Inertialsysteme sind, dann müssen ds^2 und ds'^2 gleich sein.

Diese Anforderung, die hier postuliert wird, wurde oben heuristisch hergeleitet für den Fall zweier Ereignisse, welche den aufeinanderfolgenden Emission und Detektion von Licht entsprechen.

X.2.2 Lorentz-Transformationen

In der Transformation der Zeit- und Ortskoordinaten zwischen zwei Inertialsystemen muss das oben definierte Linienelement ds^2 invariant bleiben.

Definition: Die linearen Transformationen

$$\begin{pmatrix} c dt' \\ d\vec{r}' \end{pmatrix} = \Lambda \begin{pmatrix} c dt \\ d\vec{r} \end{pmatrix} \quad (\text{X.4})$$

mit Λ einer 4×4 -Matrix, die das Linienelement $ds^2 \equiv -c^2 dt^2 + d\vec{r}^2$ invariant lassen, heißen *Lorentz-Transformationen*.

Es gibt zwei Haupttypen solcher Transformationen, und zwar einerseits Drehungen (§ X.2.2 a) — unter welchen das Zeitintervall dt sich nicht ändert — und Lorentz-Boosts (§ X.2.2 b).

X.2.2 a Drehungen

Einfache Lorentz-Transformationen sind die Drehungen der Raumkoordinaten, welche die Zeit invariant lassen, und somit den Drehungen im dreidimensionalen Ortsraum der nicht-relativistischen Mechanik entsprechen. Solche Transformationen gelten für zwei Koordinatensysteme, die sich relativ zueinander nicht bewegen, deren Achsen aber nicht parallel sind.

Die assoziierte Transformationsmatrix lautet

$$\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & \mathcal{R} \end{pmatrix} \quad (\text{X.5})$$

mit \mathcal{R} einer 3×3 -Drehmatrix. In der Tat führt

$$\begin{pmatrix} c dt' \\ d\vec{r}' \end{pmatrix} = \begin{pmatrix} c dt \\ \mathcal{R} d\vec{r} \end{pmatrix} \quad (\text{X.6})$$

mit der Bedingung $\mathcal{R}^T \mathcal{R} = \mathbb{1}_3$ sofort zu $d\vec{r}'^2 - c^2 dt'^2 = d\vec{r}^T \mathcal{R}^T \mathcal{R} d\vec{r} - c^2 dt^2 = d\vec{r}^2 - c^2 dt^2$.

X.2.2 b Lorentz-Boosts

Sei jetzt angenommen, dass die Achsen des Koordinatensystems von \mathcal{B}' parallel zu den Achsen von \mathcal{B} sind. Jetzt bewegt sich \mathcal{B}' relativ zu \mathcal{B} mit Geschwindigkeit \vec{u} . Der Einfachheit halber wird angenommen, dass die $x = x'$ -Richtung der Koordinaten entlang der Bewegungsrichtung liegt, d.h. $\vec{u} = u \vec{e}_x$.

Da nichts in den Richtungen $y = y'$ und $z = z'$ senkrecht zur Bewegung passiert, bleiben die Koordinaten eines Punkts entlang dieser Achsen unverändert in der Transformation von \mathcal{B} nach \mathcal{B}' . Demzufolge gelten $dy' = dy$ und $dz' = dz$ für das Intervall zwischen benachbarten Punkten. Diese senkrechten Richtungen sollen sich auch nicht mit den anderen Koordinaten mischen, d.h. dt' , dx' sind unabhängig von dy , dz .

Daher ist die gesuchte Matrix der Transformation der Art

$$\begin{pmatrix} c dt' \\ dx' \\ dy' \\ dz' \end{pmatrix} = \begin{pmatrix} A & B & 0 & 0 \\ C & D & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c dt \\ dx \\ dy \\ dz \end{pmatrix}$$

mit A, B, C, D drei reellen Zahlen. Dann lautet das Linienelement in \mathcal{B}'

$$\begin{aligned} ds'^2 &= -(Ac dt + B dx)^2 + (Cc dt + D dx)^2 + dy^2 + dz^2 \\ &= -(A^2 - C^2)c^2 dt^2 + (D^2 - B^2) dx^2 + 2(CD - AB)c dt dx + dy^2 + dz^2, \end{aligned}$$

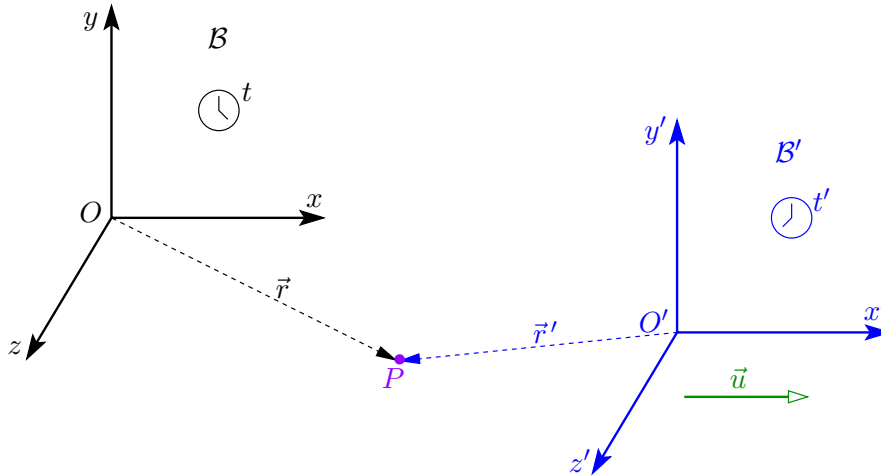


Abbildung X.1

was gleich $ds^2 = -c^2 dt^2 + dx^2 + dy^2 + dz^2$ sein muss. Beliebige dt und $d\vec{r}$ genügen dieser Gleichung nur genau dann, wenn das System

$$\begin{cases} A^2 - C^2 = 1 \\ D^2 - B^2 = 1 \\ CD - AB = 0 \end{cases}$$

erfüllt wird.

Seien $A \equiv \gamma$ und $B \equiv -\beta\gamma$. Dann stellen sie mit $C = B = -\beta\gamma$ und $D = A = \gamma$ eine Lösung des Systems dar, wenn $\gamma = 1/\sqrt{1 - \beta^2}$ ist. Dazu soll $\beta = |\vec{u}|/c$, um den richtigen nicht-relativistischen Limes $u = |\vec{u}| \ll c$ wiederzufinden (s. unten). Schließlich lautet die Matrix für ein *Lorentz-Boost*, auch *spezielle Lorentz-Transformation* genannt,

$$\Lambda = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{X.7a}$$

mit

$$\gamma = \frac{1}{\sqrt{1 - \beta^2}} \quad , \quad \beta \equiv \frac{u}{c} \tag{X.7b}$$

γ heißt *Lorentz-Faktor* des Boosts.

Im *nicht-relativistischen Limes* $u \ll c$, d.h. $\beta \ll 1$, gibt eine Taylor-Entwicklung

$$\frac{1}{\sqrt{1 - \beta^2}} \underset{\beta \ll 1}{\sim} 1 + \frac{\beta^2}{2} + \mathcal{O}(\beta^4), \tag{X.8}$$

so dass $\beta\gamma \sim \beta + \mathcal{O}(\beta^3)$ ist. Aus Gl. (X.4) und der Form (X.7a) des Lorentz-Boosts folgen dann

$$c dt' = c dt + \mathcal{O}(\beta) \quad \text{und} \quad dx' = -\beta c dt' + dx + \mathcal{O}(\beta^2).$$

Die erste Gleichung gibt $t' = t + \text{Konstante}$, während die zweite mit $\beta = u/c$ zu $dx' = dx - u dt$ oder äquivalent $x' = x - ut$ führt, entsprechend dem bekannten nicht-relativistischen Ergebnis für ein Galilei-Boost entlang der x -Achse.

Bemerkung: *Rapidity* des Lorentz-Boosts... später!

$$\xi_u \equiv \text{artanh} \frac{u}{c} = \frac{1}{2} \ln \frac{c + u}{c - u}. \tag{X.9}$$

X.2.2c Lorentz-Gruppe

Allgemeiner bilden die linearen Transformationen $(ct, \vec{r}) \rightarrow (ct', \vec{r}')$, die das Linienelement ds^2 [Gl. (X.3a)] invariant lassen, eine Gruppe, die *Lorentz-Gruppe*.

Wie wir unten (§ X.2.4b) sehen werden, lassen sich diese Transformationen noch äquivalent dadurch charakterisieren, dass ihre Matrixdarstellung Λ die Gleichung

$$\eta = \Lambda^T \eta \Lambda \quad (\text{X.10})$$

erfüllen, wobei η in Gl. (X.18) definiert ist. Dass die Menge solcher Matrizen versehen mit der üblichen Matrixmultiplikation eine Gruppe bilden ist dann schnell bewiesen:

- Die 4×4 -Einheitsmatrix $\mathbb{1}_4$ — entsprechend der Identitätstransformation — genügt die Beziehung (X.10).
- Wenn Λ_1, Λ_2 die Gl. (X.10) erfüllen, dann gilt das noch für ihr Produkt $\Lambda_1 \Lambda_2$.
- Aus Gl. (X.10) folgt $|\det \Lambda| = 1$, d.h. Λ ist invertierbar; die inverse Matrix Λ^{-1} erfüllt trivial die definierende Beziehung.

Die Lorentz-Gruppe wird oft mit $O(1,3)$ bezeichnet.

Sowohl die Drehungen des § X.2.2a als die Lorentz-Boosts des § X.2.2b, die je eine Untergruppe der Lorentz-Gruppe bilden, sind sog. *eigentliche orthochrone Lorentz-Transformationen*, Elemente der *eigentlichen orthochronen Lorentz-Gruppe* $SO^+(1,3)$.⁽⁸¹⁾ Bei den *eigentlichen Lorentz-Transformationen* handelt es sich um solche mit Determinante +1; diese bilden die *eigentliche Lorentz-Gruppe*, die mit $SO(1,3)$ bezeichnet wird.

Wiederum sind die *orthochronen Lorentz-Transformationen* diejenigen, welche die Richtung der Zeitkomponente nicht ändern. Das heißt, das Matrixelement oben links in der Matrixdarstellung Λ — das in Abschn. X.3 mit Λ^0_0 bezeichnet wird — muss positiv, und in der Tat größer gleich 1, sein. Diese Transformationen bilden die *orthochrone Lorentz-Gruppe* $O^+(1,3)$.⁽⁸¹⁾

Beispiele von uneigentlichen ($\det \Lambda = -1$) Lorentz-Transformationen sind einerseits die *Raumspiegelung* $(ct, \vec{r}) \rightarrow (ct' = ct, \vec{r}' = -\vec{r})$ mit der Matrixdarstellung

$$\Lambda_P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (\text{X.11})$$

und andererseits die *Zeitumkehr* $(ct, \vec{r}) \rightarrow (ct' = -ct, \vec{r}' = \vec{r})$ mit Matrixdarstellung

$$\Lambda_T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (\text{X.12})$$

die offensichtlich auch nicht orthochron ist. Dann ist die (kommutierende) Verknüpfung beider Transformationen, mit Matrixdarstellung $\Lambda_P \Lambda_T = -\mathbb{1}_4$, eine nicht-orthochrone eigentliche Lorentz-Transformation.

Mathematisch stellen die eigentliche orthochrone Lorentz-Gruppe $SO^+(1,3)$ und die drei Mengen $SO^-(1,3) \equiv \{\Lambda_P \Lambda_T \Lambda, \Lambda \in SO^+(1,3)\}$, $\{\Lambda_P \Lambda, \Lambda \in SO^+(1,3)\}$ und $\{\Lambda_T \Lambda, \Lambda \in SO^+(1,3)\}$ die vier Zusammenhangskomponenten der Lorentz-Gruppe $O(1,3)$ dar, entsprechend jeweils den vier Möglichkeiten $++, +-, -+, --$ für die Vorzeichen der Determinante $\det \Lambda$ und des Matrixelements Λ^0_0 .

⁽⁸¹⁾ Statt $O^+(1,3)$ bzw. $SO^+(1,3)$ wird auch die Notation $SO^\uparrow(1,3)$ bzw. $SO^\dagger(1,3)$ benutzt.

X.2.3 Folgerungen

X.2.3 a Zeitdilatation

X.2.3 b Längenkontraktion

→ Aufgabe 30

X.2.3 c Additionstheorem für Geschwindigkeiten

→ Aufgabe 34

X.2.4 Minkowski-Raum

Wegen der Absolutheit von Zeit und Raum in der Newton'schen Mechanik faktorisiert sich die zugehörige nicht-relativistische Raumzeit in das Produkt einer eindimensionalen Zeit-Geraden mit dem dreidimensionalen euklidischen Ortsraum \mathcal{E}_3 . Dagegen mischen die Lorentz-Boosts die Zeit und die Ortskomponenten eines Beobachters miteinander. Demzufolge lohnt es sich, die Zeit und die drei Ortskoordinaten als Komponenten eines Vektors in einer vierdimensionalen reellen Raumzeit \mathcal{M}_4 .

X.2.4 a Viererortsvektor

Ein Punkt in \mathcal{M}_4 , entsprechend einem Zeitpunkt und einem Ort, wird *Ereignis* genannt. Sei \mathcal{B} ein Bezugssystem, versehen mit einem Koordinatensystem, wobei die räumliche Koordinaten kartesisch sind. Nach deren Angabe lässt sich ein Ereignis durch seinen *Viererortsvektor* \mathbf{x} mit Komponenten x^μ mit $\mu = 0, 1, 2, 3$ charakterisieren, wobei $x^0 = ct$, während die x^i mit $i = 1, 2, 3$ die Koordinaten des Ortsvektors \vec{r} sind. Der Kurze halber wird hiernach für die Zerlegung in Zeit- und Ortskoordinaten die Notation

$$\mathbf{x} \equiv \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} ct \\ \vec{r} \end{pmatrix} \quad (\text{X.13})$$

benutzt.

Definition: Die Trajektorie eines Massenpunktes in der Raumzeit heißt *Weltlinie*.

Betrachte man einen bewegten Massenpunkt und \mathcal{B}_0 ein Inertialsystem, das momentan mit dem Ruhesystem des Massenpunktes übereinstimmt. Das heißt, dass der Massenpunkt relativ zu \mathcal{B}_0 momentan ruht. Sei τ die Zeit in \mathcal{B}_0 ; τ heißt *Eigenzeit* des Massenpunktes, und das Linienelement entlang der Weltlinie des Massenpunktes lässt sich momentan als $ds^2 = -c^2 d\tau^2$ schreiben.

Sei t die Zeit in einem Inertialsystem \mathcal{B} , relativ zu welchem der Massenpunkt (oder äquivalent das Inertialsystem \mathcal{B}_0) sich mit Geschwindigkeit \vec{v} bewegt. Das Eigenzeitelement $d\tau$ hängt mit dem infinitesimalen Zeitintervall dt in \mathcal{B} über

$$d\tau = \sqrt{1 - \frac{\vec{v}^2}{c^2}} dt \quad (\text{X.14})$$

zusammen, wie sich mithilfe der Lorentz-Transformation von \mathcal{B}_0 nach \mathcal{B} finden lässt.

X.2.4 b Metrischer Tensor

Bezeichnet man dann mit dx die Verschiebung zwischen zwei benachbarten Ereignissen in \mathcal{M}_4 , dann lässt sich das Linienelement (X.3a) als

$$ds^2 = -(dx^0)^2 + (dx^1)^2 + (dx^2)^2 + (dx^3)^2 \quad (\text{X.15})$$

schreiben, mit dx^μ , $\mu = 0, 1, 2, 3$ den Koordinaten von dx .

Für $\mu, \nu = 0, 1, 2, 3$ definiert man $\eta_{\mu\nu}$ durch

$$\begin{cases} \eta_{00} = -1 \\ \eta_{ii} = +1 \text{ für } i = 1, 2, 3 \\ \eta_{\mu\nu} = 0 \text{ für } \mu \neq \nu. \end{cases} \quad (\text{X.16})$$

Mithilfe dieser Zahlen lautet das Linienelement (X.15) noch

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu, \quad (\text{X.17})$$

wobei die Summe über alle Werte von μ und ν von 0 bis 3 nicht geschrieben wurde.

In Matrixdarstellung wird $\eta_{\mu\nu}$ zum Element der $(\mu+1)$ -ten Zeile und $(\nu+1)$ -ten Spalte einer 4×4 -Matrix:

$$\eta = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (\text{X.18})$$

und das Linienelement wird zu

$$ds^2 = dx^T \eta dx \quad (\text{X.19})$$

mit dx^T dem Zeilenvektor mit Elementen dx^μ .

Die 16 Zahlen $\eta_{\mu\nu}$ sind die Komponenten des *metrischen Tensors*⁽⁸²⁾ $\boldsymbol{\eta}$. Dieser dient dazu, ein (pseudo-)skalares Produkt zu definieren, vgl. § X.3.2 c, welches unter Lorentz-Transformationen der Koordinaten invariant bleibt. Wiederum bedeutet dies, dass der metrische Tensor unter solchen Koordinatentransformationen $x \rightarrow x' = \Lambda x$ bzw. $dx \rightarrow dx' = \Lambda dx$ invariant bleibt, d.h. dass seine Komponenten gleich bleiben. nach Einsatz von dx' und dx'^T in Beziehung (X.19) ergibt sich die Matrixgleichung

$$\Lambda^T \eta \Lambda = \eta. \quad (\text{X.20})$$

Die Raumzeit \mathcal{M}_4 versehen mit dem metrischen Tensor $\boldsymbol{\eta}$ heißt *Minkowski*^(aw)-*Raum*. Wiederum ist ein System von *Minkowski-Koordinaten* ein Koordinatensystem in \mathcal{M}_4 , in welchem die Komponenten des metrischen Tensors die einfache Form (X.16) annehmen⁽⁸³⁾ — was im Folgenden immer der Fall sein wird.

Bemerkungen:

* Ein Merkmal des Minkowski-Raums \mathcal{M}_4 der Speziellen Relativitätstheorie ist die Existenz von *globalen* Minkowski-Koordinatensystemen, deren Koordinaten überall und zu jeder Zeit gelten — obwohl \mathcal{M}_4 mathematisch kein Vektorraum, sondern eine vierdimensionale Mannigfaltigkeit ist.

* In diesem Skript wurde der metrische Tensor nach den Lorentz-Transformationen eingeführt; jedoch die logische Herangehensweise, die oben nur implizit benutzt wurde, geht eher in die andere Richtung: der metrische Tensor wird zunächst postuliert, dann sind die Lorentz-Transformationen die zugehörigen *Isometrien*, d.h. die linearen Koordinaten, die das entsprechende Skalarprodukt invariant lassen. Dementsprechend kann die Matrixgleichung (X.20) eigentlich als die definierende Beziehung der Lorentz-Transformationen betrachtet werden.

* Hier wurden für den metrischen Tensor und seine natürliche Matrixdarstellung unterschiedliche Notationen $\boldsymbol{\eta}$ bzw. η verwendet, um den Unterschied zwischen den beiden mathematischen Objekten zu betonen.

⁽⁸²⁾Oder kurz *Metrik*.

⁽⁸³⁾Das heißt, dass die zugehörigen Basisvektoren ein orthonormiertes System bilden.

^(aw)H. MINKOWSKI, 1864–1909

X.3 Vierervektoren und Vierertensoren

Die in Abschn. X.2 eingeführten Lorentz-Transformationen geben die Beziehungen zwischen den Minkowski-Koordinaten eines gegebenen Ereignisses bezüglich zwei verschiedener Inertialsysteme \mathcal{B} und \mathcal{B}' . In diesem Übergang von \mathcal{B} nach \mathcal{B}' können sich die mathematischen Darstellungen physikalischer Größen nicht nur wie der Viererortsvektor x , sondern auch anders transformieren.

Der Kürze halber wird hiernach die Redensart „unter Lorentz-Transformationen“ statt „im Übergang von einem Inertialsystem zu einem anderen“ verwendet, wobei die besagte Transformation diejenige für die Viererortsvektoren [Gl. (X.22)] ist.

X.3.1 Lorentz-Skalare

Definition: Ein *Lorentz-Skalar* oder *Viererskalar* ist eine Größe, die invariant unter Transformationen von einem Inertialsystem zu einem anderen ist.

Ein erstes Beispiel davon ist (definitionsgemäß!) das Linienelement (X.3a) bzw. (X.17). Demzufolge ist das Eigenzeitelement $d\tau$ ebenfalls ein Lorentz-Skalar.

Ein weiteres Beispiel ist das *Vierervolumenelement* d^4x , definiert durch

$$d^4x \equiv dx^0 dx^1 dx^2 dx^3. \quad (\text{X.21})$$

Wenn der 4-Ortsvektor sich gemäß $x \rightarrow x' = \Lambda x$ transformiert, entsprechend Gl. (X.22b) für die Koordinaten, dann transformiert sich das 4-Volumenelement gemäß

$$d^4x \rightarrow d^4x' = |\det \Lambda| d^4x,$$

woraus das Ergebnis dank $|\det \Lambda| = 1$ folgt.

Bemerkung: Man definiert auch *Lorentz-Pseudoskalare*, die invariant unter der eigentlichen Lorentz-Gruppe sind, jedoch unter uneigentlichen Transformationen wie Raumspiegelung oder Zeitumkehr ihr Vorzeichen ändern.

X.3.2 Vierervektoren

Der Viererortsvektor x mit Komponenten x^μ [Gl. (X.13)] ist ein erstes Beispiel von Vierervektor, der sich im Übergang von einem Inertialsystem \mathcal{B} nach einem zweiten Inertialsystem \mathcal{B}' gemäß

$$x \rightarrow x' = \Lambda x \quad (\text{X.22a})$$

mit Λ einer Lorentz-Transformation transformiert, entsprechend für die Koordinaten

$$x^\mu \rightarrow x'^{\mu'} = \Lambda^{\mu'}_{\nu} x^\nu \quad \text{für } \mu' = 0', 1', 2', 3'. \quad (\text{X.22b})$$

Wie es sich herausstellt, gibt es zwei unterschiedliche Arten von Vierervektoren, die sich entweder wie x oder wie der damit assoziierte Vierergradient transformieren.

Bemerkungen:

* Aus Gl. (X.22b) folgt die Beziehung

$$\Lambda^{\mu'}_{\nu} = \frac{\partial x^{\mu'}}{\partial x^\nu} \quad \forall \mu', \nu. \quad (\text{X.23})$$

* In der Tat sind die Lorentz-Transformationen nicht die einzigen linearen Transformationen, die das Linienelement (X.3a) invariant lassen. Daneben gibt es noch die affinen Transformationen der *Poincaré*^(ax)-Gruppe (oder *inhomogenen Lorentz-Gruppe*) $ISO(1,3)$, der Form

$$x \rightarrow x' = \Lambda x + a \quad (\text{X.24a})$$

^(ax)H. POINCARÉ, 1854–1912

mit \mathbf{a} einem Vierervektor; komponentenweise lautet diese Transformation

$$x^\mu \rightarrow x^{\mu'} = \Lambda^{\mu'}_\nu x^\nu + a^{\mu'}. \quad (\text{X.24b})$$

Die Poincaré-Transformationen mit $\Lambda = \mathbb{1}_4$ und beliebigem \mathbf{a} sind offensichtlich die Raumzeit-Translationen.

X.3.2a Kontravariante Vierervektoren

Definition: Ein sog. *kontravarianter Vierervektor* \mathbf{V} ist eine Menge aus vier Zahlen V^μ , die sich unter Lorentz-Transformationen wie die Koordinaten x^μ des Viererortsvektors transformieren, d.h.

$$V^\mu \rightarrow V^{\mu'} = \Lambda^{\mu'}_\nu V^\nu, \quad (\text{X.25a})$$

oder in Matrixform

$$\mathbf{V} \rightarrow \mathbf{V}' = \Lambda \mathbf{V}. \quad (\text{X.25b})$$

Ein Beispiel ist die *Vierergeschwindigkeit* \mathbf{u} , eines massiven Teilchens (Massenpunktes) entlang seiner Weltlinie $\mathbf{x}(s)$, die als

$$\mathbf{u} \equiv \frac{d\mathbf{x}(s)}{d\tau} \quad (\text{X.26a})$$

definiert ist, mit s einer Parametrisierung der Weltlinie und τ der Eigenzeit des Teilchens. Komponentenweise lautet diese Definition

$$u^\mu \equiv \frac{dX^\mu(s)}{d\tau} \quad (\text{X.26b})$$

mit u^μ bzw. $x^\mu(s)$ den Koordinaten der Vierergeschwindigkeit bzw. der Weltlinie. Dass \mathbf{u} ein kontravarianter Vierervektor ist, folgt aus der Tatsache, dass das Eigenzeitelement $d\tau$ im Zähler ein Lorentz-Skalar ist.

Wenn v^i , $i = 1, 2, 3$ die Komponenten der dreidimensionalen Geschwindigkeit \vec{v} des Teilchens und γ den damit assoziierten Lorentz-Faktor bezeichnen, dann lauten die Koordinaten der Vierergeschwindigkeit $u^0 = \gamma c$ und $u^i = \gamma v^i$ für $i = 1, 2, 3$, d.h.

$$\mathbf{u} = \begin{pmatrix} \gamma c \\ \gamma \vec{v} \end{pmatrix}. \quad (\text{X.27})$$

Bemerkung: Für ein masseloses Teilchen, z.B. ein Photon, wird die Vierergeschwindigkeit nicht definiert, da solche Teilchen kein Ruhesystem haben, in welchem die Eigenzeit definiert werden kann: sie bewegen sich entlang lichtartige Weltlinien (s. unten).

X.3.2b Kovariante Vierervektoren

Seien $\partial_\mu \equiv \partial / \partial x^\mu$ die Komponenten des *Vierergradient*-Operators:

$$\partial_0 \equiv \frac{\partial}{\partial x^0} = \frac{1}{c} \frac{\partial}{\partial t}, \quad \partial_i \equiv \frac{\partial}{\partial x^i} \quad \text{für } i = 1, 2, 3, \quad (\text{X.28})$$

wobei die drei räumlichen Komponenten genau die Koordinaten des dreidimensionalen Gradienten sind.

Mithilfe der Kettenregel gilt unter der Transformation (X.22)

$$\partial_\nu \equiv \frac{\partial}{\partial x^\nu} = \frac{\partial x^{\rho'}}{\partial x^\nu} \frac{\partial}{\partial x^{\rho'}} = \Lambda^{\rho'}_\nu \frac{\partial}{\partial x^{\rho'}} = \Lambda^{\rho'}_\nu \partial_{\rho'}. \quad (\text{X.29})$$

Sei $\Lambda^\mu_{\rho'}$ das (μ, ρ') -Element der inversen Matrix Λ^{-1} , d.h.

$$\Lambda^\mu_{\rho'} \Lambda^{\rho'}_\nu = \delta^\mu_\nu \quad \text{und} \quad \Lambda^{\mu'}_\rho \Lambda^\rho_{\nu'} = \delta^{\mu'}_{\nu'}. \quad (\text{X.30})$$

Multipliziert man Gl. (X.29), die für jedes $\nu \in \{0, 1, 2, 3\}$ gilt, mit $\Lambda^\nu_{\mu'}$ und summiert man über ν ,

so ergibt sich

$$\Lambda^{\nu}_{\mu'} \partial_{\nu} = \Lambda^{\nu}_{\mu'} \Lambda^{\rho'}_{\nu} \partial_{\rho'} = \delta^{\rho'}_{\mu'} \partial_{\rho'},$$

d.h.

$$\partial_{\mu'} = \Lambda^{\nu}_{\mu'} \partial_{\nu}. \quad (\text{X.31})$$

Somit transformiert sich der Vierergradient mit Λ^{-1} , wenn der Ortsvektor sich mit Λ transformiert.

Definition: Ein *kovarianter Vierervektor* ist eine Menge aus vier Größen W_{μ} , die sich unter Lorentz-Transformationen wie die Komponenten ∂_{μ} des Vierergradienten verhalten, d.h.

$$W_{\mu} \rightarrow W_{\mu'} = \Lambda^{\nu}_{\mu'} W_{\nu}. \quad (\text{X.32})$$

Jedem kontravarianten Vierervektor mit Koordinaten V^{μ} kann man einen kovarianten Vierervektor mit Komponenten V_{μ} gemäß

$$V_{\mu} = \eta_{\mu\nu} V^{\nu} \quad (\text{X.33})$$

zuordnen.

In einer Basistransformation ändern sich V^{ν} bzw. $\eta_{\mu\nu}$ gemäß

$$V^{\nu} \rightarrow V^{\nu'} = \Lambda^{\nu'}_{\rho} V^{\rho} \quad \text{bzw.} \quad \eta_{\mu\nu} \rightarrow \eta_{\mu'\nu'} = \Lambda^{\mu}_{\mu'} \Lambda^{\nu}_{\nu'} \eta_{\mu\nu}$$

vgl. Gl. (X.25a) bzw. (X.42). Dann kommt

$$V_{\mu} \rightarrow V_{\mu'} = \eta_{\mu'\nu'} V^{\nu'} = \Lambda^{\mu}_{\mu'} \Lambda^{\nu}_{\nu'} \eta_{\mu\nu} \Lambda^{\nu'}_{\rho} V^{\rho} = \Lambda^{\mu}_{\mu'} \eta_{\mu\nu} \delta^{\nu}_{\rho} V^{\rho},$$

wobei die letzte Gleichung aus Gl. (X.30) folgt. Schreibt man dann $\eta_{\mu\nu} \delta^{\nu}_{\rho} V^{\rho} = \eta_{\mu\nu} V^{\nu} = V_{\mu}$, so ergibt sich

$$V_{\mu} \rightarrow V_{\mu'} = \Lambda^{\mu}_{\mu'} V_{\mu},$$

was zu beweisen war. □

Umgekehrt kann mit jedem kovarianten Vierervektor mit Komponenten V_{μ} der kontravariante Vierervektor mit Komponenten

$$V^{\mu} = \eta^{\mu\nu} V_{\nu} \quad (\text{X.34})$$

assoziiert werden, mit $\eta^{\mu\nu}$ den Komponenten des inversen metrischen Tensors $\boldsymbol{\eta}^{-1}$, die numerisch gleich den $\eta_{\mu\nu}$ sind, vgl. § X.3.3 b. Unter Verwendung des metrischen Tensors (X.16) findet man

$$V_0 = -V^0, \quad V_1 = V^1, \quad V_2 = V^2, \quad V_3 = V^3. \quad (\text{X.35})$$

Bemerkungen:

* Bezeichnet man den 4-komponentigen Zeilenvektor mit Koordinaten W_{μ} mit \underline{W} , so lässt sich die Transformation (X.32) noch in Matrixdarstellung als

$$\underline{W} \rightarrow \underline{W}' = \underline{W} \Lambda^{-1} \quad (\text{X.36})$$

schreiben. Dann lautet Gl. (X.33) bzw. (X.34)

$$\underline{V} = \mathbf{V}^T \boldsymbol{\eta}^T \quad \text{bzw.} \quad \mathbf{V} = \boldsymbol{\eta}^{-1} \underline{V}^T, \quad (\text{X.37})$$

wobei $\boldsymbol{\eta}^T$ in der ersten Gleichung durch $\boldsymbol{\eta}$ ersetzt werden kann, weil der metrische Tensor — und damit seine Matrixdarstellung — symmetrisch ist. Mit dieser Änderung ist die zweite Gleichung die transponierte der ersten.

* Kontravariante und kovariante Vierervektoren sind Elemente unterschiedlicher Vektorräume, auch wenn sie die gleiche physikalische Dimension haben — wie z.B. kontravariante und kovariante Vierergeschwindigkeiten. Der metrische Tensor $\boldsymbol{\eta}$ und sein Inverse $\boldsymbol{\eta}^{-1}$ bilden zwar eine Bijektion zwischen diesen Vektorräumen, so dass man kurz nur von „Vierervektoren“ sprechen kann, ohne kontra- oder kovariant zu präzisieren. Man darf aber kontra- und kovariante Vektoren nicht miteinander addieren oder gleich setzen: Ausdrücke wie $a^{\mu} + b_{\mu}$ oder $a^{\mu} = b_{\mu}$ sind sinnlos — wenn es sich dabei nicht um numerische Gleichungen handelt, die nur in einem Bezugssystem gelten.

X.3.2c Viererprodukt

Seien V^μ bzw. W_μ die Komponenten eines kontravarianten bzw. kovarianten Vierervektors. Die Transformationsgesetze (X.25a) und (X.32) zeigen, dass die Kombination $W_\mu V^\mu$, mit Summe über $\mu = 0, 1, 2, 3$, ein Lorentz-Skalar ist. Dieses Skalar wird *Viererprodukt* der beiden Vierervektoren genannt, oder auch *Lorentz-Skalarprodukt* — etwa uneigentlich, da es sich eher um ein Pseudo-Skalarprodukt handelt.

Da $W_\mu V^\mu = W^\mu \eta_{\mu\nu} V^\nu = W^\nu V_\nu = W^\mu V_\mu$ ist, spielt es hier keine Rolle, welcher Index oben und welcher unten ist, so lange es einen oben und einen unten gibt. Komponentenweise gilt

$$W_\mu V^\mu = -W^0 V^0 + W^1 V^1 + W^2 V^2 + W^3 V^3 = W^\mu V_\mu. \quad (\text{X.38a})$$

Unter Verwendung der Matrixdarstellungen \underline{W} bzw. \underline{V} für den kovarianten bzw. kontravarianten Vierervektor, lässt sich das Viererprodukt noch als $W_\mu V^\mu = \underline{W}\underline{V}$ schreiben. Ein Nachteil dieser Schreibweise ist, dass die Multiplikation zweier Matrizen nicht-kommutativ ist: $\underline{W}\underline{V}$ ist eine Zahl, $\underline{V}\underline{W}$ eine 4×4 -Matrix. Das heißt, der Gleichung $W_\mu V^\mu = W^\mu V_\mu$ entspricht $\underline{W}\underline{V} = \underline{V}\underline{W}$, wobei die Symmetrie bezüglich des Austauschs der Vierervektoren verloren wird. Somit führt man eine weitere Notation, ähnlich dem Skalarprodukt von zwei Dreiervektoren im euklidischen Raum:

$$W_\mu V^\mu = \underline{W}\underline{V} \equiv \mathbf{W} \cdot \mathbf{V}. \quad (\text{X.38b})$$

Dann gilt problemlos $\mathbf{W} \cdot \mathbf{V} = \mathbf{V} \cdot \mathbf{W}$. Man sollte aber dabei nicht vergessen, dass es sich trotz der Notation um das *Lorentz-Skalarprodukt* handelt, nicht um das euklidische, d.h. dass es ein Minus-Zeichen vor dem Produkt der zeitlichen Komponenten gibt.

Bemerkung: Im Gegensatz zum Viererprodukt $W_\mu V^\mu$ sind die Kombinationen $W^\mu V^\mu$ und $W_\mu V_\mu$ keine Skalare unter Lorentz-Transformation, d.h. sie nehmen unterschiedliche Werte in unterschiedlichen Inertialsystemen an.

Ein erstes Beispiel von Viererprodukt ist die Kombination $\partial_\mu j^\mu(\mathbf{x})$, mit ∂_μ den Komponenten des Vierergradienten und $j^\mu(\mathbf{x})$ die Koordinaten eines Feldes, das in jedem Punkt \mathbf{x} eines Gebiets der Raumzeit \mathcal{M}_4 definiert ist:

$$\partial_\mu j^\mu(\mathbf{x}) = \frac{\partial j^\mu(\mathbf{x})}{\partial x^\mu} = \frac{\partial j^0(\mathbf{x})}{\partial x^0} + \frac{\partial j^1(\mathbf{x})}{\partial x^1} + \frac{\partial j^2(\mathbf{x})}{\partial x^2} + \frac{\partial j^3(\mathbf{x})}{\partial x^3} = \frac{1}{c} \frac{\partial j^0(t, \vec{r})}{\partial t} + \vec{\nabla} \cdot \vec{j}(t, \vec{r}), \quad (\text{X.39})$$

wobei der Dreiervektor \vec{j} die räumlichen Komponenten j^i des Vierervektors \mathbf{j} zusammenfasst, während $\vec{\nabla} \cdot$ die (Dreier-)Divergenz dieses Vektorfeldes darstellt. In Analogie mit dem dreidimensionalen Fall heißt $\partial_\mu j^\mu(\mathbf{x}) \equiv \partial \cdot \mathbf{j}(\mathbf{x})$ *Viererdivergenz* des Vierervektorfeldes \mathbf{j} .

Lorentz-Quadrat

Ein weiteres, wichtiges Beispiel von Viererprodukt ist der skalare *Lorentz-Quadrat* eines Vierervektors:

$$\mathbf{V}^2 \equiv V_\mu V^\mu = -(V^0)^2 + (V^1)^2 + (V^2)^2 + (V^3)^2, \quad (\text{X.40})$$

der denselben Wert in allen Inertialsystemen annimmt. Beispielsweise beträgt der Lorentz-Quadrat einer Vierergeschwindigkeit (X.26) immer $\mathbf{u}^2 = u_\mu u^\mu = -c^2$. Das Linienelement ds^2 ist ebenfalls ein Lorentz-Quadrat.

Definition: Je nach dem Vorzeichen des Lorentz-Quadrats unterscheidet man zwischen drei Arten von Vierervektoren:

- *zeitartige* Vierervektoren, mit negativem Lorentz-Quadrat $\mathbf{V}^2 < 0$;
- *lichtartige* Vierervektoren, auch *Nullvektoren* genannt, mit Lorentz-Quadrat $\mathbf{V}^2 = 0$;
- *raumartige* Vierervektoren, mit positivem Lorentz-Quadrat $\mathbf{V}^2 > 0$.

Zum Beispiel ist die Vierergeschwindigkeit eines Massenpunktes ein zeitartiger Vierervektor.

Der Unterschied zwischen den drei Arten ist insbesondere wichtig wenn der Vierervektor dem Abstandsvektor zwischen zwei Ereignissen P_1 und P_2 entspricht. Im Fall einer zeitartigen Separation kann man immer ein Inertialsystem \mathcal{B} derart finden, dass die Weltlinie seines Ursprungspunkts durch P_1 und P_2 geht. Somit ist das eine Ereignis in der Zukunft des anderen, z.B. $t_2 > t_1$: die zwei Ereignisse sind miteinander kausal verknüpft, d.h. was im einen passiert kann das andere beeinflussen.

Falls der Abstandsvektor raumartig ist, kann ein solches Inertialsystem nicht gefunden werden. Angenommen, dass das Vakuumlichtgeschwindigkeit die maximale Ausbreitungsgeschwindigkeit von Signalen ist, kann kein aus P_1 ausgesandte Signal P_2 erreichen, und umgekehrt. Eigentlich gibt es in diesem Fall Inertialsystemen, in denen $t_1 < t_2$, während in anderen Inertialsystemen $t_2 < t_1$ gilt: keines der Ereignisse liegt auf absoluter Weise in der Zukunft des anderen. Dementsprechend darf nicht mehr von der Gleichzeitigkeit zweier Ereignisse geredet werden, denn $t_1 = t_2$ in einem ersten Inertialsystem wird in einem anderen Bezugssystem zu $t'_1 < t'_2$ oder $t''_1 > t''_2$.

Schließlich kann im Fall eines lichtartigen Abstandsvektors zwischen Ereignissen Licht von dem einen zum anderen propagieren, so dass die Ereignisse noch im kausalen Zusammenhang stehen.

Vom Standpunkt eines Ereignisses P_0 kann somit die Raumzeit in unterschiedlichen Bereichen zerlegt werden (vgl. Abb. X.2) Die Ereignisse, die in lichtartigem Abstand von P_0 sind, bilden die *Lichtkegel*.⁽⁸⁴⁾ Innerhalb der letzteren, d.h. in zeitartigem Abstand von P_0 , befinden sich die Ereignisse, die im kausalen Zusammenhang mit P_0 sind, und zwar entweder in seiner Zukunft — wie P_1 in Abb. X.2 — oder in der Vergangenheit, wie P_2 . Dagegen sind die Ereignisse außerhalb der Lichtkegel — d.h. in raumartigem Abstand von P_0 , wie P_3 — kausal entkoppelt von P_0 .

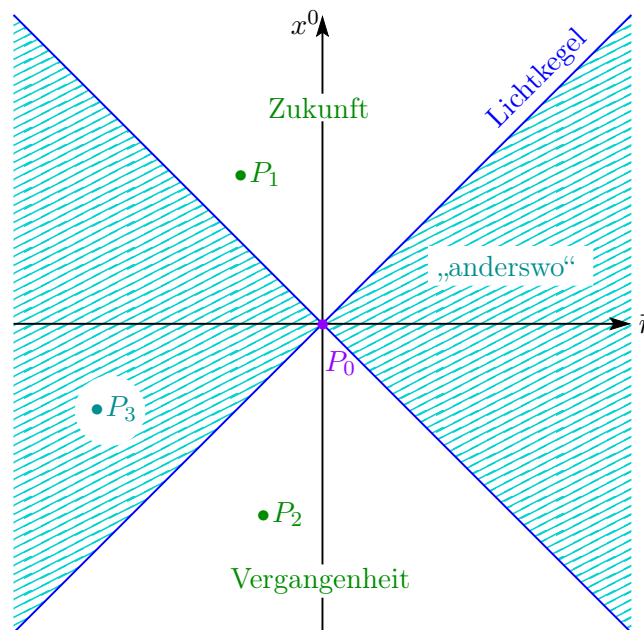


Abbildung X.2 – Kausalitätsstruktur der Raumzeit: die horizontale Achse steht für die drei räumlichen Dimensionen, die vertikale Achse für die Zeit im Ruhesystem von P_0 .

Bemerkung: Entsprechend dieser Kausalitätsstruktur hängt die Physik in P_0 nur von den Ereignissen innerhalb der „vergangenen Lichtkegel“ ab: instantane Fernwirkung wird somit ausgeschlossen.

⁽⁸⁴⁾Um die Kegel zu erkennen, muss sich die Leserin erstens eine Raumzeit mit einer zeitlichen Richtung — die der Achse der Kegel entspricht — und nur zwei räumlichen Richtungen vorstellen; dann kann sie die dritte Raumdimension hinzufügen.

X.3.3 Vierertensoren

Auf dem euklidischen Raum \mathcal{E}_3 der nicht-relativistischen Mechanik gibt es „Dreier-“Tensoren höherer Stufe, die sich unter Drehungen alle gleich transformieren. Ähnlich existieren auf der Minkowski-Raumzeit \mathcal{M}_4 *Vierertensoren* bzw. *Lorentz-Tensoren* mit bestimmten Transformationsgesetzen unter Lorentz-Transformationen.

Diese Vierertensoren können vom Typ $\binom{m}{n}$, oder „ m -fach kontravariant und n -fach kovariant“, sein, d.h. ihre Komponenten besitzen $m \geq 0$ kontravariante und $n \geq 0$ kovariante Indizes, wobei jeder Index sich wie ein entsprechender Vierervektor transformiert.⁽⁸⁵⁾ Beispielsweise transformiert sich der Vierertensor dritter Stufe mit Komponenten $T^{\mu\nu\rho}$ wie $V^\mu V^\nu V^\rho$.

Insbesondere lautet das Transformationsgesetz für einen Lorentz-Tensor vom Typ $\binom{0}{2}$, unter Verwendung der Gl. (X.32),

$$T_{\mu\nu} \rightarrow T_{\mu'\nu'} = \Lambda^\mu_{\mu'} \Lambda^\nu_{\nu'} T_{\mu\nu}. \quad (\text{X.41a})$$

In Matrixform kann ein solcher Vierertensor zweiter Stufe durch eine quadratische 4×4 -Matrix \mathbb{T} mit Elementen $T_{\mu\nu}$ dargestellt werden; dann lautet die Transformation

$$\mathbb{T} \rightarrow \mathbb{T}' = \Lambda^\top \mathbb{T} \Lambda, \quad (\text{X.41b})$$

wie aus der Gleichung $\Lambda^\mu_{\mu'} \Lambda^\nu_{\nu'} T_{\mu\nu} = \Lambda^\mu_{\mu'} T_{\mu\nu} \Lambda^\nu_{\nu'} = (\Lambda^\top)_{\mu'}^\mu T_{\mu\nu} \Lambda^\nu_{\nu'}$ sofort folgt.

X.3.3a Metrischer Tensor

Der metrische Tensor $\boldsymbol{\eta}$ wurde schon in Gl. (X.16) über die Angabe seiner Komponenten $\eta_{\mu\nu}$ in einem bestimmten System von Minkowski-Koordinaten definiert. Jede dieser Komponenten nimmt in allen solchen Systemen denselben Wert an, d.h. der metrische Tensor ist *invariant* unter Lorentz-Transformationen.

Unter Verwendung des Transformationsgesetzes (X.41a) transformieren sich die Komponenten von $\boldsymbol{\eta}$ gemäß

$$\eta_{\mu\nu} \rightarrow \eta_{\mu'\nu'} = \Lambda^\mu_{\mu'} \Lambda^\nu_{\nu'} \eta_{\mu\nu}, \quad (\text{X.42a})$$

entsprechend in Matrixdarstellung

$$\boldsymbol{\eta} \rightarrow \boldsymbol{\eta}' = \Lambda^\top \boldsymbol{\eta} \Lambda. \quad (\text{X.42b})$$

Fördert man die Invarianz des metrischen Tensors unter Lorentz-Transformationen, so sollen die Matrizen $\boldsymbol{\eta}$ und $\boldsymbol{\eta}'$ gleich sein, woraus sich Bedingung (X.20) ergibt.

X.3.3b Inverser metrischer Tensor

Dem metrischen Tensor wird der *inverse metrische Tensor* $\boldsymbol{\eta}^{-1}$ zugeordnet, dessen Komponenten $\eta^{\mu\nu}$ bzw. Matrixdarstellung $\boldsymbol{\eta}^{-1}$ derart sind, dass

$$\eta_{\mu\rho} \eta^{\rho\nu} = \eta^{\nu\rho} \eta_{\rho\mu} = \delta_\mu^\nu \quad \forall \mu, \nu = 0, 1, 2, 3 \quad \text{bzw.} \quad \boldsymbol{\eta} \boldsymbol{\eta}^{-1} = \boldsymbol{\eta}^{-1} \boldsymbol{\eta} = \mathbb{1}_4 \quad (\text{X.43})$$

gelten, mit δ_μ^ν dem üblichen Kronecker-Symbol und $\mathbb{1}_4$ der 4×4 -Einheitsmatrix. Dabei ist eigentlich $\boldsymbol{\eta}^{-1} = \boldsymbol{\eta}$, d.h. numerisch gilt $\eta^{\mu\nu} = \eta_{\mu\nu}$ für alle möglichen Werte von μ und ν .

Wie in § X.3.2 b schon gesehen wurde lassen sich mithilfe des metrischen bzw. des inversen metrischen Tensors mit Komponenten $\eta_{\mu\nu}$ bzw. $\eta^{\mu\nu}$ kontra- bzw. kovariante Lorentz-Indizes herauf- bzw. herabziehen, ähnlich den Gleichungen (X.33) und (X.34). Zum Beispiel kann man aus einem Tensor vom Typ $\binom{0}{2}$ mit Komponenten $T_{\mu\nu}$ verwandte Tensoren von den Typen $\binom{1}{1}$ und $\binom{2}{0}$ erhalten:

$$T_\mu^\sigma \equiv \eta^{\sigma\nu} T_{\mu\nu}, \quad T^\rho_\nu \equiv \eta^{\rho\mu} T_{\mu\nu} \quad \text{und} \quad T^{\rho\sigma} \equiv \eta^{\rho\mu} \eta^{\sigma\nu} T_{\mu\nu} = \eta^{\rho\mu} T_\mu^\sigma = \eta^{\sigma\nu} T^\rho_\nu. \quad (\text{X.44})$$

Wenn $T_{\mu\nu}$ nicht symmetrisch ist, sind die Tensoren mit Komponenten T_μ^σ und T^ρ_ν a priori unterschiedlich.⁽⁸⁶⁾

⁽⁸⁵⁾Vgl. Anhang A.

⁽⁸⁶⁾Bezeichnet \mathbb{T} die Matrixdarstellung des Tensors mit Komponenten $T_{\mu\nu}$, so sind die drei anderen Tensoren jeweils durch $\mathbb{T}\boldsymbol{\eta}$, $\boldsymbol{\eta}\mathbb{T}$ und $\boldsymbol{\eta}\mathbb{T}\boldsymbol{\eta}$ dargestellt, wie die Leserin einfach nachprüfen kann.

Angewandt auf $\eta^{\mu\nu}$ selber lautet die zweite der Gleichungen (X.44)

$$\eta^\rho{}_\nu \equiv \eta^{\rho\mu} \eta_{\mu\nu}. \quad (\text{X.45})$$

Der Vergleich mit Gl. (X.43) gibt dann $\eta^\rho{}_\nu = \delta^\rho_\nu$.

X.3.3c Levi-Civita-Tensor

Ein anderer invarianter Vierertensor — genauer *Pseudotensor*, da sein Vorzeichen sich unter uneigentlichen Lorentz-Transformationen ändert — ist der vollständig antisymmetrische *Levi-Civita-Tensor* 4. Stufe⁽⁸⁷⁾

$$\epsilon^{\mu\nu\rho\sigma} = \begin{cases} +1 & \text{falls } (\mu, \nu, \rho, \sigma) \text{ eine gerade Permutation von } (0,1,2,3) \text{ ist} \\ -1 & \text{falls } (\mu, \nu, \rho, \sigma) \text{ eine ungerade Permutation von } (0,1,2,3) \text{ ist} \\ 0 & \text{sonst.} \end{cases} \quad (\text{X.46})$$

Hier sollte beachtet werden, dass $\epsilon_{0123} = -\epsilon^{0123}$, während für den dreidimensionalen Levi-Civita Tensor $\epsilon_{123} = \epsilon^{123}$ gilt.

Unter einer Lorentz-Transformation Λ transformiert sich dieser Tensor gemäß

$$\epsilon^{\mu\nu\rho\sigma} \rightarrow \epsilon^{\mu'\nu'\rho'\sigma'} \equiv \Lambda^{\mu'}{}_\mu \Lambda^{\nu'}{}_\nu \Lambda^{\rho'}{}_\rho \Lambda^{\sigma'}{}_\sigma \epsilon^{\mu\nu\rho\sigma} = (\det \Lambda) \epsilon^{\mu\nu\rho\sigma}, \quad (\text{X.47})$$

wobei die letzte Gleichung aus der Definition der Determinante folgt.

X.3.3d Kontraktion zweier Tensoren

Das Herauf- oder Herabziehen von Indizes mithilfe des metrischen Tensors oder seines Inversen sind Beispiele von Tensorverjüngungen (vgl. § A.1.3 b). Allgemeiner können zwei Tensoren \mathbf{T} und \mathbf{T}' kontrahiert werden, vorausgesetzt der eine (mindestens) einen kovarianten und der andere einen kontravarianten Index hat. Dann wird über solche Indizes kontrahiert, wie z.B. $T^{\mu\nu} T'_{\mu\rho}$ (einfache Kontraktion) oder $\epsilon_{\mu\nu\rho\sigma} T^{\mu\nu}$ (doppelte Kontraktion).

X.3.4 Kovariante Formulierung eines physikalischen Gesetzes

Laut dem ersten Einstein'schen Postulat (X.1) sollen die Naturgesetze mathematisch so formuliert werden, dass die entsprechenden Gleichungen in allen Inertialsystemen die gleiche Form annehmen. Theorien, die diesem Prinzip genügen, werden als Lorentz- oder relativistisch *kovariant* bezeichnet.

Infolgedessen können solche Gleichungen nur Identitäten zwischen Lorentz-Tensoren gleicher Stufe sein, nachdem alle möglichen Kontraktionen von Indizes berücksichtigt wurden. Solche Gleichungen können entweder „geometrisch“, in Matrixdarstellung (für Vierertensoren der Stufe 1 oder 2) oder komponentenweise geschrieben werden.

Beispiele kovarianter Gleichungen sind somit $\mathbf{V} = \mathbf{W}$ oder äquivalent $V^\mu = W^\mu$; oder $\mathbf{T}\mathbf{V} = \mathbf{W}$ bzw. $T^\mu{}_\nu V^\nu = W^\mu$. Unter einer Lorentz-Transformation transformieren sie sich in $\mathbf{V}' = \mathbf{W}'$ bzw. $V^{\mu'} = W^{\mu'}$ oder $\mathbf{T}'\mathbf{V}' = \mathbf{W}'$ bzw. $T^{\mu'}{}_{\nu'} V^{\nu'} = W^{\mu'}$, d.h. sie nehmen die gleiche Form an.

Dagegen sind Identitäten wie $V^\mu = T^\mu{}_\nu$ oder $V^\mu = W_\mu$ keine gültige relativistisch kovariante Gleichungen, sondern können einen (Tipp-?)Fehler signalisieren...

Bemerkung: Manchmal wird statt Lorentz-kovariant die Redensart „Lorentz-invariant“ verwendet. Streng genommen bedeutet aber die Letztere, dass die Gleichungen unverändert unter den Lorentz-Transformationen bleiben: dies stellt eine stärkere Bedingung dar, als die Invarianz der Form der Gleichungen, die nur im Fall einer Gleichung zwischen zwei skalaren Größen erfüllt ist, wie z.B. $u^2 = -c^2$ (für den Lorentz-Quadrat einer Vierergeschwindigkeit).

⁽⁸⁷⁾ Einige Autoren benutzen die Konvention $\epsilon_{0123} = +1$, entsprechend $\epsilon^{0123} = -1$...

Literatur zum Kapitel X

- Feynman, *Vorlesungen über Physik. Band 1* [24] = *Lectures on Physics. Vol. I* [25], Kap. 15–17.
- Fließbach, *Mechanik* [2], Kap. IX.
- Greiner, *Klassische Mechanik I* [6] Kap. III.
- Landau & Lifschitz, *Band II: Klassische Feldtheorie* [14] = *The classical theory of fields* [28], Kap. I.
- Nolting, *Spezielle Relativitätstheorie, Thermodynamik* [32], 1. Teil.
- Scheck, *Mechanik* [18] Kap. 4.

KAPITEL XI

Relativistische Mechanik

Der Formalismus des vorigen Kapitels wird nun angewandt, um die charakteristischen Größen und Funktionen zur Beschreibung der Bewegung eines freien relativistischen Massenpunktes auszudrücken (Abschn. XI.1). Dann wird die relativistische Verallgemeinerung des zweiten Newton'schen Gesetzes vorgestellt (Abschn. XI.2).

XI.1 Bewegung eines freien relativistischen Teilchens

XI.1.1 Lagrange-Funktion und Wirkung eines freien Teilchens

Die Wirkung S eines physikalischen Systems, insbesondere eines freien Teilchens, ist eine reelle Zahl. Sie hängt von Größen — Positionen, Geschwindigkeiten — ab, die vom Bezugssystem abhängen. Die einfachste Möglichkeit um sicherzustellen, dass die Extrema der Wirkung immer für die gleichen physikalischen Trajektorien erreicht werden, entsprechend den physikalisch realisierten Bewegungen, ist, dass die Wirkung ein Lorentz-Skalar ist, d.h. den gleichen Zahlenwert in allen Inertialsystemen annimmt.

Betrachte nun einen einzelnen, wechselwirkungsfreien Massenpunkt mit Masse m . Wegen der Homogenität der Raumzeit — d.h. die Invarianz der Physik unter zeitlichen und räumlichen Translationen — kann seine Lagrange-Funktion $L(t, \vec{x}, \dot{\vec{x}})$ weder von der Zeit noch von seiner Position abhängen; d.h. sie ist nur Funktion von seiner Geschwindigkeit $\vec{v} = \dot{\vec{x}}$. Die letztere ist zwar nicht Lorentz-invariant, jedoch die Lagrange-Funktion L selbst ist auch nicht Lorentz-invariant: nur das Produkt $L dt$, das zwischen zwei Zeitpunkten integriert wird, soll Lorentz-invariant sein, damit S ein Lorentz-Skalar ist.

Aus dem vorigen Kapitel kennen wir eine „einfache“ Größe, proportional zum Zeitintervallelement dt und dazu Funktion der Geschwindigkeit, die Lorentz-invariant ist, und zwar das infinitesimal Eigenzeitintervall $d\tau = \sqrt{1 - \vec{v}^2/c^2} dt$, vgl. Gl. (X.14). Somit ist ein möglicher Ansatz

$$S = \alpha \int_A^B d\tau \quad (\text{XI.1})$$

mit α einer reellen Zahl und A, B zwei Ereignissen der Raumzeit. Diese Wirkung lässt sich noch als

$$S = \alpha \int_{t_A}^{t_B} \sqrt{1 - \frac{\vec{v}^2}{c^2}} dt$$

umschreiben; daher entspricht der Ansatz einer Lagrange-Funktion

$$L = \alpha \sqrt{1 - \frac{\vec{v}^2}{c^2}}. \quad (\text{XI.2})$$

Wird eine Taylor-Entwicklung dieser Abhängigkeit für den nicht-relativistischen Limes $|\vec{v}|/c \ll 1$ durchgeführt, so ergibt sich

$$L = \alpha \left[1 - \frac{\vec{v}^2}{2c^2} + \mathcal{O}\left(\frac{\vec{v}^4}{c^4}\right) \right].$$

Nach Ausmultiplizieren der Klammer kommt zuerst eine Konstante (α), die keine Rolle bei den Bewegungsgleichungen spielt. Der nächste Term ist $-\alpha\vec{v}^2/2c^2$: damit er mit der Lagrange-Funktion eines nicht-relativistischen freien Teilchens $\frac{1}{2}m\vec{v}^2$ übereinstimmt, soll α gleich $-mc^2$ sein. Somit lautet die Lagrange-Funktion bzw. die Wirkung eines freien relativistischen Massenpunktes

$$L = -mc^2 \sqrt{1 - \frac{\vec{v}^2}{c^2}} \quad (\text{XI.3a})$$

bzw.

$$S = -mc^2 \int_A^B d\tau. \quad (\text{XI.3b})$$

Erfahrung zeigt, dass der Ansatz (XI.3a) die richtige Physik wiedergibt, insbesondere den richtigen Impuls und die richtige Energie, wie sie im nächsten Paragraphen eingeführt werden.

XI.1.2 Impuls und Energie eines freien Teilchens

XI.1.2a Impuls

Gemäß der Definition (III.17) ist die i -te Komponente des konjugierten Impulses des Massenpunktes mit Lagrange-Funktion (XI.3a) durch

$$p_i \equiv \frac{\partial L}{\partial \dot{x}^i}$$

gegeben, wobei hier die Komponente mit einem kovarianten Index geschrieben wurde, im Einklang mit der Division durch eine kontravariante Komponente im rechten Glied. Eine einfache Berechnung gibt dann

$$\vec{p} = \frac{m\vec{v}}{\sqrt{1 - \frac{\vec{v}^2}{c^2}}} = \gamma m\vec{v} \quad (\text{XI.4})$$

mit γ dem Lorentz-Faktor entsprechend der Geschwindigkeit \vec{v} .

XI.1.2b Energie

Benutzt man nun die Definition (III.35a), so lässt sich die Energie des freien relativistischen Massenpunktes als

$$E = \sum_{i=1}^3 p_i \dot{x}^i - L$$

bestimmen, d.h.

$$E = \frac{mc^2}{\sqrt{1 - \frac{\vec{v}^2}{c^2}}} = \gamma mc^2. \quad (\text{XI.5})$$

In Fall eines ruhenden Massenpunktes ergibt sich $E_0 = mc^2$, entsprechend der *Massenenergie*. Wiederum ist die Differenz $(\gamma - 1)mc^2$ die kinetische Energie eines bewegten Massenpunktes.

XI.1.2c Viererimpuls

Definiert man $p^0 \equiv E/c$, so bilden p^0, p^1, p^2, p^3 ein 4-Tupel von Zahlen, die laut den Gl. (XI.4) und (XI.5) gleich m mal den Komponenten u^0, u^1, u^2, u^3 der Vierergeschwindigkeit (X.27) sind. Da die Masse m ein Lorentz-Skalar ist, sind die p^μ die Komponenten eines Vierervektors \mathbf{p} :

$$p^\mu = mu^\mu \quad (\text{XI.6a})$$

bzw.

$$\mathbf{p} = m\mathbf{u} \quad (\text{XI.6b})$$

mit

$$\mathbf{p} = \begin{pmatrix} E/c \\ \vec{p} \end{pmatrix}. \quad (\text{XI.7})$$

Dieser Vierervektor heißt *Viererimpuls*.

Vergleicht man die Gl. (XI.4) und (XI.5), so findet man, dass sich die Geschwindigkeit als

$$\vec{v} = \frac{\vec{p}c^2}{E} \quad (\text{XI.8})$$

schreiben lässt. Somit ist der nicht-relativistische Limes $|\vec{v}| \ll c$ äquivalent zu $|\vec{p}|c \ll E$.

Unter Verwendung der Definition (X.40) und der Gl. (XI.4) und (XI.5) beträgt der Lorentz-Quadrat \mathbf{p}^2 des Viererimpulses (XI.7)

$$\mathbf{p}^2 \equiv -\frac{E^2}{c^2} + \vec{p}^2 = -m^2c^2. \quad (\text{XI.9})$$

Dies folgt auch sofort aus der Beziehung (XI.6b), die zu $\mathbf{p}^2 = m^2\mathbf{u}^2$ führt, oder aus einer Berechnung im Ruhesystem des Massenpunktes.

Dieses Ergebnis wird auch oft in der Form

$$E^2 = \vec{p}^2c^2 + m^2c^4 \quad (\text{XI.10})$$

geschrieben. Ist der erste Term auf der rechten Seite viel kleiner als der zweite, d.h. $|\vec{p}| \ll mc$ — was wiederum äquivalent zu $|\vec{p}| \ll E/c$ ist, entsprechend dem nicht-relativistischen Limes —, folgt aus einer Taylor-Entwicklung

$$E = \sqrt{m^2c^4 + \vec{p}^2c^2} = mc^2 \sqrt{1 + \frac{\vec{p}^2}{m^2c^2}} = mc^2 \left[1 + \frac{1}{2} \frac{\vec{p}^2}{m^2c^2} - \left(\frac{1}{8} \frac{\vec{p}^2}{m^2c^2} \right)^2 + \dots \right]$$

d.h.

$$E = mc^2 + \frac{\vec{p}^2}{2m} - \frac{(\vec{p}^2)^2}{8m^3c^2} + \dots \quad (\text{XI.11})$$

Dabei erkennt man als ersten, führenden Term die Massenenergie, dann als nächsten Term die nicht-relativistische kinetische Energie, während der dritte Term (und die nächsten) eine relativistische Korrektur darstellt. Insbesondere ist im nicht-relativistischen Fall die kinetische Energie viel kleiner als die Massenenergie — die im nicht-relativistischen Rahmen aber nie ins Betracht gezogen wird, weil sie nur eine additive Konstante darstellt!

Schließlich folgt aus der Invarianz der Wirkung (XI.3b) unter zeitlichen und räumlichen Translationen, dass die Energie und der Impuls Konstanten der Bewegung sind, vgl. § III.3.1 a–III.3.1 b. Folglich ist der Viererimpuls erhalten, wie man auch direkt aus der Invarianz der Wirkung (XI.3b) unter den Raumzeittranslationen $x^\mu \rightarrow x^\mu + a^\mu$ herleiten kann.

Mathematisch lässt sich diese Viererimpulserhaltung als

$$\boxed{\frac{d\mathbf{p}}{d\tau} = 0} \quad \text{bzw.} \quad \boxed{\frac{dp^\mu}{d\tau} = 0} \quad (\text{XI.12})$$

ausdrücken: dabei wird die Ableitung nach der Eigenzeit betrachtet, um auf der linken Seite einen Lorentz-Vektor zu erhalten. Somit ergibt sich eine Gleichung, die in jedem Inertialsystem die gleiche Form annimmt, wie gemäß dem ersten Einstein'schen Postulat (X.1) der Fall sein soll.

XI.2 Kovariante Formulierung des Grundgesetzes der Mechanik

Wir haben gerade gesehen, dass der Viererimpuls eines freien Massenpunktes erhalten ist, was sich „relativistisch kovariant“ als Gl. (XI.12) schreiben lässt.

Unterliegt der Massenpunkt Kräften, so sollte sich dieses Ergebnis ändern, denn es ändert sich im nicht-relativistischen Fall, der als Grenzwert in den relativistischen Gleichungen erhalten ist. Um die Newton'sche Mechanik zu verallgemeinern, wird ein Vierervektor F eingeführt, die *Viererkraft*, mit Komponenten F^μ . Im Einklang mit dem nicht-relativistischen Fall sollte eine solche Viererkraft nur von der Raumzeit-Position x und der Vierergeschwindigkeit u des Massenpunktes abhängen:

$$F(x, u) \quad \text{bzw.} \quad F^\mu(x^\nu, u^\nu). \quad (\text{XI.13})$$

Insbesondere wird in Kap. XII die Form der relativistisch kovarianten Formulierung der Lorentz-Kraft angegeben.

Gegeben eine solche Viererkraft — oder die Resultierende solcher Viererkräfte —, die natürliche relativistisch kovariante Verallgemeinerung des zweiten Newton'schen Gesetzes, die auch in Abwesenheit von Kraft zur Viererimpulserhaltung (XI.12) führen wird, ist

$$\frac{dp}{d\tau} = F \quad (\text{XI.14a})$$

oder äquivalent, komponentenweise

$$\frac{dp^\mu}{d\tau} = F^\mu. \quad (\text{XI.14b})$$

Unter Einführung der *Viererbeschleunigung* des Massenpunktes

$$a \equiv \frac{du}{d\tau} \quad \text{bzw.} \quad a^\mu \equiv \frac{du^\mu}{d\tau} \quad (\text{XI.15})$$

lässt sich Gl. (XI.14a) bzw. (XI.14b) als

$$ma = m \frac{du}{d\tau} = F \quad \text{bzw.} \quad ma^\mu = m \frac{du^\mu}{d\tau} = F^\mu. \quad (\text{XI.16})$$

Bemerkung: Damit die Beziehungen (XI.14), (XI.16) erfüllt werden können, dürfen die Komponenten einer Viererkraft nicht beliebig sein. In der Tat folgt aus der Konstanz des Lorentz-Quadrats u^2 der Vierergeschwindigkeit, dass diese senkrecht auf die Viererbeschleunigung ist: $u \cdot a = u_\mu a^\mu = 0$. Somit soll $u \cdot F = u_\mu F^\mu = 0$ gelten, d.h. nur drei Komponenten der Viererkraft sind unabhängig voneinander — z.B. die drei räumlichen Komponenten, die im nicht-relativistischen Limes die Dreierkraft wiedergeben müssen.

Literatur zum Kapitel XI

- Fließbach, *Mechanik* [2], Kap. IX.
- Greiner, *Klassische Mechanik I* [6] Kap. III.
- Landau & Lifschitz, *Band II: Klassische Feldtheorie* [14] = *The classical theory of fields* [28], Kap. II.
- Nolting, *Spezielle Relativitätstheorie, Thermodynamik* [32], 1. Teil.
- Scheck, *Mechanik* [18] Kap. 4.

KAPITEL XII

Relativistisch kovariante Formulierung der Elektrodynamik

Die spezielle Relativitätstheorie wurde entwickelt, um die Invarianz der Geschwindigkeit im Vakuum von Licht, d.h. allgemeiner von elektromagnetischen Wellen, in Transformationen zwischen Inertialsystemen zu berücksichtigen. Dementsprechend ist Elektrodynamik automatisch eine „relativistische“ Theorie, was aber an den Maxwell-Gleichungen nicht sofort erkennbar ist.

In diesem Kapitel werden die Gesetze und einige Resultate der Elektrodynamik in relativistisch kovarianter Notation formuliert. Somit werden erstens Lorentz-Vierervektoren und -tensoren eingeführt, welche die unterschiedlichen elektromagnetischen Größen beschreiben (Abschn. XII.1). Diese werden dann in Abschn. XII.2 benutzt, um die Grundgleichungen der Elektrodynamik auszudrücken. Schließlich werden weitere Ergebnisse des Kapitels IX in relativistisch kovarianter Schreibweise angegeben (Abschn. XII.3).

Hiernach bezeichnet x einen Punkt der Raumzeit mit kontravarianten Koordinaten x^μ , und der metrische Tensor η hat die Signatur $-, +, +, +$ (vgl. § X.3.3 a). Die Einsteinsche Summenkonvention über doppelt auftretende Lorentz-Indizes wird durchaus benutzt.

XII.1 Lorentz-kovariante elektromagnetische Größen

Zur Formulierung der Elektrodynamik in relativistisch kovarianter Notation sollen zuerst passende Lorentz-kovariante Größen — Viererskalare, -vektoren oder -tensoren — definiert werden, welche das elektromagnetische Feld (§ XII.1.1) mit den zugehörigen Potentialen (§ XII.1.2), dessen Quellen (§ XII.1.3), sowie dessen Energie und Impuls (§ XII.1.4) beschreiben.

XII.1.1 Elektromagnetischer Feldstärketensor

In relativistisch kovarianter Schreibweise gewinnt die Bezeichnung „elektromagnetisches Feld“ ihre volle Bedeutung, indem die elektrische und magnetische Dreivektorfelder miteinander in ein einzelnes mathematisches Objekt kombiniert werden.

XII.1.1 a Definition

Um eine relativistisch kovariante Größe zu erhalten, werden das elektrische und das magnetische Feld \vec{E} und \vec{B} zum *elektromagnetischen Feldstärketensor* zweiter Stufe \mathbf{F} kombiniert. Letzterer kann durch die Angabe seiner kontravarianten Komponenten

$$\begin{aligned} F^{00}(x) &= F^{ii}(x) = 0, \\ F^{0i}(x) &= -F^{i0}(x) \equiv \frac{E^i(x)}{c}, \\ F^{ij}(x) &= -F^{ji}(x) \equiv \sum_{k=1}^3 \epsilon^{ijk} B^k(x) \end{aligned} \tag{XII.1a}$$

bezüglich eines gegebenen Inertialsystems \mathcal{B} definiert werden. Dabei sind $i, j \in \{1, 2, 3\}$, während

ϵ^{ijk} das dreidimensionale völlig antisymmetrische Levi-Civita-Symbol bezeichnet. Die Komponenten E^i, B^j der Felder \vec{E} und \vec{B} sind solche, die ein bezüglich \mathcal{B} ruhender Beobachter messen würde.

Die Matrixdarstellung des elektromagnetischen Feldstärketensors \mathbf{F} lautet

$$F^{\mu\nu}(x) = \begin{pmatrix} 0 & \frac{E^1(x)}{c} & \frac{E^2(x)}{c} & \frac{E^3(x)}{c} \\ -\frac{E^1(x)}{c} & 0 & B^3(x) & -B^2(x) \\ -\frac{E^2(x)}{c} & -B^3(x) & 0 & B^1(x) \\ -\frac{E^3(x)}{c} & B^2(x) & -B^1(x) & 0 \end{pmatrix}, \quad (\text{XII.1b})$$

wobei der erste bzw. zweite Index des Tensors dem Zeilen- bzw. Spaltenindex der Matrix entspricht. $F^{\mu\nu}(x)$ ist offensichtlich antisymmetrisch unter dem Austausch der zwei Indizes

$$F^{\mu\nu}(x) = -F^{\nu\mu}(x) \quad \forall \mu, \nu \in \{0, 1, 2, 3\}. \quad (\text{XII.2})$$

Unter Lorentz-Transformationen $x^\mu \rightarrow x^{\mu'} = \Lambda^{\mu'}_{\mu} x^\mu$ zwischen Inertialsystemen bzw. Minkowski-Koordinaten transformiert sich der Feldstärketensor wie jeder Tensor zweiter Stufe, d.h. gemäß

$$F^{\mu\nu} \rightarrow F^{\mu'\nu'} = \Lambda^{\mu'}_{\mu} \Lambda^{\nu'}_{\nu} F^{\mu\nu}.$$

Daraus kann man die entsprechenden Transformationen für die Felder \vec{E} und \vec{B} folgern.

Für einen drehungsfreien Boost (§ X.2.2 b) mit Geschwindigkeit $\vec{v} = v \vec{e}$ findet man

$$\begin{aligned} \vec{E}' &= (\vec{e} \cdot \vec{E}) \vec{e} + \frac{\vec{e} \times (\vec{E} \times \vec{e}) - \vec{v} \times \vec{B}}{\sqrt{1 - \vec{v}^2/c^2}} \\ \vec{B}' &= (\vec{e} \cdot \vec{B}) \vec{e} + \frac{\vec{e} \times (\vec{B} \times \vec{e}) + \vec{v} \times \vec{E}/c^2}{\sqrt{1 - \vec{v}^2/c^2}} \end{aligned}$$

mit \vec{e} dem Einheitsvektor in Richtung des Boosts.

Sei $u = \{u^\mu\}_{\mu=0\dots3}$ die Vierergeschwindigkeit eines Beobachters \mathcal{B}' relativ zum Inertialsystem \mathcal{B} . Man kann Vierervektoren $E_u(x)$ und $B_u(x)$ durch die Angabe ihrer jeweiligen kontravarianten Komponenten

$$E_u^\mu(x) \equiv F^{\mu\nu}(x) u_\nu \quad (\text{XII.3a})$$

$$B_u^\mu(x) \equiv \frac{1}{2c} \epsilon^{\mu\nu\rho\sigma} u_\nu F_{\rho\sigma}(x) \quad (\text{XII.3b})$$

bezüglich \mathcal{B} definieren, wobei $\epsilon^{\mu\nu\rho\sigma}$ den vierdimensionalen Levi-Civita-Tensor (X.46) bezeichnet. Falls \mathcal{B}' in \mathcal{B} ruht — d.h. $u^0 = -u_0 = 1$ und $u^i = 0$ für $i = 1, 2, 3$ —, dann sind die räumlichen Komponenten der entsprechenden Vierervektoren $E_0(x)$ bzw. $B_0(x)$ gleich den Komponenten in \mathcal{B} der Dreiervektoren $\vec{E}(x)$ bzw. $\vec{B}(x)$, während die 0-Komponenten von $E_0(x)$ und $B_0(x)$ Null sind. Allgemeiner stellt $E_u(x)$ bzw. $B_u(x)$ das elektrische bzw. magnetische Feld relativ zum Ruhesystem des Beobachters \mathcal{B}' dar, wie es vom Bezugssystem \mathcal{B} aus gesehen aussieht.

XII.1.1 b Dualer Feldstärketensor

Mithilfe des vierdimensionalen Levi-Civita-Tensors definiert man noch den *dualen elektromagnetischen Feldstärketensor* $\tilde{\mathbf{F}}$, dessen kontravariante Komponenten durch

$$\tilde{F}^{\mu\nu}(x) \equiv \frac{1}{2} \epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma}(x) \quad (\text{XII.4a})$$

definiert sind. In Matrixdarstellung ergibt sich⁽⁸⁸⁾

$$\tilde{F}^{\mu\nu}(\mathbf{x}) = \begin{pmatrix} 0 & B^1(\mathbf{x}) & B^2(\mathbf{x}) & B^3(\mathbf{x}) \\ -B^1(\mathbf{x}) & 0 & -\frac{E^3(\mathbf{x})}{c} & \frac{E^2(\mathbf{x})}{c} \\ -B^2(\mathbf{x}) & \frac{E^3(\mathbf{x})}{c} & 0 & -\frac{E^1(\mathbf{x})}{c} \\ -B^3(\mathbf{x}) & -\frac{E^2(\mathbf{x})}{c} & \frac{E^1(\mathbf{x})}{c} & 0 \end{pmatrix}. \quad (\text{XII.4b})$$

Relativ zu den Komponenten $F^{\mu\nu}$ von \mathbf{F} folgen die $\tilde{F}^{\mu\nu}$ aus den Substitutionen $E^i/c \rightarrow B^i$ und $B^i \rightarrow -E^i/c$.

Der duale Feldstärketensor ist deutlich antisymmetrisch.

Bemerkungen:

* Die Komponenten des in Gl. (XII.3) eingeführten Vierervektors $B_u(\mathbf{x})$ lassen sich einfacher als

$$B_u^\mu(\mathbf{x}) \equiv \frac{1}{c} \tilde{F}^{\mu\nu}(\mathbf{x}) u_\nu$$

schreiben. Wiederum gilt $E_u^\mu(\mathbf{x}) \equiv -\frac{1}{2} \epsilon^{\mu\nu\rho\sigma} u_\nu \tilde{F}_{\rho\sigma}(\mathbf{x})$.

* Der „biduale“ Feldstärketensor mit Komponenten $\tilde{\tilde{F}}^{\mu\nu} = \frac{1}{2} \epsilon^{\mu\nu\rho\sigma} \tilde{F}_{\rho\sigma}$ ist einfach $-\mathbf{F}$.

XII.1.1 c Invarianten des elektromagnetischen Feldes

Aus dem Feldstärketensor (XII.1) und seinem Dual (XII.4) lassen sich durch Kontraktionen zwei unabhängige Größen finden, die sich unter Transformationen zwischen Inertialsystemen nicht ändern, d.h. zwei Lorentz-Skalaren, und zwar

$$F_{\mu\nu}(\mathbf{x}) F^{\mu\nu}(\mathbf{x}) = 2 \left[\vec{B}(\mathbf{x})^2 - \frac{\vec{E}(\mathbf{x})^2}{c^2} \right] \quad (\text{XII.5a})$$

und

$$\tilde{F}_{\mu\nu}(\mathbf{x}) F^{\mu\nu}(\mathbf{x}) = -4 \frac{\vec{E}(\mathbf{x}) \cdot \vec{B}(\mathbf{x})}{c}. \quad (\text{XII.5b})$$

Die dritte natürliche Möglichkeit $\tilde{F}_{\mu\nu}(\mathbf{x}) \tilde{\tilde{F}}^{\mu\nu}(\mathbf{x})$ ist nicht unabhängig von den zwei ersten, sondern gleich $-F_{\mu\nu}(\mathbf{x}) F^{\mu\nu}(\mathbf{x})$.

XII.1.2 Viererpotential

Das Skalarpotential Φ und das Vektorpotential \vec{A} bilden einen Vierervektor, das *Viererpotential*, mit den kontravarianten Komponenten

$$\mathbf{A}(\mathbf{x}) = \begin{pmatrix} \Phi(\mathbf{x}) \\ c \\ \vec{A}(\mathbf{x}) \end{pmatrix}. \quad (\text{XII.6})$$

Der elektromagnetische Feldstärketensor (XII.1) lässt sich dadurch ausdrücken gemäß

$$F^{\mu\nu}(\mathbf{x}) = \partial^\mu A^\nu(\mathbf{x}) - \partial^\nu A^\mu(\mathbf{x}), \quad (\text{XII.7})$$

wobei $\partial^\mu \equiv \partial / \partial x_\mu$, d.h. $\partial^0 = -(1/c) \partial / \partial t$ während ∂^i die i -te Komponente des Gradienten ist.

⁽⁸⁸⁾ Dabei muss die Leserin darauf aufpassen, dass numerisch $F_{0i} = -F_{i0} = -F^{0i} = -E^i/c$ und $F_{ij} = F^{ij}$ gelten, wie sich aus $F_{\rho\sigma} = \eta_{\rho\mu} \eta_{\sigma\nu} F^{\mu\nu}$ berechnen lässt.

Dementsprechend stellen die Beziehungen

$$\frac{E^i}{c} = F^{0i} = \partial^0 A^i - \partial^i A^0 = -\frac{1}{c} \frac{\partial A^i}{\partial t} - \frac{\partial}{\partial x^i} \left(\frac{\Phi}{c} \right)$$

$$B^i = \frac{1}{2} \sum_{j,k=1}^3 \epsilon^{ijk} F^{jk} = \sum_{j,k=1}^3 \epsilon^{ijk} \partial^j A^k = \sum_{j,k=1}^3 \epsilon^{ijk} \frac{\partial A^k}{\partial x^j}$$

jeweils die i -te Komponente der Gleichung (IX.13) und (IX.12) dar.

Eichinvarianz

In relativistisch kovarianter Schreibweise lautet die Eichtransformation (IX.14)

$$A^\mu(x) \rightarrow A'^\mu(x) = A^\mu(x) + \partial^\mu \chi(x), \tag{XII.8}$$

mit $\chi(x)$ einer zweimal kontinuierlich differenzierbaren skalaren Funktion. Für solche Funktionen gilt $\partial^\mu \partial^\nu \chi(x) = \partial^\nu \partial^\mu \chi(x)$, so dass die Viererpotentiale A und A' über Beziehung (XII.7) zum gleichen Feldstärketensor \mathbf{F} führen, wie zu erwarten war.

Unter Berücksichtigung der Beziehung $\epsilon_0 \mu_0 = 1/c^2$ [Gl. (IX.29c)] wird die Bedingung (IX.16), welche die Lorenz-Eichung definiert, zu

$$\partial_\mu A^\mu(x) = 0. \tag{XII.9}$$

Dabei ist deutlich, dass diese Eichbedingung Lorentz-invariant ist, d.h. sie bleibt in allen Inertialsystemen erfüllt. Dagegen gilt diese Eigenschaft nicht für die Coulomb-Eichbedingung, für die es keine Lorentz-kovariante Formulierung gibt.

XII.1.3 Elektrischer Viererstrom

Die elektrischen Ladungs- und Stromdichte bilden einen Vierervektor $J_{\text{el.}}(x)$, der als *elektrischer Viererstrom* bzw. *Viererstromdichte* — was genauer ist — bezeichnet wird:

$$J_{\text{el.}}(x) = \begin{pmatrix} \rho_{\text{el.}}(x)c \\ \vec{j}_{\text{el.}}(x) \end{pmatrix}. \tag{XII.10a}$$

Komponentenweise entspricht dies

$$J_{\text{el.}}^0(x) = \rho_{\text{el.}}(x)c, \quad J_{\text{el.}}^i(x) = j_{\text{el.}}^i(x) \text{ für } i = 1, 2, 3. \tag{XII.10b}$$

Dabei bezeichnen $\rho_{\text{el.}}$ und $\vec{j}_{\text{el.}}$ die bekannten „nicht-relativistischen“ Größen in einem festen Bezugssystem, das sich möglicherweise relativ zu den Ladungen bewegt.

Für eine Punktladung q mit Weltlinie bzw. Vierergeschwindigkeit $(t, \vec{x}(t))$ bzw. $u(\tau) = dx(\tau)/d\tau$, mit τ der Eigenzeit der Punktladung, gilt

$$J_{\text{el.}}(x) = q \delta^{(3)}(\vec{r} - \vec{x}(t)) u(x) = \varrho_0(x) u(x), \tag{XII.11a}$$

wobei $\varrho_0(x)$ die Ladungsdichte im Ruhesystem der Punktladung bezeichnet. Unter Berücksichtigung des Ausdrucks (X.27) der Vierergeschwindigkeit lautet dies noch

$$J_{\text{el.}}(x) = \begin{pmatrix} \varrho_0(x) \gamma c \\ \varrho_0(x) \gamma \vec{v} \end{pmatrix} \tag{XII.11b}$$

mit $\gamma = 1/\sqrt{1 - \vec{v}^2/c^2}$ dem Lorentz-Faktor.

XII.1.4 Energieimpulstensor

Die Energiedichte $e_{\text{e.m.}}$ [Gl. (IX.23a)], der Poynting-Vektor $\vec{S}_{\text{e.m.}}$ [Gl. (IX.23b)], die Impulsdichte $\vec{g}_{\text{e.m.}}$ [Gl. (IX.26a)] und die Impulsstromdichte $\mathbf{T}_{\text{e.m.}} = -\boldsymbol{\sigma}_{\text{e.m.}}$ [Gl. (IX.26b)–(IX.26c)] des elektromagnetischen Feldes lassen sich in einen Vierertensor zweiter Stufe $\mathbf{T}_{\text{e.m.}}$, den *Energieimpulstensor* des Feldes, kombinieren. In Matrixdarstellung lautet dieser

$$T_{\text{e.m.}}^{\mu\nu}(\mathbf{x}) = \left(\begin{array}{c|c} e_{\text{e.m.}}(\mathbf{x}) & c\vec{g}_{\text{e.m.}}(\mathbf{x}) \\ \hline \frac{\vec{S}_{\text{e.m.}}(\mathbf{x})}{c} & \mathbf{T}_{\text{e.m.}}(\mathbf{x}) \end{array} \right). \quad (\text{XII.12a})$$

Der Energieimpulstensor lässt sich auch durch den Feldstärketensor (XII.1) ausdrücken, und zwar komponentenweise

$$T_{\text{e.m.}}^{\mu\nu}(\mathbf{x}) = \frac{1}{\mu_0} \left[-F^\mu{}_\rho(\mathbf{x}) F^{\rho\nu}(\mathbf{x}) - \frac{1}{4} \eta^{\mu\nu} F_{\rho\sigma}(\mathbf{x}) F^{\rho\sigma}(\mathbf{x}) \right], \quad (\text{XII.12b})$$

mit $\eta^{\mu\nu}$ den Komponenten des inversen metrischen Tensors. In dieser Form prüft man einfach, dass der Energieimpulstensor symmetrisch ist.

XII.2 Relativistisch kovariante Formulierung der Grundgesetze

Mit den in Abschn. XII.1 eingeführten Größen lassen sich die Maxwell-Gleichungen (§ XII.2.1), die Kontinuitätsgleichung (§ XII.2.2), die Lorentz-Kraftdichte (§ XII.2.3) und die Energie- und Impulsbilanz (§ XII.2.4) in relativistisch kovarianter Form schreiben.

XII.2.1 Maxwell-Gleichungen

Mit dem elektromagnetischen Feldstärketensor (XII.1) und dem elektrischen Viererstrom (XII.10) lauten die Maxwell-Gleichungen

$$\partial_\nu F^{\mu\nu}(\mathbf{x}) = \mu_0 J_{\text{el.}}^\mu(\mathbf{x}) \quad \forall \mu \in \{0, 1, 2, 3\}, \quad (\text{XII.13a})$$

$$\partial^\mu F^{\nu\rho}(\mathbf{x}) + \partial^\nu F^{\rho\mu}(\mathbf{x}) + \partial^\rho F^{\mu\nu}(\mathbf{x}) = 0 \quad \forall \mu, \nu, \rho \in \{0, 1, 2, 3\}. \quad (\text{XII.13b})$$

Dabei stellen die Gl. (XII.13a) mit einem Quellterm die inhomogenen Maxwell–Gauß- und Maxwell–Ampère-Gleichungen (IX.1a) und (IX.1d) dar.

Wiederum stehen die Gl. (XII.13b) — wobei nur die vier Fälle $(\mu, \nu, \rho) = (0, 1, 2), (0, 1, 3), (0, 2, 3)$ und $(1, 2, 3)$ unterschiedlich sind — für die homogenen Maxwell-Gleichungen (IX.1b) und (IX.1c). Unter Verwendung des dualen elektromagnetischen Feldstärketensors (XII.4) lassen sich die homogenen Maxwell-Gleichungen noch als

$$\partial_\mu \tilde{F}^{\mu\nu}(\mathbf{x}) = 0 \quad \forall \nu \in \{0, 1, 2, 3\} \quad (\text{XII.13c})$$

umschreiben. In dieser Form ist sofort klar, dass es nur vier Gleichungen gibt.

XII.2.2 Kontinuitätsgleichung

Ausgedrückt durch den elektrischen Viererstrom (XII.10) lautet die Kontinuitätsgleichung (IX.7)

$$\partial_\mu J_{\text{el.}}^\mu(\mathbf{x}) = 0. \quad (\text{XII.14})$$

Genau wie in § IX.1.1 b lässt sich diese Bilanzgleichung aus den homogenen Maxwell-Gleichungen herleiten: somit lautet die Viererdivergenz der Gl. (XII.13a)

$$\partial_\mu \partial_\nu F^{\mu\nu}(\mathbf{x}) = \mu_0 \partial_\mu J_{\text{el.}}^\mu(\mathbf{x}).$$

Dabei sind auf der linken Seite $\partial_\mu \partial_\nu$ symmetrisch und $F^{\mu\nu}$ antisymmetrisch unter dem Austausch von μ und ν , so dass die zweifache Kontraktion $\partial_\mu \partial_\nu F^{\mu\nu}$ automatisch Null ist, entsprechend Gl. (XII.14).

XII.2.3 Lorentz-Kraft und -Kraftdichte

Das elektromagnetische Feld übt auf eine elektrische Ladungs- und Stromverteilung eine Lorentz-Kraftdichte \vec{f}_L [Gl. (VI.2)] aus. Mit dieser kann man eine Viererkraftdichte f_L assoziieren, die sich durch den elektromagnetischen Feldstärketensor (XII.1) und den elektrischen Viererstrom (XII.10) ausdrücken lässt. Komponentenweise gilt

$$f_L^\nu(\mathbf{x}) = J_{\text{el.},\mu}(\mathbf{x}) F^{\mu\nu}(\mathbf{x}). \quad (\text{XII.15a})$$

Mit den Ausdrücken des Feldstärketensors und des Viererstroms prüft man nämlich schnell, dass die räumlichen Komponenten

$$f_L^i(\mathbf{x}) = \rho_{\text{el.}}(\mathbf{x}) E^i(\mathbf{x}) + \sum_{k,l=1}^3 \epsilon^{ikl} j_{\text{el.}}^k(\mathbf{x}) B^l(\mathbf{x}) \quad \text{für } i = 1, 2, 3 \quad (\text{XII.15b})$$

betragen. Wiederum lautet die zeitliche Komponente

$$f_L^0(\mathbf{x}) = \frac{1}{c} \vec{j}_{\text{el.}}(\mathbf{x}) \cdot \vec{E}(\mathbf{x}). \quad (\text{XII.15c})$$

Im Fall einer Punktladung q mit Vierergeschwindigkeit $\mathbf{u} = (\gamma c, \gamma \vec{v})$, entsprechend dem elektrischen Viererstrom (XII.11a), kann die Lorentz-Kraftdichte sofort über die Ortskoordinaten integriert werden. Daraus ergibt sich die Lorentz-Viererkraft

$$F_L^\nu = q u_\mu F^{\mu\nu}(\mathbf{x}). \quad (\text{XII.16})$$

Mit dieser Viererkraft lautet die relativistisch kovariante Verallgemeinerung (XI.14) des zweiten Newton'schen Gesetzes

$$\frac{dp^\nu}{d\tau} = q u_\mu F^{\mu\nu}(\mathbf{x}) \quad (\text{XII.17})$$

mit $\tau = t/\gamma$ der Eigenzeit der Punktladung und $\mathbf{p} = m\mathbf{u}$ ihrer (kinetischen) Viererimpuls. Dabei lautet die zeitliche Komponente $dp^0/d\tau = q\gamma \vec{v} \cdot \vec{E}/c$, d.h. $d(p^0 c)/dt = q\vec{E} \cdot \vec{v}$: auf der linken Seite steht die Rate der zeitlichen Änderung von der Energie $p^0 c$ der Punktladung; auf der rechten, die instantane Leistung der Lorentz-Kraft.

Bemerkung: Die Lorentz-Viererkraft (XII.16) ist „orthogonal“ zur Vierergeschwindigkeit, $u_\nu F_L^\nu = 0$.

Dies folgt aus der Antisymmetrie des elektromagnetischen Feldstärketensors und der Symmetrie des Produkts $u_\nu u_\mu$.

XII.2.4 Energie- und Impulsbilanzgleichungen

Schließlich lassen sich die lokalen Bilanzgleichungen (IX.23) und (IX.26d) für die Energie und den Impuls des elektromagnetischen Feldes unter Verwendung des Energieimpulstensors (XII.12) und der Lorentz-Viererkraftdichte (XII.15a) in kompakter Form ausdrücken:

$$\partial_\mu T_{\text{e.m.}}^{\mu\nu}(\mathbf{x}) = -f_L^\nu(\mathbf{x}). \quad (\text{XII.18})$$

XII.3 Weitere Resultate in relativistisch kovarianter Form

XII.3.1 Bewegungsgleichung für das Viererpotential

Setzt man den Ausdruck (XII.7) des Feldstärketensors durch das Viererpotential in die inhomogenen Maxwell-Gleichungen (XII.13a) ein, so findet man

$$\partial_\nu \partial^\mu A^\nu(x) - \partial_\nu \partial^\nu A^\mu(x) = \mu_0 J_{\text{el.}}^\mu(x).$$

Dabei ist $\partial_\nu \partial^\nu$ gerade der d'Alembert-Operator \square [Gl. (IX.10)], so dass diese Gleichung in der Form

$$\square A^\mu(x) = -\mu_0 J_{\text{el.}}^\mu(x) + \partial^\mu [\partial_\nu A^\nu(x)] \quad \forall \mu \in \{0, 1, 2, 3\} \quad (\text{XII.19})$$

umgeschrieben werden kann. Dies stellt die Bewegungsgleichung für das Viererpotential $A(x)$ dar. Die Komponente $\nu = 0$ und ist äquivalent zur Bewegungsgleichungen (IX.20) des Skalarpotentials, die räumlichen Komponenten entsprechend der Gl. (IX.18) für das Vektorpotential.

Bemerkung: Aus der Beziehung (XII.7) folgt

$$\tilde{F}^{\mu\nu}(x) = \frac{1}{2} \epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma}(x) = \epsilon^{\mu\nu\rho\sigma} \partial_\rho A_\sigma(x),$$

woraus die homogene Gleichung $\partial_\mu \tilde{F}^{\mu\nu}(x) = \epsilon^{\mu\nu\rho\sigma} \partial_\mu \partial_\rho A_\sigma(x) = 0$ sofort folgt, da $\epsilon^{\mu\nu\rho\sigma}$ antisymmetrisch und $\partial_\mu \partial_\rho$ symmetrisch unter dem Austausch von μ und ρ ist.

XII.3.2 Klassische Wellengleichung und ebene Wellen

In Abwesenheit von äußeren Quellen des Feldes und in der Lorenz-Eichung $\partial_\nu A^\nu(x) = 0$ wird die Bewegungsgleichung (XII.19) zur klassischen Wellengleichung (vgl. § IX.4.1)

$$\square A^\mu(x) = 0 \quad \forall \mu \in \{0, 1, 2, 3\} \quad (\text{XII.20a})$$

oder äquivalent, in geometrischer Form

$$\square A(x) = 0. \quad (\text{XII.20b})$$

Eine ebene Welle ist eine Lösung dieser partiellen Differentialgleichung der Form [vgl. § IX.4.1 a)

$$A^\mu(x) = \epsilon^\mu f(\vec{k} \cdot \vec{r} - \omega_{\vec{k}} t) \quad (\text{XII.21a})$$

wobei $\omega_{\vec{k}} = c|\vec{k}|$ [Gl. (IX.43c)]. Dabei ist \vec{k} ein Vektor, der die Propagationsrichtung angibt, und ϵ^μ die Komponente eines *Polarisationsvierervektors* ϵ . Definiert man einen (lichtartigen) Vierervektor $k = (k^0 \equiv \omega_{\vec{k}}/c, \vec{k})$, so lässt sich diese Lösung als

$$A^\mu(x) = \epsilon^\mu f(k \cdot x) \quad (\text{XII.21b})$$

schreiben, oder auch geometrisch

$$A(x) = \epsilon f(k \cdot x). \quad (\text{XII.21c})$$

Wegen der Lorenz-Eichbedingung (XII.9) sollen ϵ und k i Allgemeinen orthogonal zueinander sein, $k \cdot \epsilon = 0$. In einem Bezugssystem, wo A^0 verschwindet,⁽⁸⁹⁾ gilt automatisch $\epsilon^0 = 0$ — woraus man erkennt, dass der Polarisationsvierervektor raumartig ist, $\epsilon^2 = \epsilon_\mu \epsilon^\mu > 0$. In diesem Bezugssystem gilt somit $k \cdot \epsilon = \vec{k} \cdot \vec{\epsilon} = 0$, und man findet die Transversalität des Polarisationsvektors wieder.

⁽⁸⁹⁾Diese Bedingung definiert die *temporale Weyl^(ay)-Eichung*, die hier, kombiniert mit der Lorenz-Eichung $\partial_\mu A^\mu = 0$, äquivalent zur Coulomb-Eichung $\vec{\nabla} \cdot \vec{A} = 0$ ist.

^(ay)H. WEYL, 1885–1955

XII.3.3 Retardiertes Viererpotential

In der Lorenz-Eichung (XII.9) verschwindet der letzte Term der Bewegungsgleichung (XII.19) für das Viererpotential, die sich somit zu

$$\square A^\mu(\mathbf{x}) = -\mu_0 J_{\text{el.}}^\mu(\mathbf{x}) \quad \forall \nu \quad (\text{XII.22})$$

vereinfacht.

Wie im § IX.5.2 führt die Lösung dieser inhomogenen partiellen Differentialgleichung mithilfe der retardierten Green'schen Funktion (IX.58a) für die klassische Wellengleichung zum *retardierten Viererpotential*

$$A_{\text{ret.}}(\mathbf{x}) = \frac{\mu_0}{4\pi} \int \frac{1}{|\vec{r} - \vec{r}'|} J_{\text{el.}} \left(t - \frac{|\vec{r} - \vec{r}'|}{c}, \vec{r}' \right) d^3 \vec{r}', \quad (\text{XII.23})$$

entsprechend den Gl. (IX.64).

Im Fall des elektrischen Viererstroms einer bewegten Punktladung mit Weltlinie $x_q(t) = (t, \vec{x}(t))$ ergibt sich das Liénard–Wiechert-Viererpotential

$$A(\mathbf{x}) = \frac{\mu_0}{4\pi} \frac{dx_q(t_{\text{ret.}})}{dt} \frac{q}{|\vec{r} - \vec{x}(t_{\text{ret.}})| - [\vec{r} - \vec{x}(t_{\text{ret.}})] \cdot \frac{\vec{v}_{\text{ret.}}}{c}}, \quad (\text{XII.24a})$$

mit $t_{\text{ret.}}$ der retardierten Zeit und $\vec{v}_{\text{ret.}}$ der entsprechenden Dreiergeschwindigkeit der Punktladung, vgl. § IX.5.4 a.

Man kann eine relativistisch kovariante Form dieses Viererpotentials finden. Sei $\mathbf{u}_{\text{ret.}}$ die Vierergeschwindigkeit im retardierten Punkt $\mathbf{x}_{\text{ret.}} \equiv x_q(t_{\text{ret.}})$ und $\mathbf{X} \equiv \mathbf{x} - \mathbf{x}_{\text{ret.}} = (c(t - t_{\text{ret.}}), \vec{r} - \vec{x}(t_{\text{ret.}}))$. Dann gilt der deutlich kovariante Ausdruck

$$A(\mathbf{x}) = \frac{\mu_0 c}{4\pi} \frac{q \mathbf{u}_{\text{ret.}}}{\mathbf{X} \cdot \mathbf{u}_{\text{ret.}}}. \quad (\text{XII.24b})$$

Beweis: Im Inertialsystem \mathcal{B}_0 , das sich zur retardierten Zeit mit der Geschwindigkeit $\vec{v}_{\text{ret.}}$ bewegt, d.h. in welchem $\vec{v}_{\text{ret.}} = \vec{0}$, vereinfacht sich der Nenner des Viererpotentials (XII.24a) zu

$$|\vec{r} - \vec{x}(t_{\text{ret.}})| = c(t - t_{\text{ret.}}) = \frac{\mathbf{X} \cdot \mathbf{u}_{\text{ret.}}}{c},$$

wobei die erste Gleichung der Definition der retardierten Zeit entspricht. Die Zeit t in \mathcal{B}_0 ist die Eigenzeit τ der Punktladung, so dass $dx_q(t_{\text{ret.}})/dt$ genau $dx_q(t_{\text{ret.}})/d\tau = \mathbf{u}_{\text{ret.}}$ entspricht. Somit stimmt Gl. (XII.24a) in \mathcal{B}_0 mit Gl. (XII.24b) überein. Da die letztere eine Gleichung zwischen Vierervektoren ist, bleibt sie in jedem Inertialsystem gültig. \square

Literatur zum Kapitel XII

- Feynman, *Vorlesungen über Physik. Band 2* [22] = *Lectures on Physics. Vol. II* [23] Kap. 25, 26 & 31.8
- Fließbach, *Elektrodynamik* [3] Teil IV Kap. 18 & Teil V Kap. 22
- Greiner, *Klassische Elektrodynamik* [8] Teil IV Kap. 22.
- Griffiths, *Elektrodynamik* [9] = *Introduction to Electrodynamics* [10] Kap. 12.3
- Jackson, *Klassische Elektrodynamik* [11] = *Classical Electrodynamics* [12] Kap. 11.9–11.10 & 12.11.
- Landau & Lifschitz, *Klassische Feldtheorie* [14] = *The classical theory of fields* [28] Kap. III § 16–18 & 23–25.
- Nolting, *Spezielle Relativitätstheorie, Thermodynamik* [32] Erster Teil, Kap. 2.3.
- Scheck, *Klassische Feldtheorie* [19] Kap. 2.2.

KAPITEL XIII

Lagrange-Formulierung der Elektrodynamik

Die Gesetze der Elektrodynamik — und zwar die Maxwell-Gleichungen und die Form der Lorentz-Kraft — sind dadurch motiviert, dass die daraus folgenden Vorhersagen im Einklang mit experimentellen Ergebnissen sind. Diese Gesetze können aber auch hergeleitet werden, und zwar aus einem Extremalprinzip: ähnlich wie in Kap. III führt die Forderung, dass eine Wirkung ihr Extremum für die physikalisch realisierte Konfiguration von Feld und Quellen erreicht, zu Euler-Lagrange-Gleichungen, die genau die üblichen Bewegungsgleichungen sind. Hiernach wird einem System aus bewegten Punktladungen — oder allgemeiner aus elektrischen Ladungs- und Stromverteilungen — und einem elektromagnetischen Feld eine Lagrange-Funktion zugeordnet. Die Letztere besteht aus drei Beiträgen, die jeweils die freien Ladungen, das freie elektromagnetische Feld und deren Wechselwirkungsterm beschreiben.

XIII.1 Ladungen und Ströme in einem elektromagnetischen Feld

Wir betrachten zuerst die Lagrange-Formulierung der Dynamik einer Punktladung, oder allgemeiner von Punktladungen bzw. von einer Ladungs- und Stromverteilung, in einem nicht-dynamischen elektromagnetischen Feld.

XIII.1.1 Wiederholung: Lagrange-Funktion einer freien Punktladung

Die Lagrange-Funktion eines freien punktförmigen Teilchens der Masse m wurde in früheren Kapiteln dieser Vorlesung schon diskutiert. Sei $\vec{x}(t)$ bzw. $\vec{v}(t)$ die Position bzw. die Geschwindigkeit des Teilchens relativ zu einem festen Bezugssystem.

Falls das Teilchen nicht-relativistisch ist (§ III.2.3 a) lautet seine Lagrange-Funktion [Gl. (III.18)]

$$\mathcal{L}_M^{(n.-r.)}(t, \vec{x}(t), \vec{v}(t)) = \frac{m}{2} \vec{v}(t)^2, \quad (\text{XIII.1a})$$

wobei das tiefgestellte Kürzel M für Materie steht. Für ein relativistisches Teilchen gilt [§ XI.1.1, Gl. (XI.3a)]

$$\mathcal{L}_M^{(\text{rel.})}(t, \vec{x}(t), \vec{v}(t)) = -mc^2 \sqrt{1 - \frac{\vec{v}(t)^2}{c^2}}. \quad (\text{XIII.1b})$$

Entsprechend diesen Lagrange-Funktionen wird der kanonisch konjugierte Impuls $\vec{p} \equiv \partial \mathcal{L}_M / \partial \vec{v}$ definiert, und zwar

$$\vec{p} = m\vec{v} \quad (\text{XIII.2a})$$

im nicht-relativistischen Fall bzw.

$$\vec{p} = \frac{m\vec{v}}{\sqrt{1 - \vec{v}^2/c^2}} \quad (\text{XIII.2b})$$

[Gl. (XI.4)] für das relativistische Teilchen.

Das Integral der Lagrange-Funktion (XIII.1) über die Zeit entlang einer Bahn $\vec{x}(t)$ gibt die zugehörige Wirkung

$$S_M = \int_{t_a}^{t_b} \mathcal{L}_M^{(n-r)}(t, \vec{x}(t), \vec{v}(t)) dt \quad (\text{XIII.3a})$$

bzw.

$$S_M = \int_{t_a}^{t_b} \mathcal{L}_M^{(\text{rel.})}(t, \vec{x}(t), \vec{v}(t)) dt = -mc^2 \int_{\tau_a}^{\tau_b} d\tau, \quad (\text{XIII.3b})$$

wobei τ die Eigenzeit im des Teilchens bezeichnet.

Die physikalisch realisierte Bahnkurve ergibt sich laut dem Hamilton-Prinzip (III.10) aus der Extremierung der Wirkung. Dies führt zu den üblichen Euler-Lagrange-Gleichungen (III.11)

$$\frac{\partial \mathcal{L}_M(t, \vec{x}(t), \vec{v}(t))}{\partial \vec{x}} = \frac{d}{dt} \frac{\partial \mathcal{L}_M(t, \vec{x}(t), \vec{v}(t))}{\partial \vec{v}}, \quad (\text{XIII.4a})$$

wobei die Schreibweise mit Ableitungen nach Vektoren eine günstige Notation für

$$\frac{\partial \mathcal{L}_M(t, \vec{x}(t), \vec{v}(t))}{\partial x^i} = \frac{d}{dt} \frac{\partial \mathcal{L}_M(t, \vec{x}(t), \vec{v}(t))}{\partial v^i} \quad \text{für } i = 1, 2, 3, \quad (\text{XIII.4b})$$

darstellt. Für die Lagrange-Funktionen (XIII.1) ergibt sich einfach

$$\frac{d\vec{p}(t)}{dt} = \vec{0} \quad (\text{XIII.5})$$

mit \vec{p} dem kanonisch konjugierten Impuls (XIII.2).

XIII.1.2 Punktladung in einem äußeren elektromagnetischen Feld

Sei jetzt angenommen, dass sich die Punktladung mit Masse m und elektrischer Ladung q in einem *äußeren*, vorgegebenen elektromagnetischen Feld befindet. Das Letztere wird durch Potentiale $\Phi(t, \vec{r})$ und $\vec{A}(t, \vec{r})$ oder äquivalent ein Viererpotential $A(t, \vec{r})$ [Gl. (XII.6)] beschrieben.

XIII.1.2a Lagrange-Funktion und Wirkung

Der Beitrag zur Lagrange-Funktion zur Beschreibung der Wechselwirkung zwischen Punktladung und elektromagnetischem Feld lautet [33]

$$\mathcal{L}_{M+F}(t, \vec{x}(t), \vec{v}(t)) = -q\Phi(t, \vec{x}(t)) + q\vec{v}(t) \cdot \vec{A}(t, \vec{x}(t)) \quad (\text{XIII.6a})$$

oder äquivalent, unter Einführung der Komponenten $x^\mu(t)$ bzw. $A^\mu(t, \vec{r})$ der Weltlinie der Punktladung bzw. des Viererpotentials

$$\mathcal{L}_{M+F}(t, \vec{x}(t), \vec{v}(t)) = q \frac{dx_\mu(t)}{dt} A^\mu(t, \vec{x}(t)). \quad (\text{XIII.6b})$$

Dabei steht das Kürzel M+F für die Wechselwirkung zwischen Materie und Feld.

Aus dieser Lagrange-Funktion folgt die Wirkung

$$S_{M+F} = \int_{t_a}^{t_b} \mathcal{L}_{M+F}(t, \vec{x}(t), \vec{v}(t)) dt = q \int_a^b A^\mu(t, \vec{x}(t)) dx_\mu = q \int_{\tau_a}^{\tau_b} u_\mu(t) A^\mu(t, \vec{x}(t)) d\tau, \quad (\text{XIII.7})$$

mit $u_\mu(t)$ den kovarianten Koordinaten der Vierergeschwindigkeit des Teilchens. Diese Wirkung ist deutlich Lorentz-invariant, wie die Wirkung (XIII.3b) des relativistischen freien Teilchens. Dagegen ist die Lagrange-Funktion (XIII.6) nicht Lorentz-invariant.

Die Lagrange-Funktion (XIII.6) ist auch nicht invariant unter Eichtransformationen der Potentiale. Unter einer Transformation (IX.14) — oder äquivalent (XII.8) für das Viererpotential — ändert sich die Lagrange-Funktion gemäß

$$\begin{aligned}\mathcal{L}_{M+F}(\vec{x}(t), \vec{v}(t)) &\rightarrow \mathcal{L}'_{M+F}(\vec{x}(t), \vec{v}(t)) = \mathcal{L}_{M+F}(\vec{x}(t), \vec{v}(t)) + q \left\{ \frac{\partial \chi(t, \vec{x}(t))}{\partial t} + [\vec{v}(t) \cdot \vec{\nabla}] \chi(t, \vec{x}(t)) \right\} \\ &= \mathcal{L}_{M+F}(\vec{x}(t), \vec{v}(t)) + q \frac{d\chi(t, \vec{x}(t))}{dt}.\end{aligned}$$

Da \mathcal{L}_{M+F} und \mathcal{L}'_{M+F} nur um eine totale Zeitableitung abweichen, unterscheiden sich die zugehörigen Wirkungen (XIII.7) nur um eine Konstante, die später keine Rolle für die Bewegungsgleichungen spielt [vgl. Gl. (III.12) und den entsprechenden Beweis]. Daher werden die Bewegungsgleichungen in § XIII.1.2 c eichinvariant sein.

XIII.1.2 b Kanonischer Impuls und Hamilton-Funktion

Der *kanonische Impuls* der Punktladung im elektromagnetischen Feld folgt aus der üblichen Definition, angewandt auf die gesamte Lagrange-Funktion $\mathcal{L}^{(1)} = \mathcal{L}_M + \mathcal{L}_{M+F}$:

$$\frac{\partial \mathcal{L}^{(1)}(t, \vec{x}(t), \vec{v}(t))}{\partial \vec{v}} \equiv \vec{\Pi} = \vec{p} + q\vec{A}(t, \vec{x}(t)), \quad (\text{XIII.8})$$

wobei der erste Term im rechten Glied der Impuls (XIII.2) der freien Punktladung ist, entsprechend deren *kinetischen Impuls*. Somit sind kanonischer und kinetischer Impuls einer Punktladung in Anwesenheit eines elektromagnetischen Feldes unterschiedlich.

Die Hamilton-Funktion für eine Punktladung in einem äußeren elektromagnetischen Feld ist

$$H \equiv \vec{\Pi} \cdot \vec{v} - \mathcal{L}^{(1)}. \quad (\text{XIII.9a})$$

Dies ergibt für ein nicht-relativistisches Teilchen

$$H^{(n.r.)} = \frac{m\vec{v}(t)^2}{2} + q\Phi(t, \vec{r}) \quad (\text{XIII.9b})$$

und im allgemeineren Fall einer möglicherweise relativistischen Punktladung

$$H^{(rel)} = \frac{mc^2}{\sqrt{1 - \vec{v}(t)^2/c^2}} + q\Phi(t, \vec{r}). \quad (\text{XIII.9c})$$

Dabei wird die Hamilton-Funktion aber noch nicht durch die kanonisch konjugierten Variablen $(\vec{r}, \vec{\Pi})$ ausgedrückt, wie es für den Hamilton-Formalismus nötig ist.

Führt man den Vierervektor Π mit kontravarianten Komponenten $\Pi^\mu \equiv (H/c, \vec{\Pi})$ ein, so lautet der kinetische Viererimpuls der Punktladung, bestehend aus ihrer Energie und ihrem kinetischen Impuls, $\mathbf{p} = \Pi - q\mathbf{A}(x)$ bzw. komponentenweise $p^\mu = \Pi^\mu - qA^\mu(x)$.⁽⁹⁰⁾ Bildet man den Lorentz-Quadrat dieser Identität, so kommt

$$-m^2c^2 = p_\mu p^\mu = -\frac{1}{c^2} [H - q\Phi(t, \vec{r})]^2 + [\vec{\Pi} - q\vec{A}(t, \vec{r})]^2,$$

d.h. nach trivialer Umformung

$$H(\vec{r}, \vec{\Pi}) = \sqrt{c^2 [\vec{\Pi} - q\vec{A}(t, \vec{r})]^2 + m^2c^4} + q\Phi(t, \vec{r}). \quad (\text{XIII.10})$$

Nun wird die Hamilton-Funktion durch die Variablen $(\vec{r}, \vec{\Pi})$ ausgedrückt.

XIII.1.2 c Bewegungsgleichungen

⁽⁹⁰⁾Diese Gleichungen zeigen im Nachhinein, dass Π ein wohldefinierter Vierervektor ist.

Kombiniert man die aus der Kettenregel folgende Identität

$$\frac{d\vec{A}(t, \vec{x}(t))}{dt} = \frac{\partial \vec{A}(t, \vec{x}(t))}{\partial t} + [\vec{v}(t) \cdot \vec{\nabla}] \vec{A}(t, \vec{x}(t))$$

mit der Ableitung

$$\frac{\partial \mathcal{L}^{(1)}(t, \vec{x}(t), \vec{v}(t))}{\partial \vec{x}} = -q \vec{\nabla} \Phi(t, \vec{x}) + q \vec{\nabla} [\vec{v}(t) \cdot \vec{A}(t, \vec{x}(t))],$$

so führt die Euler–Lagrange-Gleichung

$$\frac{\partial \mathcal{L}^{(1)}(t, \vec{x}(t), \vec{v}(t))}{\partial \vec{x}} = \frac{d}{dt} \frac{\partial \mathcal{L}^{(1)}(t, \vec{x}(t), \vec{v}(t))}{\partial \vec{v}}$$

zu

$$\frac{d\vec{p}(t)}{dt} = q \left\{ -\frac{\partial \vec{A}(t, \vec{x}(t))}{\partial t} - [\vec{v}(t) \cdot \vec{\nabla}] \vec{A}(t, \vec{x}(t)) - \vec{\nabla} \Phi(t, \vec{x}(t)) + \vec{\nabla} [\vec{v}(t) \cdot \vec{A}(t, \vec{x}(t))] \right\}.$$

Aus der Formel $\vec{v} \times (\vec{\nabla} \times \vec{A}) = \vec{\nabla}(\vec{v} \cdot \vec{A}) - (\vec{v} \cdot \vec{\nabla}) \vec{A}$ und den Beziehungen (IX.12) und (IX.13) ergibt sich dann die Bewegungsgleichung

$$\frac{d\vec{p}(t)}{dt} = q \left[\vec{E}(t, \vec{x}(t)) + \vec{v}(t) \times \vec{B}(t, \vec{x}(t)) \right]. \quad (\text{XIII.11})$$

Auf der linken Seite steht die Zeitableitung des kinetischen Impulses der Punktladung, auf der rechten Seite erkennt man die darauf wirkende Lorentz-Kraft: man findet die übliche Bewegungsgleichung wieder.

Bemerkung: Ausgehend aus der Hamilton-Funktion (XIII.10) führen die Hamilton-Gleichungen⁽⁹¹⁾

$$\begin{cases} \frac{\partial H}{\partial \Pi_i} = \dot{x}^i = v^i, \\ \frac{\partial H}{\partial x_i} = -\dot{\Pi}^i \end{cases} \quad \text{für } i = 1, 2, 3$$

zur gleichen Bewegungsgleichung.

XIII.1.3 Ladungs- und Stromverteilungen in einem elektromagnetischen Feld

Die Lagrange-Funktion (XIII.6) bzw. die Wirkung (XIII.7) lässt sich auf den Fall mehrerer Punktladungen $\{q_a\}$ in einem elektromagnetischen Feld sofort verallgemeinern. Somit gilt

$$\mathcal{L}_{M+F}(t, \{\vec{x}_a(t)\}, \{\vec{v}_a(t)\}) = -\sum_a q_a \Phi(t, \vec{x}_a(t)) + \sum_a q_a \vec{v}_a(t) \cdot \vec{A}(t, \vec{x}_a(t)) \quad (\text{XIII.12a})$$

$$= \sum_a q_a \frac{dx_{a,\mu}(t)}{dt} A^\mu(t, \vec{x}_a(t)). \quad (\text{XIII.12b})$$

Dabei sind die $\{\vec{x}_a(t)\}$ die Bahnkurven der Punktladungen, die $\{x_{a,\mu}(t)\}$ die kovarianten Koordinaten der entsprechenden Weltlinien, und die $\{\vec{v}_a(t)\}$ ihre Geschwindigkeiten.

Eine Integration über die Zeit liefert die zugehörige Wirkung

$$S_{M+F} = -\int_{t_a}^{t_b} \sum_a q_a \left[\Phi(t, \vec{x}_a(t)) - \vec{v}_a(t) \cdot \vec{A}(t, \vec{x}_a(t)) \right] dt \quad (\text{XIII.13a})$$

$$= \int_{t_a}^{t_b} \sum_a q_a \frac{dx_{a,\mu}(t)}{dt} A^\mu(t, \vec{x}_a(t)) dt. \quad (\text{XIII.13b})$$

⁽⁹¹⁾Die Position der Indizes in diesen Gleichungen ist nicht willkürlich!

Diese Ausdrücke können noch transformiert werden, und zwar in der Form

$$S_{M+F} = - \int_{t_a}^{t_b} \int_{\mathbb{R}^3} \sum_a q_a \delta^{(3)}(\vec{r} - \vec{x}_a(t)) \left[\Phi(t, \vec{r}) - \vec{v}_a(t) \cdot \vec{A}(t, \vec{r}) \right] d^3\vec{r} dt.$$

Unter Verwendung der elektrischen Ladungsdichte $\rho_{\text{el.}}$ und -stromdichte $\vec{j}_{\text{el.}}$ [Gl. (??)] der Punktladungen ergibt sich dann

$$S_{M+F} = - \int_{t_a}^{t_b} \int_{\mathbb{R}^3} \left[\rho_{\text{el.}}(t, \vec{r}) \Phi(t, \vec{r}) - \vec{j}_{\text{el.}}(t, \vec{r}) \cdot \vec{A}(t, \vec{r}) \right] d^3\vec{r} dt. \quad (\text{XIII.14a})$$

Ausgedrückt durch den aus Gl. (XII.11) folgenden elektrischen Viererstrom $J_{\text{el.}}$ lautet dies noch

$$S_{M+F} = \int J_{\text{el.}}(\mathbf{x}) \cdot \mathbf{A}(\mathbf{x}) \frac{d^4\mathbf{x}}{c}, \quad (\text{XIII.14b})$$

wobei $d^4\mathbf{x}$ das Vierervolumenelement bezeichnet [Gl. (X.21)], während das Integrationsgebiet nicht präzisiert wurde.

Diese Ausdrücke der Wirkung — und dabei der Lagrange-Funktion, die als Integrand des Integrals über die Zeit in Gl. (XIII.14a) auftaucht — gelten noch für beliebige elektrische Ladungs- und Stromverteilungen $\rho_{\text{el.}}$, $\vec{j}_{\text{el.}}$ in einem elektromagnetischen Feld.

Man prüft ähnlich wie in § XIII.1.2 a nach, dass die Wirkung (XIII.14) eichinvariant ist. An der Gl. (XIII.14b) erkennt man auch deren Invarianz unter Lorentz-Transformationen.

Das Integrand $J_{\text{el.}}(\mathbf{x}) \cdot \mathbf{A}(\mathbf{x})$ ist nämlich ein Lorentz-Skalar, und so ist das Integrationsmaß $d^4\mathbf{x}$.

XIII.2 Elektromagnetisches Feld mit Quellen

In diesem Abschnitt werden zuerst Elemente der klassischen Feldtheorie — Lagrange-Dichte, mit den zugehörigen Euler–Lagrange-Gleichungen, Noether-Theorem. . . — eingeführt (§ XIII.2.1). Diese Ideen werden dann auf das Beispiel des elektromagnetischen Feldes angewandt (§ XIII.2.2–XIII.2.3). Hiernach werden die relativistisch kovarianten Notationen durchaus benutzt.

XIII.2.1 Einführung in die klassische Feldtheorie

Es seien $\varphi_k(\mathbf{x})$, $k = 1, \dots, N$ die Komponenten in einem gegebenen Bezugssystem von einem Feld bzw. von einem System von Feldern auf dem Minkowski-Raum,⁽⁹²⁾ wobei N von der tensoriellen Natur des Feldes abhängt: $N = 1$ für ein einzelnes Skalarfeld, $N = 4$ für ein Vierervektorfeld, usw.

Die in diesem Paragraph eingeführten Begriffe und Ergebnisse stellen eine Verallgemeinerung jener des § III.2 auf den Fall einer unendlichen Anzahl von Freiheitsgraden dar.

XIII.2.1 a Definitionen

Diesem Feld wird ein Funktional der Komponenten $\varphi_k(\mathbf{x})$ und deren Ableitungen $\partial_\mu \varphi_k(\mathbf{x})$ zugeordnet, die *Lagrange-Dichte*

$$\mathcal{L}[\varphi_k(\mathbf{x}), \partial_\mu \varphi_k(\mathbf{x})]. \quad (\text{XIII.15a})$$

Dieses Funktional bestimmt die ganze Dynamik des Feldes. Die entsprechende Lagrange-Funktion ist durch

$$\mathcal{L} = \int \mathcal{L}[\varphi_k(\mathbf{x}), \partial_\mu \varphi_k(\mathbf{x})] d^3\vec{r}, \quad (\text{XIII.15b})$$

gegeben, wobei die Integration auf den ganzen dreidimensionalen Raum durchgeführt wird — wes-

⁽⁹²⁾Tatsächlich darf man auch Felder auf einem beliebigen Zeitraum betrachten.

halb \mathcal{L} als Dichte bezeichnet wird. Die resultierende Wirkung lautet

$$S = \int_{t_a}^{t_b} L dt = \int \mathcal{L}[\varphi_k(\mathbf{x}), \partial_\mu \varphi_k(\mathbf{x})] \frac{d^4x}{c} \quad (\text{XIII.15c})$$

mit d^4x dem Vierervolumenelement (X.21).

Bemerkungen:

* Ist die Lagrange-Dichte eines Systems von Feldern ein Lorentz-Skalar, so ist die resultierende Wirkung automatisch Lorentz-invariant.

* Zwei Lagrange-Dichten \mathcal{L} und \mathcal{L}' , die nur um eine Viererdivergenz $\partial \cdot \mathcal{M}(\mathbf{x}) = \partial_\mu \mathcal{M}^\mu(\mathbf{x})$ abweichen, wobei \mathcal{M} ein Vierervektorfeld bezeichnet, führen jeweils zu Wirkungen S und S' , die dieselben Bewegungsgleichungen liefern. Daher sind solche Lagrange-Dichten äquivalent: die Lagrange-Dichte eines gegebenen Systems von Feldern ist nicht eindeutig.

XIII.2.1 b Hamilton-Prinzip. Euler-Lagrange-Gleichungen

Wie im § III.2.2 gilt das *Hamilton-Prinzip*, laut welchem die physikalisch realisierte „Konfiguration“ der Felder solche ist, welche die Wirkung extremal macht — wobei die Konfiguration in der Raumzeit zu betrachten ist, d.h. sie entspricht der Zeitentwicklung einer räumlichen Anordnung.

Betrachtet man die Variation der Wirkung (XIII.15c) für eine Variation $\delta\varphi_k$ des Feldes, so lautet sie

$$\delta S = \int \left[\frac{\partial \mathcal{L}}{\partial \varphi_k} \delta \varphi_k + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi_k)} \delta \partial_\mu \varphi_k \right] \frac{d^4x}{c},$$

wobei über doppelt auftretende Indizes $k = 1 \dots N$ und μ summiert wird.

Da $\delta \partial_\mu \varphi_k = \partial_\mu \delta \varphi_k$, kann der zweite Term im Integranden durch partielle Integration berechnet werden. Der resultierende integrierte Term verschwindet⁽⁹³⁾ und es bleibt

$$\delta S = \int \left[\frac{\partial \mathcal{L}}{\partial \varphi_k} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi_k)} \right] \delta \varphi_k \frac{d^4x}{c}$$

übrig.

Die Wirkung ist extremal wenn $\delta S = 0$ für beliebige Variationen $\delta\varphi_k$, d.h. wenn das Feld und seine Ableitungen den Euler-Lagrange-Gleichungen

$$\frac{\partial \mathcal{L}[\varphi_k, \partial_\nu \varphi_k]}{\partial \varphi_k} = \partial_\mu \left[\frac{\partial \mathcal{L}[\varphi_k, \partial_\nu \varphi_k]}{\partial (\partial_\mu \varphi_k)} \right] \quad (\text{XIII.16})$$

genügen. Diese stellen die Bewegungsgleichungen des Feldes dar.

XIII.2.2 Elektromagnetisches Feld in Anwesenheit fester Quellen

XIII.2.2 a Lagrange-Dichte des elektromagnetischen Feldes

Der Vergleich der Gl. (XIII.14b) und (XIII.15c) zeigt, dass die Lagrange-Funktion für den Wechselwirkungsterm zwischen einem elektrischen Strom und dem elektromagnetischen Feld als Integral der Lagrange-Dichte

$$\mathcal{L}_{M+F} \equiv J_{\text{el.}}^\nu(\mathbf{x}) A_\nu(\mathbf{x}) = \mathbf{J}_{\text{el.}}(\mathbf{x}) \cdot \mathbf{A}(\mathbf{x}) \quad (\text{XIII.17})$$

betrachtet werden kann.

⁽⁹³⁾An den „Endpunkten“, entsprechend in der Raumzeit einem Oberflächenterm, werden die Felder nicht variiert.

In dieser Lagrange-Dichte tritt das elektromagnetische Feld in der Form des Viererpotentials eher als des Feldstärketensors auf: die relevanten Freiheitsgrade sind also die Komponenten $A_\nu(\mathbf{x})$. Dazu kann die Lagrange-Dichte auch Funktion der Ableitungen $\partial_\mu A_\nu(\mathbf{x})$ sein.

Die „Standard“-Lagrange-Dichte für das freie elektromagnetische Feld ist

$$\mathcal{L}_F[A_\nu, \partial_\mu A_\nu] = -\frac{1}{4\mu_0} F_{\mu\nu}(\mathbf{x}) F^{\mu\nu}(\mathbf{x}) = \frac{\epsilon_0}{2} \left[\vec{E}(t, \vec{r})^2 - c^2 \vec{B}(t, \vec{r})^2 \right] \quad (\text{XIII.18})$$

mit $F_{\mu\nu}(\mathbf{x})$ den kovarianten Komponenten des elektromagnetischen Feldstärketensors (XII.1).

Bemerkungen:

* Da die Lagrange-Dichte (XIII.18) des freien elektromagnetischen Feldes nur vom Feldstärketensor abhängt, ist sie wie dieser eichinvariant.

Im Gegensatz hängt die Lagrange-Dichte (XIII.17) für den Wechselwirkungsterm zwischen Feld und Punktladung von der Eichung ab. Wie oben schon bemerkt wurde ist die entsprechende Wirkung aber eichinvariant.

* Wie immer kann man andere Lagrange-Dichten für das freie elektromagnetische Feld postulieren, die zu den gleichen Bewegungsgleichungen führen. Daher wird die Wahl (XIII.18) oft als Standard bezeichnet.

XIII.2.2b Bewegungsgleichungen

Ein dynamisches elektromagnetisches Feld in Anwesenheit von festen Quellen lässt sich dann durch die Lagrange-Dichte $\mathcal{L}^{(2)}[A_\nu, \partial_\mu A_\nu] = \mathcal{L}_F[A_\nu, \partial_\mu A_\nu] + \mathcal{L}_{M+F}[A_\nu, \partial_\mu A_\nu]$ beschreiben.

Die entsprechenden Euler-Lagrange-Gleichungen folgen aus

$$\frac{\partial \mathcal{L}_{M+F}}{\partial A_\nu} = J_{\text{el.}}^\nu(\mathbf{x}), \quad \frac{\partial \mathcal{L}_{M+F}}{\partial(\partial_\mu A_\nu)} = 0, \quad \frac{\partial \mathcal{L}_F}{\partial A_\nu} = 0$$

und

$$\frac{\partial \mathcal{L}_F}{\partial(\partial_\mu A_\nu)} = -\frac{1}{4\mu_0} \frac{\partial}{\partial(\partial_\mu A_\nu)} [F_{\mu\nu}(\mathbf{x}) F^{\mu\nu}(\mathbf{x})] = -\frac{1}{2\mu_0} \frac{\partial}{\partial(\partial_\mu A_\nu)} [\partial_\mu A_\nu(\mathbf{x}) F^{\mu\nu}(\mathbf{x})] = -\frac{1}{\mu_0} F^{\mu\nu}(\mathbf{x}).$$

Laut Gl. (XIII.16) ergibt sich

$$J_{\text{el.}}^\nu(\mathbf{x}) = -\frac{1}{\mu_0} \partial_\mu F^{\mu\nu}(\mathbf{x}),$$

d.h. unter Berücksichtigung der Antisymmetrie des Feldstärketensors ($\partial_\mu F^{\mu\nu} = -\partial_\nu F^{\mu\nu}$)

$$\partial_\nu F^{\mu\nu}(\mathbf{x}) = \mu_0 J_{\text{el.}}^\mu(\mathbf{x}). \quad (\text{XIII.19})$$

Das heißt, man findet die inhomogenen Maxwell-Gleichungen (XII.13a) wieder, was die Wahl der Lagrange-Dichte (XIII.18) rechtfertigt.⁽⁹⁴⁾

Bemerkung: In der obigen Herleitung der Bewegungsgleichungen wurde angenommen, dass die Komponenten $A_\nu(\mathbf{x})$ mit $\nu = 0, 1, 2, 3$ unabhängig voneinander sind. Wegen der Eichinvarianz der Elektrodynamik ist dies aber nicht der Fall: das elektromagnetische Feld im Vakuum besitzt nicht vier Freiheitsgrade, sondern nur zwei — entsprechend den zwei möglichen linearen Polarisierungen. Die Eichinvarianz der Theorie kann später durchgesetzt werden, wie sich am folgenden Beispiel des Energieimpulstensors beobachten lässt.

XIII.2.3 Energieimpulstensor

Aus der Invarianz der Elektrodynamik unter einer beliebigen Translation $x^\mu \rightarrow x'^\mu = x^\mu + a^\mu$ in der Raumzeit folgt die Existenz einer Erhaltungsgröße, des Energieimpulstensors.

⁽⁹⁴⁾Wie schon in § XII.3.1 bemerkt folgen die homogenen Maxwell-Gleichungen (XII.13b) automatisch aus der Beziehung $F_{\mu\nu}(\mathbf{x}) = \partial_\mu A_\nu(\mathbf{x}) - \partial_\nu A_\mu(\mathbf{x})$, d.h. sie sind nicht dynamisch.

XIII.2.3 a Noether-Theorem

Betrachten wir wieder die allgemeine Lagrange-Dichte $\mathcal{L}[\varphi_k(\mathbf{x}), \partial_\mu \varphi_k(\mathbf{x})]$ des § XIII.2.1. Es wird angenommen, dass die resultierende Wirkung

$$S = \int_{\Omega} \mathcal{L}[\varphi_k(\mathbf{x}), \partial_\mu \varphi_k(\mathbf{x})] \frac{d^4x}{c} \quad (\text{XIII.20})$$

invariant ist unter den gleichzeitigen infinitesimalen Transformationen

$$x^\mu \rightarrow x'^\mu = x^\mu + \delta x^\mu, \quad (\text{XIII.21a})$$

$$\varphi_k(\mathbf{x}) \rightarrow \varphi'_k(\mathbf{x}) = \varphi_k(\mathbf{x}) + \tilde{\delta} \varphi_k(\mathbf{x}). \quad (\text{XIII.21b})$$

Unter diesen Transformationen ändert sich das vierdimensionale Integrationsvolumen in Gl. (XIII.20) gemäß $\Omega \rightarrow \Omega'$. Dann lautet die transformierte Wirkung (XIII.20)

$$S' = \int_{\Omega'} \mathcal{L}[\varphi'_k(\mathbf{x}'), \partial_\mu \varphi'_k(\mathbf{x}')] \frac{d^4x'}{c} = \int_{\Omega} \mathcal{L}[\varphi'_k(\mathbf{x}), \partial_\mu \varphi'_k(\mathbf{x})] \frac{d^4x}{c} + \oint_{\partial\Omega} \mathcal{L}[\varphi'_k(\mathbf{x}), \partial_\mu \varphi'_k(\mathbf{x})] \delta x^\mu \frac{d^3\sigma_\mu}{c},$$

mit $\partial\Omega$ der dreidimensionalen Oberfläche („Hyperfläche“) des Integrationsvolumens Ω und $d^3\sigma_\mu$ einem Hyperflächenelement.

Die Taylor-Entwicklung zur ersten Ordnung der rechten Seite dieser Gleichung liefert die Variation der Wirkung

$$\begin{aligned} S' - S &= \int_{\Omega} \left[\frac{\partial \mathcal{L}}{\partial \varphi_k} \tilde{\delta} \varphi_k + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi_k)} \partial_\mu \tilde{\delta} \varphi_k \right] \frac{d^4x}{c} + \oint_{\partial\Omega} \mathcal{L} \delta x^\mu \frac{d^3\sigma_\mu}{c} \\ &= \int_{\Omega} \left[\frac{\partial \mathcal{L}}{\partial \varphi_k} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi_k)} \right) \right] \tilde{\delta} \varphi_k \frac{d^4x}{c} + \oint_{\partial\Omega} \left[\frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi_k)} \tilde{\delta} \varphi_k + \mathcal{L} \delta x^\mu \right] \frac{d^3\sigma_\mu}{c}, \end{aligned}$$

wobei der Übergang von der ersten zur zweiten Zeile einer partiellen Integration entspricht. In allen Termen ist das nicht-geschriebene Argument der Lagrange-Dichte $[\varphi_k(\mathbf{x}), \partial_\mu \varphi_k(\mathbf{x})]$.

Wegen der Euler-Lagrange-Gleichungen (XIII.16) ist das Volumenintegral in der zweiten Zeile gleich Null. Somit ist die Wirkung nur dann invariant unter der Transformation (XIII.21), wenn das Oberflächenintegral ebenfalls verschwindet. Im Letzteren kann $\tilde{\delta} \varphi_k(\mathbf{x})$ durch die Variation des Feldes

$$\delta \varphi_k \equiv \varphi'_k(\mathbf{x}') - \varphi_k(\mathbf{x}) \simeq \tilde{\delta} \varphi_k(\mathbf{x}) + \delta x_\nu \partial^\nu \varphi_k(\mathbf{x}) \quad (\text{XIII.22})$$

ausgedrückt werden. Dann gilt

$$\begin{aligned} \oint_{\partial\Omega} \left[\frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi_k)} (\delta \varphi_k - \delta x_\nu \partial^\nu \varphi_k) + \mathcal{L} \delta x^\mu \right] \frac{d^3\sigma_\mu}{c} = \\ \oint_{\partial\Omega} \left[\frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi_k)} \delta \varphi_k + \left(\eta^{\mu\nu} \mathcal{L} - \frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi_k)} \partial^\nu \varphi_k \right) \delta x_\nu \right] \frac{d^3\sigma_\mu}{c} = 0. \end{aligned}$$

Man definiert den (kanonischen) *Energieimpulstensor* $\mathbf{T}_{\text{kan.}}$ durch seine Komponenten

$$T_{\text{kan.}}^{\mu\nu} \equiv \eta^{\mu\nu} \mathcal{L} - \frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi_k)} \partial^\nu \varphi_k \quad (\text{XIII.23a})$$

sowie den *Noether-Strom* \mathbf{N} mit Komponenten

$$N^\mu \equiv \frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi_k)} \delta \varphi_k + T_{\text{kan.}}^{\mu\nu} \delta x_\nu, \quad (\text{XIII.23b})$$

wobei es sich eigentlich um eine Stromdichte handelt. Dann liefert der Satz von Stokes

$$\oint_{\partial\Omega} N^\mu(\mathbf{x}) d^3\sigma_\mu = 0 = \int_{\Omega} \partial_\mu N^\mu(\mathbf{x}) d^4x. \quad (\text{XIII.23c})$$

Sei Ω das Volumen zwischen den Hyperflächen $t_1 = \text{Konstante}$ und $t_2 = \text{Konstante}$, entsprechend dem zeitartigen Hyperflächenelement $d^3\sigma_\mu = (-d^3\vec{r}, \vec{0})$. Wenn die Felder φ_k im räumlichen Unendlichen verschwinden, dann ist $\int N^0(x) d^3\vec{r}$ eine Konstante der Bewegung.

Dies gilt insbesondere für jede der vier Komponenten

$$P^\nu \equiv \int_t T_{\text{kan.}}^{0\nu}(x) d^3\vec{r} \quad \text{für } \nu = 0, 1, 2, 3,$$

die sich aus dem Einsetzen von $\delta\varphi_k = 0$ und $\delta x_\nu = a_\nu$ — entsprechend der Invarianz der Lagrange-Dichte unter Translationen in der Raumzeit — in den Ausdruck (XIII.23b) des Noether-Stroms ergeben.

Die Erhaltung der vier Größen P^ν stellt einen Sonderfall des Noether-Theorems dar, laut dem zu jeder kontinuierlichen Gruppe von Transformationen der Felder und Koordinaten, welche die Wirkung invariant lassen, eine Erhaltungsgröße zugeordnet werden kann.

Bemerkungen:

* Die Invarianz der Wirkung unter Translationen der Raumzeit ist nur Teil der nötigen Invarianz relativistischer Theorien unter allen Elementen [Gl. (X.24)] der Poincaré-Gruppe. Aus der Invarianz unter Lorentz-Transformationen der Koordinaten folgt die Erhaltung des Tensors dritter Stufe mit Komponenten $\mathcal{M}^{\mu\nu\rho} = x^\nu T_{\text{kan.}}^{\mu\rho} - x^\rho T_{\text{kan.}}^{\mu\nu}$ — entsprechend einem „verallgemeinerten Drehimpuls“ —, mit $T_{\text{kan.}}^{\mu\nu}$ dem kanonischen Energieimpulstensor (XIII.23a).

* In manchen Büchern wird statt der Invarianz der Wirkung unter den Transformationen (XIII.21) die strengere Invarianz der Lagrange-Dichte erfordert. Diese Bedingung ist aber zu beschränkend und gilt beispielsweise nicht für die Lagrange-Dichte einer Punktladung in einem elektromagnetischen Feld! Für eine Diskussion s. Lévy-Leblond [34].

* In Analogie zur Mechanik einer endlichen Zahl von Freiheitsgraden ordnet man dem Feld φ_k mit der Lagrange-Dichte \mathcal{L} einen kanonisch konjugierten Impuls

$$\pi_k \equiv \frac{\partial \mathcal{L}}{\partial(\partial_0 \varphi_k)} \quad \text{(XIII.24)}$$

zu. Damit ergibt sich die *Hamilton-Dichte*

$$\mathcal{H}[\pi_k, \varphi_k] = \pi_k \partial_0 \varphi_k - \mathcal{L}, \quad \text{(XIII.25)}$$

dessen Integral über den Raum die Hamilton-Funktion liefert. Aus Gl. (XIII.23a) mit $\eta^{00} = -1$ und $\partial^0 \varphi_k = -\partial_0 \varphi_k$ folgt die Identität der Hamilton-Dichte mit der 00-Komponente des kanonischen Energieimpulstensors, $\mathcal{H} = T_{\text{kan.}}^{00}$.

XIII.2.3 b Elektromagnetischer Energieimpulstensor

Im Vakuum, d.h. in Abwesenheit von Quellen, ist die Wirkung des (freien) elektromagnetischen Feldes invariant unter Raumzeittranslationen. Aus Gl. (XIII.23a) und der Standard-Lagrange-Dichte (XIII.18) des freien elektromagnetischen Feldes ergibt sich der *kanonische* Energieimpulstensor

$$T_{\text{kan.}}^{\mu\nu} = \eta^{\mu\nu} \mathcal{L}_F - \frac{\partial \mathcal{L}_F}{\partial(\partial_\mu A_\rho)} \partial^\nu A_\rho = \eta^{\mu\nu} \mathcal{L}_F + \frac{1}{\mu_0} F^{\mu\rho} \partial^\nu A_\rho, \quad \text{(XIII.26)}$$

der laut den Ergebnissen des vorigen Paragraphen „erhalten“ sein soll:⁽⁹⁵⁾ damit Gl. (XIII.23c) für ein beliebiges Ω erfüllt ist, soll

$$\partial_\mu T_{\text{kan.}}^{\mu\nu}(x) = 0 \quad \text{(XIII.27)}$$

in jedem Punkt x und für jedes $\nu = 0, 1, 2, 3$ gelten.

⁽⁹⁵⁾Eigentlich ist die zugehörige Noether-Ladung erhalten.

Der Tensor (XIII.26) ist aber nicht eichinvariant! Ersetzt man A^μ durch $A'^\mu = A^\mu + \partial^\mu \chi$, so transformiert sich $T_{\text{kan.}}^{\mu\nu}$ in $T_{\text{kan.}}'^{\mu\nu} = T_{\text{kan.}}^{\mu\nu} + (1/\mu_0)F^{\mu\rho}\partial^\nu\partial_\rho\chi$.

Allgemein existieren neben dem kanonischen Tensor (XIII.23a) weitere Energieimpulstensenoren, die ebenfalls erhalten sind, und zwar der Art

$$T^{\mu\nu}(\mathbf{x}) = T_{\text{kan.}}^{\mu\nu}(\mathbf{x}) + \partial_\rho K^{\mu\rho\nu}(\mathbf{x}),$$

wobei $K^{\mu\rho\nu}$ antisymmetrisch in μ und σ ist: $K^{\mu\rho\nu} = -K^{\rho\mu\nu}$. Man überprüft einfach

$$\partial_\mu T^{\mu\nu}(\mathbf{x}) = \partial_\mu T_{\text{kan.}}^{\mu\nu}(\mathbf{x}) + \partial_\mu \partial_\rho K^{\mu\rho\nu}(\mathbf{x}) = 0,$$

wobei der zweite Term wegen der Symmetrie von $\partial_\mu \partial_\rho$ und der Antisymmetrie von $K^{\mu\rho\nu}$ Null ist. Für das elektromagnetische Feld liefert die Wahl $K^{\mu\rho\nu} = -(1/\mu_0)F^{\mu\rho}A^\nu$ einen eichinvarianten Energieimpulstensor

$$\begin{aligned} T_{\text{e.m.}}^{\mu\nu}(\mathbf{x}) &= \eta^{\mu\nu} \mathcal{L}_F + \frac{1}{\mu_0} \left[F^{\mu\rho}(\mathbf{x}) \partial^\nu A_\rho(\mathbf{x}) - \partial_\rho (F^{\mu\rho}(\mathbf{x}) A^\nu(\mathbf{x})) \right] \\ &= \eta^{\mu\nu} \mathcal{L}_F + \frac{1}{\mu_0} \left[F^{\mu\rho}(\mathbf{x}) \partial^\nu A_\rho(\mathbf{x}) - F^{\mu\rho}(\mathbf{x}) \partial_\rho A^\nu(\mathbf{x}) \right], \end{aligned}$$

wobei die Maxwell-Gleichung $\partial_\rho F^{\mu\rho} = 0$ in der letzten Gleichung benutzt wurde. Mit dem Ausdruck (XIII.18) der Lagrange-Dichte lässt sich dieser Tensor noch als

$$T_{\text{e.m.}}^{\mu\nu}(\mathbf{x}) = -\frac{1}{\mu_0} \left[F^{\mu\rho}(\mathbf{x}) F_{\rho}{}^{\nu}(\mathbf{x}) + \frac{1}{4} \eta^{\mu\nu} F_{\rho\sigma}(\mathbf{x}) F^{\rho\sigma}(\mathbf{x}) \right] \quad (\text{XIII.28a})$$

umschreiben. Somit findet man Gl. (XII.12b) wieder, und die Gleichung

$$\partial_\mu T_{\text{e.m.}}^{\mu\nu}(\mathbf{x}) = 0 \quad (\text{XIII.28b})$$

stellt einen Spezialfall der Bilanzgleichung (XII.18) mit $\mathbf{J}_{\text{el.}}(\mathbf{x}) = 0$ dar.

Literatur zum Kapitel XIII

- Fließbach, *Elektrodynamik* [3] Teil IV Kap. 19.
- Greiner, *Klassische Elektrodynamik* [8] Teil IV Kap. 23.
- Jackson, *Klassische Elektrodynamik* [11] = *Classical Electrodynamics* [12] Kap. 12.1, 12.7, 12.10.
- Landau & Lifschitz, *Klassische Feldtheorie* [14] = *The classical theory of fields* [28] Kap. 4 § 27–33.
- Nolting, *Spezielle Relativitätstheorie, Thermodynamik* [32] Erster Teil, Kap. 2.4.
- Scheck, *Klassische Feldtheorie* [19] Kap. 3.

Anhang

ANHANG A

Tensoren auf einem Vektorraum

A.1 Vektoren, Linearformen und Tensoren 251

A.1.1 Vektoren 251

A.1.2 Linearformen 252

A.1.3 Tensoren 253

A.1.4 Metrischer Tensor 255

A.1.5 Warum Tensoren? 256

A.2 Basistransformation 256

In diesem Anhang werden einige Definitionen und Ergebnisse betreffend Tensoren ohne Anspruch auf mathematische Strenge zusammengestellt. Das Ziel ist, den „modernen“ geometrischen Gesichtspunkt einzuführen. Somit werden Tensoren über ihre Operation auf Vektoren oder Linearformen definiert, d.h. auf koordinatenunabhängige Weise (Abschn. A.1), im Gegensatz zur „alten“ Charakterisierung durch ihr Verhalten unter Basistransformationen (Abschn. A.2).

Dabei wird von der Leserin erwartet, dass sie schon einige Grundkenntnisse in Linearalgebra hat, wie z.B. was Linearität, Linearkombinationen, (multi)lineare Abbildungen, Vektoren, Matrizen... sind. Auf ähnlicher Weise werden Begriffe wie Gruppe, Körper, Abbildung/Funktion... ohne weitere Erklärung benutzt.

Dem durchschnittlichen Physik-Studierenden werden die in Abschn. A.1 eingeführten Begriffe vielleicht auf erster Sicht unnötig abstrakt erscheinen, während die Formeln des Abschn. A.2 mehr praktischer Natur sind. Dem Autor scheint⁽⁹⁶⁾ die hier adoptierte Darstellung die natürlichste zu sein, welche die Existenz von Tensoren „mit Indizes oben oder unten oder sogar beides“ irgendwie erklärt. Und auch wenn der Unterschied anscheinend unwichtig ist für die Physik im drei-dimensionalen euklidischen Raum der nicht-relativistischen klassischen Mechanik — solange man in kartesischen Koordinaten arbeitet —, wird er unvermeidbar im relativistischen Kontext, oder wenn man krummlinige Koordinaten benutzt.

In den Physik-Kapiteln dieses Skripts werden in der Tat nur Tensoren auf *reellen* Vektorräumen betrachtet, d.h. der Körper \mathbb{K} der Skalare die Menge \mathbb{R} ist; in diesem Anhang wird dies nicht angenommen. Die Einstein'sche Summenkonvention über doppelt auftretende Indizes wird durchgehend verwendet.

A.1 Vektoren, Linearformen und Tensoren

A.1.1 Vektoren

... sind definitionsgemäß die Elemente \vec{c} eines *Vektorraums* \mathcal{V} , d.h. einer Menge mit 1) einer inneren zweistelligen Operation („Vektoraddition“), für welche die Menge eine Abel'sche^(az) Gruppe⁽⁹⁷⁾

⁽⁹⁶⁾erstaunlicherweise!

⁽⁹⁷⁾Eine Gruppe ist eine Menge mit einer assoziativen zweistelligen inneren Operation, für die es ein Neutralement

^(az)N. ABEL, 1802–1829

bildet, und 2) einer äußeren Verknüpfung („Skalarmultiplikation“) mit „Skalaren“ — den Elementen eines Körpers \mathbb{K} —, die assoziativ ist, eine Einheit besitzt, und distributiv bezüglich der Additionen in \mathcal{V} und in \mathbb{K} ist.

Sei $\mathcal{B} = \{\vec{e}_i\}$ eine *Basis*, d.h. eine Teilmenge von linear unabhängigen Vektoren⁽⁹⁸⁾, welche auch ein Erzeugendensystem⁽⁹⁹⁾ des gesamten Vektorraums \mathcal{V} bilden. Jedem Vektor \vec{c} von \mathcal{V} werden dann Zahlen $\{c^i\}$ des Körpers \mathbb{K} eindeutig zugeordnet, seine *Koordinaten* bezüglich der Basis, die derart sind, dass

$$\vec{c} = c^i \vec{e}_i. \quad (\text{A.1})$$

Ist die Anzahl der Vektoren einer Basis endlich — was dann der Fall aller Basen des Vektorraums ist — und gleich einer ganzen Zahl D — welche die gleiche für alle Basen ist —, so wird der Vektorraum \mathcal{V} endlich-dimensional genannt und D ist seine *Dimension* (über \mathbb{K}): $D = \dim \mathcal{V}$. Wir werden ab jetzt annehmen, dass die betrachteten Vektorräume endlich-dimensional sind.

A.1.2 Linearformen

... auf einem Vektorraum \mathcal{V} sind die *linearen* Abbildungen von \mathcal{V} in den zugrundeliegenden Körper \mathbb{K} von Skalaren. Hiernach werden diese Linearformen mit \underline{h} bezeichnet.

Die Menge der Linearformen auf \mathcal{V} , mit den „natürlichen“ Addition und Skalarmultiplikation, ist selbst ein Vektorraum über \mathbb{K} . Dieser wird mit \mathcal{V}^* bezeichnet und heißt *Dualraum* zu \mathcal{V} .

Wenn \mathcal{V} endlich-dimensional ist, dann ist es auch der Fall von \mathcal{V}^* und $\dim \mathcal{V}^* = \dim \mathcal{V}$. Gegeben eine Basis $\mathcal{B} = \{\vec{e}_i\}$ von \mathcal{V} , so kann man die *duale Basis* $\mathcal{B}^* = \{\underline{\epsilon}^j\}$ auf \mathcal{V}^* derart definieren, dass

$$\underline{\epsilon}^j(\vec{e}_i) = \delta_i^j, \quad (\text{A.2})$$

wobei δ_i^j das übliche Kronecker-Symbol bezeichnet.

Die Komponenten einer Linearform \underline{h} bezüglich einer gegebenen Basis werden im Folgenden mit $\{h_j\}$ bezeichnet:

$$\underline{h} = h_j \underline{\epsilon}^j. \quad (\text{A.3})$$

Bemerkungen:

* Die Wahl der Notationen, insbesondere die Position der Indizes, ist nicht unbedeutend! Somit lässt sich trivial nachprüfen, dass wenn $\{\underline{\epsilon}^j\}$ die duale Basis zu $\{\vec{e}_i\}$ bezeichnet, gelten

$$c^i = \underline{\epsilon}^i(\vec{c}) \quad \text{und} \quad h_j = \underline{h}(\vec{e}_j). \quad (\text{A.4})$$

* Oft werden die Vektoren von \mathcal{V} bzw. die Linearformen von \mathcal{V}^* „kontravariante Vektoren“ bzw. „kovariante Vektoren“ oder „Kovektoren“ genannt, und dementsprechend ihre Komponenten als „kontravariante“ bzw. „kovariante“ Koordinaten bezeichnet.

Die letzteren zwei, betreffend den Komponenten, sind zwar nützliche kurze Bezeichnungen, insbesondere wenn sie sich auf Tensoren beziehen (vgl. unten). Sie sind aber gleichzeitig verwirrend, denn sie sind nicht unterschiedliche Koordinaten einer einzigen mathematischen Größe, sondern Komponenten von unterschiedlichen Objekten, zwischen denen ein „natürlicher“ Zusammenhang eingeführt wurde, insbesondere mithilfe eines metrischen Tensors wie in § A.1.4.

* Werden die Vektoren von \mathcal{V} in Matrixdarstellung mit Spaltenvektoren bezeichnet, so lassen sich die Linearformen von \mathcal{V}^* durch Zeilenvektoren darstellen.

gibt, und derart, dass jedes Gruppenelement ein inverses Element besitzt. Die Gruppe heißt *kommutativ* oder *abel'sch*, wenn die Operation auch kommutativ ist

⁽⁹⁸⁾ ... d.h. keiner \vec{e}_{i_0} kann als Linearkombination der anderen Basisvektoren \vec{e}_i mit $i \neq i_0$ ausgedrückt werden.

⁽⁹⁹⁾ ... d.h. jeder Vektor von \mathcal{V} kann als Linearkombination der Basisvektoren geschrieben werden.

A.1.3 Tensoren

A.1.3 a Erste Definitionen und Resultate

Sei \mathcal{V} ein Vektorraum über dem Körper \mathbb{K} , und m, n zwei nicht-negative ganze Zahlen.

Definition: Man nennt *Tensoren vom Typ* $\binom{m}{n}$ auf \mathcal{V} die multilinearen Abbildungen, die m Linearformen — Elemente von \mathcal{V}^* — und n Vektoren — Elemente von \mathcal{V} — auf Skalare des Körpers \mathbb{K} abbilden, wobei Linearität bezüglich jedes Arguments gelten soll. Die ganze Zahl $m + n$ heißt *Stufe* (oder manchmal *Rang*, was aber relativ inkorrekt ist) des Tensors.

Schon bekannte Objekte sind Sonderfälle dieser Definition mit entweder m oder n gleich Null:

- die $\binom{0}{0}$ -Tensoren sind einfach die Skalare des zugrundeliegenden Körpers \mathbb{K} ;
- die $\binom{1}{0}$ -Tensoren stimmen mit den Vektoren überein;
Genauer sind sie die Elemente des *Bidualraums* \mathcal{V}^{**} , die sich im endlich-dimensionalen Fall mit den Elementen von \mathcal{V} dank einem „natürlichen“ Isomorphismus identifizieren lassen.
- die Tensoren vom Typ $\binom{0}{1}$ sind die Linearformen. Allgemeiner werden die $\binom{0}{n}$ -Tensoren auch *n-Multilinearformen* genannt.
- Schließlich werden Tensoren vom Typ $\binom{2}{0}$ manchmal — in der älteren Literatur — “Bivektoren” or “Dyaden” genannt.

In diesem Skript werden Tensoren allgemein mit \mathbf{T} bezeichnet, unabhängig von ihrer Stufe, so lange die letztere weder 0 noch 1 ist.

Bemerkungen:

- * Tensoren vom Typ $\binom{m}{n}$ werden oft als „ m -fach kontravariant und n -fach kovariant“ bezeichnet.
- * Falls sowohl m als n ungleich Null sind, spricht man von einem *gemischten Tensor*.
- * Sei $\mathbf{T} : (\mathcal{V}^*)^m \times (\mathcal{V})^n \rightarrow \mathbb{K}$ ein Tensor vom Typ $\binom{m}{n}$ und $m' \leq m, n' \leq n$ zwei nicht-negative ganze Zahlen. Für jedes m' -Tupel von Linearformen $\{\underline{h}_i\}$ und n' -Tupel von Vektoren $\{\vec{c}_j\}$ — und für zugehörige Tupeln von Positionen der Argumente, obwohl wir hier der Einfachheit halber die ersten Positionen annehmen — betrachtet man das Objekt

$$\mathbf{T}(\underline{h}_1, \dots, \underline{h}_{m'}, \cdot, \dots, \cdot; \vec{c}_1, \dots, \vec{c}_{n'}, \cdot, \dots, \cdot),$$

wobei die Punkte „leere“ Argumente bezeichnen. Dieses Objekt bildet $m - m'$ Linearformen und $n - n'$ Vektoren auf ein Skalar ab. Anders gesagt induziert der Tensor \mathbf{T} eine multilineare Abbildung⁽¹⁰⁰⁾ von $(\mathcal{V}^*)^{m'} \times (\mathcal{V})^{n'}$ in die Menge der Tensoren vom Typ $\binom{m-m'}{n-n'}$.

Beispielsweise hängen die $\binom{1}{1}$ -Tensoren natürlich mit den linearen Abbildungen von \mathcal{V} nach \mathcal{V} zusammen, d.h. wiederum mit den quadratischen Matrizen der Ordnung $\dim \mathcal{V}$. Aus diesem Grund werden solche Tensoren oft als Matrizen dargestellt.

Ein Tensor kann symmetrisch oder antisymmetrisch unter dem Austausch von zwei seiner Argumente sein, wobei diese entweder beide Vektoren oder beide Linearformen sind. Nach einer direkten Verallgemeinerung kann ein Tensor *total symmetrisch* — wie z.B. der metrische Tensor, den wir unten weiter diskutieren werden — oder *total antisymmetrisch* sein. Ein Beispiel für die letztere Möglichkeit ist die Determinante (von D D -dimensionalen Vektoren), welche die einzige (bis auf einem multiplikativen Faktor) total antisymmetrische D -Multilinearform auf einem Vektorraum von Dimension D ist.

⁽¹⁰⁰⁾ Genauer, so viele solche Abbildungen, wie es unabhängige Kombinationen — unter Betrachtung der möglichen Symmetrien von \mathbf{T} — aus m' bzw. n' Linearform bzw. Vektor Argumenten gibt.

A.1.3 b Operationen auf Tensoren

Versuchen mit der Addition und der Skalarmultiplikation, die sich aus den entsprechenden Operationen auf den Körper \mathbb{K} induzieren lassen, bilden die Tensoren eines gegebenen Typs auf einem Vektorraum \mathcal{V} selbst einen Vektorraum über \mathbb{K} .

Neben diesen zwei Verknüpfungen lassen sich weitere Operationen auf Tensoren definieren, insbesondere das *äußere Produkt* oder *Tensorprodukt* — welches zu Tensoren höherer Stufe führt — und die *Kontraktion* oder *Tensorverjüngung*, welche die Stufe reduziert.

Es seien zunächst zwei Tensoren \mathbf{T} und \mathbf{T}' von den jeweiligen Typen $\binom{m}{n}$ und $\binom{m'}{n'}$. Ihr äußeres Produkt $\mathbf{T} \otimes \mathbf{T}'$ ist ein Tensor vom Typ $\binom{m+m'}{n+n'}$, der für jedes $(m+m')$ -Tupel $(\underline{h}_1, \dots, \underline{h}_m, \dots, \underline{h}_{m+m'})$ von Linearformen und jedes $(n+n')$ -Tupel $(\vec{c}_1, \dots, \vec{c}_n, \dots, \vec{c}_{n+n'})$ von Vektoren die Gleichung

$$\mathbf{T} \otimes \mathbf{T}'(\underline{h}_1, \dots, \underline{h}_{m+m'}; \vec{c}_1, \dots, \vec{c}_{n+n'}) = \mathbf{T}(\underline{h}_1, \dots, \underline{h}_m; \vec{c}_1, \dots, \vec{c}_n) \mathbf{T}'(\underline{h}_{m+1}, \dots, \underline{h}_{m+m'}; \vec{c}_{n+1}, \dots, \vec{c}_{n+n'})$$

erfüllt. Dieses äußere Produkt ist assoziativ und distributiv bezüglich der Addition von Tensoren.

Zum Beispiel ist das äußere Produkt zweier Linearformen $\underline{h}, \underline{h}'$ eine Bilinearform $\underline{h} \otimes \underline{h}'$ derart, dass für jedes Paar (\vec{c}, \vec{c}') von Vektoren $\underline{h} \otimes \underline{h}'(\vec{c}, \vec{c}') = \underline{h}(\vec{c}) \underline{h}'(\vec{c}')$ gilt. Wiederum gibt die Tensormultiplikation zweier Vektoren \vec{c}, \vec{c}' einen $\binom{2}{0}$ -Tensor $\vec{c} \otimes \vec{c}'$, der für jedes Paar $(\underline{h}, \underline{h}')$ von Linearformen $\vec{c} \otimes \vec{c}'(\underline{h}, \underline{h}') = \underline{h}(\vec{c}) \underline{h}'(\vec{c}')$ erfüllt.

Tensoren vom Typ $\binom{m}{n}$, die sich als Tensorprodukt aus m Vektoren und n Linearformen schreiben lassen, werden manchmal *elementare Tensoren* genannt.

Sei nun \mathbf{T} ein gemischter Tensor vom Typ $\binom{m}{n}$ mit positiven m und n und seien $1 \leq j \leq m$ und $1 \leq k \leq n$. In der einfachsten koordinatenunabhängigen Definition der Kontraktion bezüglich des j -ten Linearform-Arguments und des k -ten vektoriellen Arguments wird \mathbf{T} als Summe von elementaren Tensoren geschrieben. Wird dann in jedem Summanden die k -te Linearform auf den j -ten Vektor angewandt, was eine Zahl (ein Skalar) gibt, so erhält man eine Summe aus elementaren Tensoren vom Typ $\binom{m-1}{n-1}$, d.h. ein Tensor dieses Typs, welcher das Ergebnis der Verjüngung ist.

Beispiele von Kontraktionen werden hiernach in § A.1.4 nach der Einführung des metrischen Tensors angegeben.

A.1.3 c Komponenten von Tensoren

Sei $\{\vec{e}_i\}$ bzw. $\{\underline{\epsilon}^j\}$ eine Basis auf einem Vektorraum \mathcal{V} der Dimension D bzw. auf dessen Dualraum \mathcal{V}^* — prinzipiell sind die Basen nicht dual zueinander, in der Praxis wird das aber implizit fast immer angenommen — und seien m, n zwei nicht-negative ganze Zahlen.

Die D^{m+n} elementaren Tensoren $\{\vec{e}_{i_1} \otimes \dots \otimes \vec{e}_{i_m} \otimes \underline{\epsilon}^{j_1} \otimes \dots \otimes \underline{\epsilon}^{j_n}\}$, wobei jeder Index i_k oder j_k jeden Wert zwischen 1 und D annimmt, bilden eine Basis der Tensoren vom Typ $\binom{m}{n}$. Die Komponenten eines Tensors \mathbf{T} bezüglich dieser Basis werden hiernach mit $\{\mathbf{T}^{i_1 \dots i_m}_{j_1 \dots j_n}\}$ bezeichnet:

$$\mathbf{T} = \mathbf{T}^{i_1 \dots i_m}_{j_1 \dots j_n} \vec{e}_{i_1} \otimes \dots \otimes \vec{e}_{i_m} \otimes \underline{\epsilon}^{j_1} \otimes \dots \otimes \underline{\epsilon}^{j_n} \quad (\text{A.5a})$$

mit

$$\mathbf{T}^{i_1 \dots i_m}_{j_1 \dots j_n} = \mathbf{T}(\underline{\epsilon}^{i_1}, \dots, \underline{\epsilon}^{i_m}; \vec{e}_{j_1}, \dots, \vec{e}_{j_n}). \quad (\text{A.5b})$$

Die mögliche Symmetrie oder Antisymmetrie eines Tensors unter dem Austausch zweier Argumente spiegelt sich in der Symmetrie oder Antisymmetrie seiner Komponenten unter dem Austausch der entsprechenden Indizes wider.

Komponentenweise liefert die Kontraktion von \mathbf{T} bezüglich seines j -ten Linearform-Arguments und seines k -ten vektoriellen Arguments den Tensor mit Komponenten $\mathbf{T}^{\dots i_{j-1}, \ell, i_{j+1}, \dots}_{\dots j_{k-1}, \ell, j_{k+1}, \dots}$, mit Summe über den doppelt auftretenden Index ℓ .

A.1.4 Metrischer Tensor

Nicht-entartete⁽¹⁰¹⁾ symmetrische Bilinearformen spielen eine wichtige Rolle, denn sie erlauben die Einführung einer weiteren Struktur auf dem Vektorraum \mathcal{V} , indem sie ein *inneres Produkt* definieren.

hier genauer sein: Bilinearform auf reellen Vektorräumen, Sesquilinearform auf komplexen Vektorräumen.

Sei $\{\underline{e}^j\}$ eine Basis des Dualraums \mathcal{V}^* . Eine Bilinearform $\mathbf{g} = g_{ij} \underline{e}^i \otimes \underline{e}^j$ heißt *metrischer Tensor* auf \mathcal{V} , wenn sie symmetrisch ist — d.h. $\mathbf{g}(\vec{a}, \vec{b}) = \mathbf{g}(\vec{b}, \vec{a})$ für alle Vektoren \vec{a}, \vec{b} , oder äquivalent $g_{ij} = g_{ji}$ für alle i, j —, und wenn die quadratische Matrix mit Elementen g_{ij} regulär ist. Das Skalar $\mathbf{g}(\vec{a}, \vec{b})$ wird dann auch als $\vec{a} \cdot \vec{b}$ bezeichnet, was insbesondere

$$g_{ij} = \mathbf{g}(\vec{e}_i, \vec{e}_j) = \vec{e}_i \cdot \vec{e}_j \quad (\text{A.6})$$

ergibt, wobei $\{\vec{e}_i\}$ die Basis von \mathcal{V} dual zu $\{\underline{e}^j\}$ ist.

Da die $D \times D$ -Matrix mit Elementen g_{ij} regulär ist, kann sie invertiert werden. Die Elemente der inversen Matrix seien mit g^{ij} bezeichnet: dann gelten $g_{ij}g^{jk} = \delta_i^k$ und $g^{ij}g_{jk} = \delta_k^i$. Die D^2 Skalaren g^{ij} definieren einen Tensor $g^{ij} \vec{e}_i \otimes \vec{e}_j$ vom Typ $\binom{2}{0}$, der als *inverser metrischer Tensor* und mit \mathbf{g}^{-1} bezeichnet wird.

Unter Verwendung eines Satzes über symmetrische Matrizen ist die quadratische Matrix mit Elementen g_{ij} diagonalisierbar — d.h. man kann eine geeignete Basis $\{\vec{e}_i\}$ von \mathcal{V} finden, für welche $\mathbf{g}(\vec{e}_i, \vec{e}_j) = 0$ für alle $i \neq j$ gilt. Da \mathbf{g} nicht-entartet ist, sind die Eigenwerte nicht Null: auf Kosten der möglichen Multiplikation der Basisvektoren $\{\vec{e}_i\}$ mit einem Skalar kann man fordern, dass jede Zahl $\mathbf{g}(\vec{e}_i, \vec{e}_i)$ gleich entweder $+1$ oder -1 sei, was die sog. *kanonische Form*

$$g_{ij} = \text{diag}(-1, \dots, -1, 1, \dots, 1) \quad (\text{A.7})$$

der Matrixdarstellung der Komponenten des metrischen Tensors ergibt.

In dieser speziellen Basis stimmen die Komponenten g^{ij} von \mathbf{g}^{-1} mit den g_{ij} , was aber in einer beliebigen Basis nicht der Fall ist.

Rolle von \mathbf{g} als Abbildung zwischen Tensoren unterschiedlicher Typen

Sei $\vec{c} = c^i \vec{e}_i$ ein beliebiger Vektor von \mathcal{V} . In Übereinstimmung mit der dritten Bemerkung vom § A.1.3 a bildet das Objekt $\mathbf{g}(\vec{c}, \cdot)$ die Vektoren von \mathcal{V} auf den zugrundeliegenden Körper \mathbb{K} ab, d.h. es entspricht einer Linearform $\underline{c} = c_j \underline{e}^j$, mit Komponenten

$$c_j = \underline{c}(\vec{e}_j) = \mathbf{g}(\vec{c}, \vec{e}_j) = \mathbf{g}(c^i \vec{e}_i, \vec{e}_j) = c^i g_{ij}. \quad (\text{A.8a})$$

Das heißt, ein metrischer Tensor \mathbf{g} stellt eine Abbildung zwischen Vektoren und Linearformen dar. Umgekehrt bildet der inverse metrische Tensor \mathbf{g}^{-1} eine Linearformen auf einen Vektor, dessen Komponenten durch die Beziehungen

$$c^i = g^{ij} c_j \quad (\text{A.8b})$$

gegeben sind.

Die Zusammenhänge der Gl. (A.8a)–(A.8b) lassen sich auf Tensoren höherer Stufe verallgemeinern: der metrische Tensor und sein Inverser werden dann für das „Herunterstellen oder Hochstellen von Indizes“ benutzt, d.h. für Operationen, die einen Tensor vom Typ $\binom{m}{n}$ auf einen Tensor vom jeweiligen Typ $\binom{m \mp 1}{n \pm 1}$ abbilden.

Bemerkungen:

* Das Herunterstellen bzw. Hochstellen eines Index entspricht eigentlich dem sukzessiven Durchführen eines äußeren Produkts mit \mathbf{g} bzw. \mathbf{g}^{-1} gefolgt durch die Kontraktion zweier Indizes. Zum

⁽¹⁰¹⁾ Das heißt, für jeden Vektor $\vec{a} \neq \vec{0}$ kann man einen Vektor \vec{b} finden, der $\mathbf{g}(\vec{a}, \vec{b}) \neq 0$ erfüllt. Dies kann auch äquivalent als Bedingung über die Matrix mit Elementen g_{ij} formuliert werden, wie 4 Zeilen unten zu finden ist.

Beispiel

$$\vec{c} = c^i \vec{e}_i \quad \xrightarrow{\text{äußeres Produkt}} \quad \vec{c} \otimes \mathbf{g} = c^i g_{jk} \vec{e}_i \otimes \underline{\epsilon}^j \otimes \underline{\epsilon}^k \quad \xrightarrow{\text{Kontraktion}} \quad \underline{c} = c^i g_{ik} \underline{\epsilon}^k = c_k \underline{\epsilon}^k$$

wobei die Verjüngung über das erste und zweite Argumente von $\vec{c} \otimes \mathbf{g}$ operiert.

* Die Notation mithilfe eines Punkts des durch den metrischen Tensors induzierten Skalarprodukts von Vektoren wird oft auf Tensoren verallgemeinert, deren Kontraktion ebenfalls mit einem Punkt bezeichnet wird. Beispielsweise schreibt man für die Verjüngung einer Bilinearform \mathbf{T} und eines Vektors \vec{c}

$$\mathbf{T} \cdot \vec{c} = (\mathbf{T}_{ij} \underline{\epsilon}^i \otimes \underline{\epsilon}^j) \cdot (c^k \vec{e}_k) = \mathbf{T}_{ij} c^j \underline{\epsilon}^i, -1mm$$

wobei Gl. (A.2) benutzt wurde. Die Notation ist meistens nützlich wenn sie eindeutig ist, d.h. wenn der Tensor \mathbf{T} symmetrisch ist, so dass alle seine Indizes die gleiche Rolle spielen.

Auf ähnlicher Weise, für einen $\binom{2}{0}$ -Tensor \mathbf{T} und eine Bilinearform \mathbf{T}'

$$\mathbf{T} \cdot \mathbf{T}' = (\mathbf{T}^{ij} \vec{e}_i \otimes \vec{e}_j) \cdot (\mathbf{T}'_{kl} \underline{\epsilon}^k \otimes \underline{\epsilon}^l) = \mathbf{T}^{ij} \mathbf{T}'_{jl} \vec{e}_i \otimes \underline{\epsilon}^l.$$

Dies ist unterschiedlich von $\mathbf{T}' \cdot \mathbf{T}$ wenn die Tensoren nicht symmetrisch sind.

Die Leserin kann so gar in der Literatur die Notation

$$\mathbf{T} : \mathbf{T}' \equiv \mathbf{T}^{ij} \mathbf{T}'_{ji}$$

finden, entsprechend zwei sukzessiven Kontraktionen.

* Wichtige Rolle: Struktur eines metrischen Raums

A.1.5 Warum Tensoren?

To do! hier ein kurzes Essay über die Notwendigkeit von Tensoren in der Physik...

A.2 Basistransformation

Seien $\mathcal{B} = \{\vec{e}_i\}$ und $\mathcal{B}' = \{\vec{e}_{j'}\}$ zwei Basen des Vektorraums \mathcal{V} , und $\mathcal{B}^* = \{\underline{\epsilon}^i\}$, $\mathcal{B}'^* = \{\underline{\epsilon}^{j'}\}$ die zugehörigen Dualbasen auf \mathcal{V}^* . Die Basisvektoren von \mathcal{B}' lassen sich durch jene von \mathcal{B} mithilfe einer regulären Matrix Λ mit Elementen $\Lambda^i_{j'}$ ausdrücken, gemäß

$$\vec{e}_{j'} = \Lambda^i_{j'} \vec{e}_i. \quad (\text{A.9})$$

Bemerkung: Die Matrix Λ ist *nicht* die Matrixdarstellung eines Tensors, denn die zwei Indizes seiner Elemente beziehen sich auf zwei unterschiedliche Basen — was mit der Verwendung eines gestrichenen und eines ungestrichenen Index betont wird. Im Gegensatz sind beide Komponenten eines Tensors vom Typ $\binom{1}{1}$ bezüglich der „gleichen“ Basis. ⁽¹⁰²⁾

Die Elemente der inversen Matrix Λ^{-1} seien mit $\Lambda^{k'}_i$ bezeichnet, das heißt

$$\Lambda^{k'}_i \Lambda^i_{j'} = \delta^{k'}_{j'} \quad \text{und} \quad \Lambda^i_{k'} \Lambda^{k'}_j = \delta^i_j.$$

Man kann dann einfach nachprüfen, dass die Zahlen $\Lambda^{k'}_i$ die Basistransformation von \mathcal{B}^* nach \mathcal{B}'^* angeben, und zwar

$$\underline{\epsilon}^{j'} = \Lambda^{j'}_i \underline{\epsilon}^i. \quad (\text{A.10})$$

Dementsprechend wird jede „Vektor“-Komponente ⁽¹⁰³⁾ mit Λ^{-1} transformiert:

$$c^{j'} = \Lambda^{j'}_i c^i, \quad \mathbf{T}^{j'_1 \dots j'_m} = \Lambda^{j'_1}_{i_1} \dots \Lambda^{j'_m}_{i_m} \mathbf{T}^{i_1 \dots i_m}. \quad (\text{A.11})$$

⁽¹⁰²⁾ Genauer beziehen sie sich auf eine Basis und ihre Dualbasis.

⁽¹⁰³⁾ Hier ist die Bezeichnung *kontravariante Komponente* nützlich.

Wiederum transformiert sich jede „Linearform“-Komponente⁽¹⁰⁴⁾ mit Λ :

$$h_{j'} = \Lambda^i{}_{j'} h_i, \quad \mathbf{T}_{j'_1 \dots j'_n} = \Lambda^{i_1}{}_{j'_1} \cdots \Lambda^{i_n}{}_{j'_n} \mathbf{T}_{i_1 \dots i_n}. \quad (\text{A.12})$$

Somit lassen sich die Komponenten eines beliebigen Tensors in irgendeiner Basis erhalten, wenn man die Transformationen der (Basis-)Vektoren und Linearformen kennt.

To do! Wichtiges Beispiel von Basistransformation: Drehmatrix

Literatur zum Anhang A

- Ihr bevorzugtes Lehrbuch zur Linearalgebra.
- Eine ziemlich ausführliche jedoch noch einfache Darstellung mit einem Schwerpunkt auf die geometrischen Anwendungen der Linearalgebra ist in Postnikov, *Lectures in Geometry* [35]⁽¹⁰⁵⁾ zu finden, insbesondere in Lectures 1 (Anfang), 4–6 & 18.

⁽¹⁰⁴⁾ ... oder *kovariante Komponente*.

⁽¹⁰⁵⁾ Manchmal werden die verwendeten mathematischen Bezeichnungen, zumindest in der englischen Übersetzung, relativ nicht Standard, z.B. (linear, bilinear) „functional“ anstatt „form“ oder „conjugate“ (space, basis) statt „dual“.

ANHANG B

Drehungen

B.1 Isometrien im euklidischen Raum 258

B.2 Infinitesimale Drehungen 260

In diesem Anhang wird erstens der Zusammenhang zwischen Isometrien, insbesondere Drehungen, im euklidischen Raum und orthogonalen Matrizen, die den Gl. (B.3) genügen, präzisiert. Dann werden infinitesimale Drehungen diskutiert, wobei einige Begriffe und Ergebnisse der Gruppentheorie eingeführt und ohne Beweis verwendet werden.

Der Anhang bezieht sich auf Transformationen in einem dreidimensionalen Raum. Isometrien, orthogonale Abbildungen, Drehungen, usw. können aber auch in einem D -dimensionalen euklidischen Raum betrachtet werden. Die zugehörige Verallgemeinerung der Ergebnisse des Abschn. B.1 ist trivial, während jene des Abschn. B.2 mehr Arbeit erfordern.

B.1 Isometrien im euklidischen Raum

Definition: Eine lineare Abbildung des dreidimensionalen (affinen) euklidischen Punktraums \mathcal{E}_3 in sich selbst heißt *Isometrie*, wenn sie den Abstand zwischen zwei Punkten invariant lässt.

Darunter sind z.B. die Translationen sowie Transformationen mit mindestens einem Fixpunkt, wie die Drehungen oder die Punktspiegelungen.

Die Verkettung zweier Isometrien gibt wieder eine Isometrie. Da die Identitätstransformation in \mathcal{E}_3 selbst eine Isometrie ist, während jede Isometrie bijektiv ist — und somit eine inverse Transformation besitzt, die ebenfalls eine Isometrie ist —, bildet die Menge der Isometrien, versehen mit der Verkettung, eine Gruppe, die mit $\text{ISO}(3)$ bezeichnet wird.

Bemerkung: Statt von Isometrien spricht man auch oft von *Bewegungen* und dementsprechend von der *Bewegungsgruppe*.⁽¹⁰⁶⁾

Beschränkt man die Diskussion auf die Transformationen mit einem Fixpunkt, der als Ursprungspunkt eines kartesischen Koordinatensystems gewählt wird, so ist eine solche Isometrie äquivalent zu einer linearen Abbildung \mathcal{O} im euklidischen Vektorraum \mathbb{R}^3 der Form

$$\vec{x} \in \mathbb{R}^3 \mapsto \vec{x}' = \mathcal{O}(\vec{x}) \in \mathbb{R}^3, \quad (\text{B.1})$$

die das euklidische Skalarprodukt $\vec{x} \cdot \vec{x}$ invariant lässt, d.h. $\vec{x}' \cdot \vec{x}' = \vec{x} \cdot \vec{x}$. Diese Abbildung \mathcal{O} wird *orthogonale Transformation* genannt.

Da die orthogonalen Transformationen linear sind, können sie in Matrixform dargestellt werden. Gegeben sei ein System von kartesischen Koordinaten. Ordnet man den Vektoren \vec{x}, \vec{x}' von \mathbb{R}^3 bzw. der orthogonalen Abbildung \mathcal{O} die reellen dreikomponentigen Spaltenvektoren ihrer Komponenten

⁽¹⁰⁶⁾ Im Rahmen einer Mechanik-Vorlesung könnten diese Bezeichnungen aber verwirrend sein!

bzw. eine 3×3 -Matrix zu, die ebenfalls mit \vec{x} , \vec{x}' bzw. \mathcal{O} bezeichnet werden, so lässt sich Gl. (B.1) als

$$\vec{x} \mapsto \vec{x}' = \mathcal{O}\vec{x} \quad (\text{B.2a})$$

umschreiben. Unter Verwendung der jeweiligen Koordinaten x^i , x'^i mit $i = 1, 2, 3$ der Vektoren und der Matrixelemente \mathcal{O}^i_j mit $i, j = 1, 2, 3$ der Abbildungsmatrix \mathcal{O} lautet dies auch

$$x'^i \mapsto x^i = \mathcal{O}^i_j x^j \quad \text{für } i = 1, 2, 3 \quad (\text{B.2b})$$

wobei die einsteinsche Summenkonvention benutzt wurde.

In Matrixdarstellung lautet das Skalarprodukt $\vec{x}'^T \vec{x}$, mit \vec{x}'^T dem zu \vec{x}' transponierten Zeilenvektor. Damit die Matrix \mathcal{O} eine orthogonale Transformation darstellt, muss für jeden $\vec{x} \in \mathbb{R}^3$

$$\vec{x}'^T \vec{x} = \vec{x}'^T \mathcal{O}\vec{x} = (\mathcal{O}\vec{x})^T \mathcal{O}\vec{x} = \vec{x}^T \mathcal{O}^T \mathcal{O}\vec{x}$$

gelten, d.h. die Matrix muss der Eigenschaft

$$\mathcal{O}^T \mathcal{O} = \mathbb{1}_3 \quad (\text{B.3a})$$

genügen, mit $\mathbb{1}_3$ der 3×3 -Einheitsmatrix. Alternativ kann man

$$\mathcal{O}^T = \mathcal{O}^{-1} \quad (\text{B.3b})$$

schreiben. Umgekehrt stellt jede Matrix \mathcal{O} , die diese Gleichung erfüllt, eine orthogonale Transformation des dreidimensionalen Raums dar.

Eine solche Matrix wird *orthogonale 3×3 -Matrix* genannt. Diese Matrizen bilden eine Gruppe, die sog. *orthogonale Gruppe* $O(3)$.

Bemerkungen:

* Dank der Beschränkung auf Isometrien von \mathcal{E}_3 mit einem festen Fixpunkt ist die Korrespondenz zwischen solchen Isometrien und den Matrizen von $O(3)$ bijektiv.

Mathematisch gesagt ist diese Korrespondenz eine sog. *lineare Darstellung* der Isometrien mit einem Fixpunkt — d.h. ein Homomorphismus von der Gruppe solcher Isometrien in die Gruppe $GL(\mathcal{V})$ der Automorphismen (d.h. der bijektiven linearen Abbildungen) eines Vektorraums \mathcal{V} , des *Darstellungsraums*, in sich selbst. Da die $O(3)$ -Matrizen den Automorphismen eines dreidimensionalen Vektorraums (\mathbb{R}^3) entsprechen, ist die Darstellung „vom Grad 3“. Diese wird noch als *unitär* bezeichnet, denn die Matrizen der Darstellung sind unitär — sie sind reell und erfüllen Gl. (B.3). Schließlich ist die Darstellung *irreduzibel*, weil es keinen nicht-trivialen Unterraum des Darstellungsraums gibt, der invariant unter der Wirkung der ganzen Gruppe $O(3)$ ist.

* In diesem Anhang werden *aktive* Transformationen der Vektoren betrachtet, entsprechend der eigentlichen Operation (Drehen...) auf ein Objekt, nicht der Änderung des Gesichtspunkts. Dementsprechend beziehen sich die Komponenten x^i , x'^i auf zwei unterschiedliche geometrische Vektoren beobachtet in demselben Koordinatensystem. Somit unterscheidet sich Gl. (B.2b) von der ersten Beziehung in Gl. (A.11), die den Zusammenhang zwischen Koordinaten eines einzigen Vektoren in zwei verschiedenen Koordinatensystemen angibt.

Bildet man die Determinante der Gleichung (B.3a), so erhält man $(\det \mathcal{O})^2 = 1$, d.h. $\det \mathcal{O} = \pm 1$. Die Matrizen mit Determinante +1 formen die *spezielle orthogonale Gruppe* $SO(3)$, deren Elemente, die *Drehmatrizen* — die in diesem Skript meistens mit dem Buchstaben \mathcal{R} bezeichnet werden —, die dreidimensionalen Drehungen um den Ursprungspunkt darstellen. Beispielsweise entspricht der Drehung um einen Winkel $\theta \in \mathbb{R}$ um die x^3 -Achse die Drehmatrix

$$\mathcal{R} = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (\text{B.4})$$

Wiederum ergeben sich die Isometrien⁽¹⁰⁷⁾ mit Determinante -1 durch Komposition der Drehungen mit der *Raumspiegelung* $\vec{x} \mapsto \vec{x}' = -\vec{x}$, d.h. der Punktspiegelung bezüglich des Ursprungspunkts des Koordinatensystems, die der Abbildungsmatrix $-\mathbb{1}_3$ entspricht.

Bemerkung: Die Bijektivität der Korrespondenz zwischen Isometrien von \mathcal{E}_3 mit einem festen Fixpunkt und orthogonalen 3×3 -Matrizen wird benutzt, um die letzteren ebenfalls als „Isometrien“ zu nennen. Dementsprechend werden die Matrizen von $SO(3)$ oft als „Drehungen“ bezeichnet, obwohl sie eigentlich nur eine der möglichen Matrixdarstellungen der Drehungen sind.

B.2 Infinitesimale Drehungen

Die Resultate dieses Abschnitts werden nirgendwo in der Vorlesung benutzt. Sie lassen sich aber problemlos nachprüfen, und werden in weiteren Vorlesungen auftauchen.

Eine Drehung um den infinitesimal kleinen Winkel $d\theta$ um die j -te Koordinatenachse transformiert einen Vektor \vec{x} von \mathbb{R}^3 in

$$\vec{x}' = \mathcal{R}(\vec{x}) = \vec{x} + d\theta \vec{e}_j \times \vec{x} \quad (\text{B.5a})$$

mit dem Einheitsvektor \vec{e}_j in Richtung j . Unter Verwendung kartesischer Koordinaten lautet diese Transformation

$$x'^k = \mathcal{R}^k_l x^l = x^k + d\theta \epsilon^k_{jl} x^l = x^k - d\theta \epsilon_j^k x^l, \quad (\text{B.5b})$$

wobei $\epsilon^k_{jl} \equiv \delta^{ki} \epsilon_{ijl}$ bzw. $\epsilon_j^k \equiv \delta^{ki} \epsilon_{jil}$, d.h. numerisch $\epsilon^k_{jl} = -\epsilon_j^k = -\epsilon_{jkl}$.⁽¹⁰⁸⁾ Nach Identifikation sind die Elemente der Darstellungsmatrix \mathcal{R} durch

$$\mathcal{R}^k_l = \delta_l^k - d\theta \epsilon_j^k \quad (\text{B.5c})$$

gegeben. Die Drehmatrix lässt sich somit als

$$\mathcal{R} = \mathbb{1}_3 - i d\theta \mathcal{J}_j \quad (\text{B.5d})$$

umschreiben, wobei \mathcal{J}_j eine 3×3 -Matrix ist, die als *Generator* oder *Erzeuger* der Gruppe $SO(3)$ bezeichnet ist, und deren Matrixelemente

$$(\mathcal{J}_j)^k_l = -i \epsilon_j^k \quad (\text{B.6a})$$

sind. In Matrixdarstellung lauten die drei Generatoren von $SO(3)$

$$\mathcal{J}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \mathcal{J}_2 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}, \quad \mathcal{J}_3 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (\text{B.6b})$$

Beispielsweise ist die Matrix einer infinitesimalen Drehung um die x^3 -Achse

$$\mathcal{R} = \mathbb{1}_3 - i d\theta \mathcal{J}_3 = \begin{pmatrix} 1 & -d\theta & 0 \\ d\theta & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (\text{B.6c})$$

wobei der letzte Ausdruck auch aus einer Taylor-Entwicklung von Gl. (B.4), mit Winkel $d\theta$ statt θ , bis zur Ordnung $\mathcal{O}(d\theta)$ folgt.

Die Matrizen (B.6b) besitzen einige wichtigen Eigenschaften. Erstens sind sie offensichtlich hermitesch.⁽¹⁰⁹⁾ Dazu prüft man einfach nach, dass die Generatoren \mathcal{J}_j den Vertauschungsrelationen

$$[\mathcal{J}_i, \mathcal{J}_j] = i \epsilon_{ijk} \mathcal{J}_k \quad \text{für alle } i, j \in \{1, 2, 3\} \quad (\text{B.6d})$$

⁽¹⁰⁷⁾ Genauer, die durch eine orthogonale Matrix mit Determinante -1 dargestellten Isometrien.

⁽¹⁰⁸⁾ In kartesischen Koordinaten spielt die Stelle der Indizes, oben oder unten, keine Rolle. In einem System beliebiger Koordinaten wird sie aber wichtig; dann sollten die Faktoren δ^{ki} durch die Komponenten g^{ki} des inversen metrischen Tensors ersetzt werden.

⁽¹⁰⁹⁾ D.h., die komplex konjugierte Matrix \mathcal{J}_j^* ist gleich der Transponierten \mathcal{J}_j^T .

genügen, wobei die rechteckigen Klammer den Kommutator zweier Matrizen bezeichnet.⁽¹¹⁰⁾

Allgemein lautet die 3×3 -Matrixdarstellung der infinitesimalen Drehung um den Winkel $d\theta$ um die Achse mit Einheitsvektor \vec{e}

$$\mathcal{R} = \mathbb{1}_3 - i d\theta \vec{e} \cdot \vec{\mathcal{J}}, \quad (\text{B.7})$$

wobei zu beachten ist, dass das Punktprodukt $\vec{e} \cdot \vec{\mathcal{J}}$ eine 3×3 -Matrix bezeichnet.

Eine Drehung um einen beliebigen Winkel θ kann dann als Produkt von N Drehungen um den Winkel θ/N um die gleiche Achse ausgedrückt werden. Somit erhält man die *Exponentialdarstellung*

$$\mathcal{R} = \lim_{N \rightarrow \infty} \left(\mathbb{1}_3 - i \frac{\theta}{N} \vec{e} \cdot \vec{\mathcal{J}} \right)^N = \exp(-i\theta \vec{e} \cdot \vec{\mathcal{J}}). \quad (\text{B.8})$$

Bemerkungen:

* Die Einführung eines Faktors i in Gl. (B.5d), und entsprechend in die Matrizen \mathcal{J}_j Gl. (B.6), ist zwar überraschend, denn die $SO(3)$ -Matrizen sind reell. Dieser Faktor erlaubt aber eine einfache Verallgemeinerung auf Gruppen komplexer Matrizen.

* Das Minus-Zeichen in Gl. (B.8) — und in den vorigen Gleichungen — gilt für eine aktive Drehung; für eine passive Drehung, entsprechend einer Basistransformation, kommt ein Plus-Zeichen.

* Die Generatoren werden oft mit T_j statt \mathcal{J}_j bezeichnet.

⁽¹¹⁰⁾ $[A, B] \equiv AB - BA$.

ANHANG C

Die δ -Distribution

C.1	Definition und erste Ergebnisse	262
C.1.1	Definition	262
C.1.2	Erste Eigenschaften	263
C.1.3	Darstellungen der δ -Distribution	263
C.1.4	Heaviside-Funktion	265
C.2	Rechenregeln	265
C.2.1	Skalierung	266
C.2.2	Ableitung	266
C.2.3	Substitution der Integrationsvariablen	266
C.3	Mehrdimensionale δ -Distributionen	267

Physiker verwenden in ihren Modellierungen oft punktförmige Objekte, wie Punktmassen oder Punktladungen. Zur Beschreibung solcher Objekte wurde die δ -Distribution eingeführt, die in diesem Anhang samt einiger ihrer Eigenschaften kurz dargestellt wird.

C.1 Definition und erste Ergebnisse

C.1.1 Definition

Definition: Als *Testfunktion* auf einem Raum Ω wird eine Funktion T bezeichnet, die einerseits glatt — d.h. beliebig oft differenzierbar — ist, und andererseits einen kompakten Träger besitzt — d.h. T verschwindet außerhalb eines beschränkten Gebiets von Ω .

Sei $\mathcal{C}_0^\infty(\Omega)$ der Raum der Testfunktionen auf Ω . In diesem und dem folgenden Abschnitt ist $\Omega = \mathbb{R}$, in Abschn. C.3 ist $\Omega = \mathbb{R}^n$ mit $n \geq 1$.

Definition: Die δ -Distribution, auch *Dirac-Distribution* oder *Dirac- δ* genannt,⁽¹¹¹⁾ ist eine stetige lineare Abbildung auf $\mathcal{C}_0^\infty(\mathbb{R})$, die einer Testfunktion T ihren Wert $T(0)$ zuordnet.

Um die Stetigkeit der δ -Distribution genauer diskutieren zu können, soll man einen Abstand auf dem Funktionenraum $\mathcal{C}_0^\infty(\mathbb{R})$ einführen.

Für die Operation der δ -Distribution auf eine Testfunktion T existieren mehrere Notationen: Mathematiker schreiben entweder $\langle \delta, T \rangle = T(0)$ oder, selten, $\delta[T] = T(0)$. Physiker verwenden aber die Schreibweise

$$\int T(x)\delta(x) dx = T(0), \quad (\text{C.1})$$

wobei das Integrationszeichen ohne Grenzen eine Integration von $-\infty$ bis ∞ bedeuten soll. Diese Notation wird hiernach weiter motiviert, muss aber mit Vorbehalt genommen werden.

Insbesondere handelt es sich beim Integrationszeichen nicht um ein (Lebesgue-)Integral mit dem üblichen Integrationsmaß, wie bei der Integration von Funktionen.

⁽¹¹¹⁾oder noch, fehlerhaft, δ - bzw. Dirac-Funktion, da es sich nicht um eine „Funktion“ handelt.

Bemerkungen:

- * Die δ -Distribution ist ein Beispiel von *Distribution*(!), d.h. von einem stetigen linearen Funktional auf einem Raum von Testfunktionen, das den üblichen Begriff einer Funktion verallgemeinert — weshalb Distributionen auch als *verallgemeinerte Funktionen* bezeichnet werden.
- * Die Beschränkung auf (Test)Funktionen mit einem kompakten Träger kann aufgehoben werden.
- * Die mithilfe einer δ -Distribution modellierten physikalischen Größen sind meistens Verteilungen, die als Gewicht in Integralen dienen, was eine erste Motivation die Schreibweise (C.1) gibt. Wichtig dabei ist, dass wenn die Variable x eine physikalische Dimension X hat, dann hat $\delta(x)$ die Dimension X^{-1} , so dass $\delta(x) dx$ dimensionslos ist.

C.1.2 Erste Eigenschaften

Definitionsgemäß ist die δ -Distribution linear: gegeben zwei beliebige Testfunktionen T_1, T_2 und zwei Zahlen λ_1, λ_2 , gilt erfüllt die Beziehung $\langle \delta, \lambda_1 T_1 + \lambda_2 T_2 \rangle = \lambda_1 \langle \delta, T_1 \rangle + \lambda_2 \langle \delta, T_2 \rangle$, d.h. mit der Notation der Physiker

$$\int [\lambda_1 T_1(x) + \lambda_2 T_2(x)] \delta(x) dx = \lambda_1 \int T_1(x) \delta(x) dx + \lambda_2 \int T_2(x) \delta(x) dx. \quad (\text{C.2})$$

Schließlich ist dies noch gleich $\lambda_1 T_1(0) + \lambda_2 T_2(0)$. Diese Eigenschaft ist eine der Motivationen hinter der Schreibweise (C.1), denn sie gilt für das übliche Integral von Funktionen.

Unter Verwendung der Integralschreibweise sieht auch die folgende *Faltungseigenschaft* trivial aus. Sei $\xi \in \mathbb{R}$; dann gilt

$$\int T(x) \delta(x - \xi) dx = T(\xi). \quad (\text{C.3})$$

Die unterliegende mathematische Konstruktion ist etwa aufwändig; dabei wird eine *translatierte δ -Distribution* $\delta_{(\xi)}$ definiert, die auf eine Testfunktion gemäß $\langle \delta_{(\xi)}, T \rangle = T(\xi)$ operiert. Dann gilt symbolisch $\langle \delta_{(\xi)}(x), T(x) \rangle \equiv \langle \delta(x), T(x + \xi) \rangle \equiv \langle \delta(x - \xi), T(x) \rangle$, wobei die letzte Gleichung als Definition von $\delta(x - \xi)$ zu betrachten ist.

Wird die „definierende“ Beziehung (C.1) auf $T(x) = e^{-ikx}$ angewandt — wobei es sich natürlich nicht um eine Testfunktion von $\mathcal{C}_0^\infty(\mathbb{R})$ handelt —, so ergibt sich

$$\int e^{-ikx} \delta(x) dx = 1. \quad (\text{C.4a})$$

Diese Gleichung sieht wie eine Formel für die Fourier-Transformation der δ -Distribution aus, und die entsprechende Rücktransformation lautet

$$\delta(x) = \int e^{ikx} \frac{dk}{2\pi}. \quad (\text{C.4b})$$

Diese Formel stellt die sog. *Fourier-Darstellung* der δ -Distribution dar.

Wieder lassen sich diese Ergebnisse mathematisch korrekt begründen, indem die herkömmliche Fourier-Transformation auf Distributionen verallgemeinert wird.

C.1.3 Darstellungen der δ -Distribution

Die δ -Distribution lässt sich als Grenzwert einer Funktionenfolge darstellen, d.h. symbolisch

$$\delta(x) = \lim_{\varepsilon \rightarrow 0} \delta_\varepsilon(x), \quad (\text{C.5a})$$

wobei δ_ε eine (integrierbare) Funktion im üblichen Sinn ist, die von einem reellen Parameter $\varepsilon > 0$ kontinuierlich abhängt. Dabei bedeutet der obige Grenzwert

$$\int f(x)\delta(x) dx = \lim_{\varepsilon \rightarrow 0} \int f(x)\delta_\varepsilon(x) dx \quad (\text{C.5b})$$

für jede Funktion f mit bestimmten Eigenschaften — z.B. jede Testfunktion von $\mathcal{C}_0^\infty(\mathbb{R})$.⁽¹¹²⁾ Dann bilden die Funktionen δ_ε eine *Darstellung der δ -Distribution*.

Da δ_ε eine Funktion ist, ist das Integral auf der rechten Seite von Gl. (C.5b) ein übliches (Lebesgue)-Integral. Somit wird das Verwenden der Integralnotation für die Operation auf f der δ -Distribution, auf der linken Seite von Gl. (C.5b), weiter motiviert.

Die Funktionen δ_ε werden oft so gewählt, dass sie die folgenden Eigenschaften haben:

- sie haben einen Peak bei $x = 0$, der mit abnehmendem ε schmaler und höher wird;
- für jeden $x \neq 0$ gilt $\lim_{\varepsilon \rightarrow 0} \delta_\varepsilon(x) = 0$;
- $\int \delta_\varepsilon(x) dx = 1$.

Man prüft leicht nach, dass die hiernach angegebenen Beispiele diese Eigenschaften erfüllen.

C.1.3a Skalierte Rechteckfunktionen

Eine einfache Darstellung, jedoch durch unstetige Funktionen, ist

$$\delta_\varepsilon(x) = \begin{cases} \frac{1}{2\varepsilon} & \text{für } |x| < \varepsilon \\ 0 & \text{für } |x| > \varepsilon. \end{cases} \quad (\text{C.6})$$

C.1.3b Skalierte Dreieckfunktionen

Eine stetige, aber nicht überall differenzierbare Darstellung, ist

$$\delta_\varepsilon(x) = \begin{cases} \frac{\varepsilon - |x|}{\varepsilon^2} & \text{für } |x| \leq \varepsilon \\ 0 & \text{für } |x| > \varepsilon. \end{cases} \quad (\text{C.7})$$

C.1.3c Exponentialfunktionen

Ebenfalls nicht überall differenzierbar — diesmal liegt das Problem in $x = 0$ — ist die Darstellung durch Exponentialfunktionen

$$\delta_\varepsilon(x) = \frac{1}{2\varepsilon} e^{-|x|/\varepsilon}. \quad (\text{C.8})$$

C.1.3d Gaußsche Verteilungen

Gaußsche Verteilungen liefern eine unendlich differenzierbare Darstellung:

$$\delta_\varepsilon(x) = \frac{1}{\varepsilon\sqrt{2\pi}} e^{-x^2/2\varepsilon^2}. \quad (\text{C.9})$$

C.1.3e Lorentz-Verteilungen

Mithilfe von Lorentz-Verteilungen ergibt sich die Darstellung:

$$\delta_\varepsilon(x) = \frac{1}{\pi} \frac{\varepsilon}{x^2 + \varepsilon^2} = \frac{1}{\pi} \operatorname{Im} \frac{1}{x - i\varepsilon}. \quad (\text{C.10})$$

⁽¹¹²⁾ Je nach der gewählten Klasse von Funktionen f können die Funktionen δ_ε eine passende Darstellung bilden oder nicht.

C.1.3f Kardinalsinus

Ebenfalls nützlich ist die Darstellung

$$\delta_\varepsilon(x) = \begin{cases} \frac{\sin(x/\varepsilon)}{\pi x} & \text{für } x \neq 0 \\ \frac{1}{\pi\varepsilon} & \text{für } x = 0, \end{cases} \quad (\text{C.11})$$

wobei die Definition des Zahlenwerts in $x = 0$ die Stetigkeit der Funktion δ_ε gewährleistet. Für $x \neq 0$ gilt nämlich

$$\delta_\varepsilon(x) = \int_{-1/\varepsilon}^{1/\varepsilon} e^{ikx} \frac{dk}{2\pi},$$

und der Limes $\varepsilon \rightarrow 0$ gibt die Fourier-Darstellung (C.4b) wieder.

C.1.4 Heaviside-Funktion

Definition: Die *Heaviside-* oder *Stufenfunktion* wird durch

$$\Theta(x) \equiv \begin{cases} 1 & \text{für } x > 0 \\ 0 & \text{für } x < 0 \end{cases} \quad (\text{C.12})$$

definiert, wobei der Wert für $x = 0$ irrelevant ist.

Für diese Funktion kann man, genau wie für die δ -Distribution, Darstellungen Θ_ε finden, mit

$$\Theta(x) = \lim_{\varepsilon \rightarrow 0} \Theta_\varepsilon(x) \quad (\text{C.13a})$$

oder genauer

$$\lim_{\varepsilon \rightarrow 0} \int_{-\infty}^{\infty} f(x) \Theta_\varepsilon(x) dx = \int_{-\infty}^{\infty} f(x) \Theta(x) dx = \int_0^{\infty} f(x) dx \quad (\text{C.13b})$$

für jede passende Funktion f , wobei die zweite Gleichung aus der Definition der Heaviside-Funktion folgt.

Dann liefert die Ableitung jeder differenzierbaren Darstellung Θ_ε der Heaviside-Funktion eine Darstellung δ_ε der δ -Distribution: symbolisch schreibt man $\Theta'_\varepsilon = \delta_\varepsilon$ oder noch

$$\Theta'(x) = \delta(x). \quad (\text{C.14})$$

Dabei handelt es sich um eine Ableitung „im Sinne der Distributionen“.

Beispiele von Darstellungen der Heaviside-Funktion sind

$$\Theta_\varepsilon(x) = \begin{cases} 0 & \text{für } x < -\varepsilon \\ \frac{x + \varepsilon}{2\varepsilon} & \text{für } |x| \leq \varepsilon \\ 1 & \text{für } x > \varepsilon, \end{cases} \quad (\text{C.15})$$

entsprechend der Darstellung (C.6) der δ -Distribution, oder

$$\Theta_\varepsilon(x) = \frac{1}{2} + \frac{1}{\pi} \arctan\left(\frac{x}{\varepsilon}\right), \quad (\text{C.16})$$

deren Ableitung die Darstellung (C.10) der δ -Distribution wiedergibt.

C.2 Rechenregeln

Neben den schon in § C.1.2 angegebenen Eigenschaften gelten noch weitere, die hiernach meistens ohne expliziten Beweis⁽¹¹³⁾ aufgelistet werden.

⁽¹¹³⁾Die interessierte Leserin kann die verschiedenen Gleichungen relativ leicht prüfen, indem sie die δ -Distribution als Grenzwert einer Darstellung schreibt.

C.2.1 Skalierung

Sei $a \in \mathbb{R}$ (oder allgemeiner $a \in \mathbb{C}$); dann gilt eine Skalierungseigenschaft, die sich symbolisch als

$$\delta(ax) = \frac{1}{|a|} \delta(x) \quad (\text{C.17a})$$

schreiben lässt, was eigentlich

$$\int T(x) \delta(ax) dx = \frac{1}{|a|} T(0), \quad (\text{C.17b})$$

für jede Testfunktion T bedeutet.

Für eine Darstellung δ_ε folgt die entsprechende Gleichung einfach aus einer Substitution der Integrationsvariablen.

Insbesondere gilt für $a = -1$

$$\delta(-x) = \delta(x), \quad (\text{C.18})$$

d.h. die δ -Distribution ist gerade.

C.2.2 Ableitung

Man definiert eine Distribution δ' , die sich bezüglich der δ -Distribution wie die Ableitung f' bezüglich einer (Stamm)Funktion f verhält.

Dann gilt für jede Testfunktion T

$$\int T(x) \delta'(x) dx = -T'(0). \quad (\text{C.19})$$

Nimmt man die Integralschreibweise ernst und führt man eine partielle Integration durch, so ergibt sich

$$\int T(x) \delta'(x) dx = \left[T(x) \delta(x) \right]_{-\infty}^{\infty} - \int T'(x) \delta(x) dx,$$

woraus das gesuchte Ergebnis „folgt“, denn T verschwindet im Unendlichen. \square

Auf ähnlicher Weise wird eine zweite Ableitung δ'' definiert, für die

$$\int T(x) \delta''(x) dx = T''(0) \quad (\text{C.20})$$

gilt, und so weiter für die höheren Ableitungen.

C.2.3 Substitution der Integrationsvariablen

Sei f eine differenzierbare reelle Funktion auf \mathbb{R} , mit einer einzigen Nullstelle in x_0 : $f(x_0) = 0$, wo ihre Ableitung nicht Null ist, $f'(x_0) \neq 0$. Dann gilt für die Verkettung von der δ -Distribution mit f

$$\delta(f(x)) = \frac{1}{|f'(x_0)|} \delta(x - x_0), \quad (\text{C.21a})$$

entsprechend für jede Testfunktion T

$$\int T(x) \delta(f(x)) dx = \frac{1}{|f'(x_0)|} \int T(x) \delta(x - x_0) dx = \frac{1}{|f'(x_0)|} T(x_0). \quad (\text{C.21b})$$

Dank den Annahmen über f gilt in der Nachbarschaft von x_0

$$f(x) \sim f(x_0) + f'(x_0)(x - x_0) + \mathcal{O}((x - x_0)^2) = f'(x_0)(x - x_0) + \mathcal{O}((x - x_0)^2).$$

Das Ergebnis folgt dann aus der Skalierungseigenschaft (C.17). \square

Für eine Funktion f mit mehreren Nullstellen x_i , in denen die Ableitung f' nicht verschwindet, gilt die Verallgemeinerung

$$\delta(f(x)) = \sum_{\text{Nullstellen } x_i} \frac{1}{|f'(x_i)|} \delta(x - x_i). \quad (\text{C.22})$$

Beispiel 1: Sei $f(x) = \arctan x$, mit einer einzigen Nullstelle in $x_0 = 0$, während die Ableitung $f'(x) = 1/(1+x^2)$ nirgendwo verschwindet. Da $f'(0) = 1$ lautet wird Gl. (C.21a) zu

$$\delta(\arctan x) = \delta(x).$$

Beispiel 2: Sei $f(x) = x^2 - a^2$ mit $a \in \mathbb{R}_+$; f hat zwei Nullstellen in $-a$ und a . Aus $f'(x) = 2x$ folgen $f'(-a) = -2a$, $f'(a) = 2a$, und somit $|f'(-a)| = |f'(a)| = 2a$. Somit ergibt Gl. (C.22)

$$\delta(x^2 - a^2) = \frac{1}{2a} [\delta(x - a) + \delta(x + a)]. \quad (\text{C.23})$$

Bemerkungen:

* Falls eine Funktion f eine Nullstelle x_0 mit $f'(x_0) = 0$ hat, kann man die Funktion oft leicht „verformen“, um statt nur x_0 zwei (oder mehr) Nullstellen zu erhalten, in denen die Ableitung nicht verschwindet.

Beispielsweise würde man $\delta(x^2)$ durch $\delta(x^2 - \eta)$ ersetzen, den letzteren Ausdruck mithilfe der Gl. (C.23) umschreiben, die daraus folgenden Berechnungen durchführen, und schließlich den Limes $\eta \rightarrow 0$ betrachten... und hoffen, dass das Ergebnis sinnvoll ist!

* Wie in Gl. (C.22) implizit ist, darf man δ -Distributionen problemlos miteinander addieren. Mit Produkten von δ -Distributionen muss man aber vorsichtig sein: insbesondere ist das Produkt von $\delta(x)$ mit sich selbst nicht definiert. Dagegen dürfen δ -Distributionen von unterschiedlichen Argumenten, wie z.B. $\delta(x)\delta(y)$ oder $\delta(x)\delta(y-x)$, miteinander multipliziert werden. Dann muss aber über alle Argumente integriert werden, um eine Zahl zu erhalten, wie im nächsten Abschnitt ausführlicher diskutiert wird.

C.3 Mehrdimensionale δ -Distributionen

Seien x^1, x^2, \dots, x^n unabhängige Variablen, die in einen n -dimensionalen Vektor \mathbf{x} zusammengefasst werden. Betrachtet man Testfunktionen $T(\mathbf{x})$ dieser n Variablen, so kann man eine n -dimensionale δ -Distribution, die oft mit $\delta^{(n)}$ bezeichnet wird,⁽¹¹⁴⁾ gemäß

$$\int T(\mathbf{x}) \delta^{(n)}(\mathbf{x}) d^n \mathbf{x} = T(\mathbf{0}), \quad (\text{C.24})$$

wobei das Integral über \mathbb{R}^n ist. Anders gesagt gilt

$$\delta^{(n)}(\mathbf{x}) \equiv \delta(x^1) \cdots \delta(x^n) = \prod_{k=1}^n \delta(x^k). \quad (\text{C.25})$$

Die für die eindimensionale δ -Distribution angegebenen Eigenschaften lassen sich einfach auf $\delta^{(n)}$ verallgemeinern. Beispielsweise gelten die Faltungseigenschaft

$$\int T(\mathbf{x}) \delta^{(n)}(\mathbf{x} - \boldsymbol{\xi}) d^n \mathbf{x} = T(\boldsymbol{\xi}), \quad (\text{C.26})$$

⁽¹¹⁴⁾ Dies sollte man nicht mit der n -ten Ableitung von der eindimensionalen δ -Distribution verwechseln — in der Physik treten höhere Ableitungen von δ aber extrem selten.

wobei $\boldsymbol{\xi} \in \mathbb{R}^n$, sowie die Fourier-Darstellung

$$\delta^{(n)}(\boldsymbol{x}) = \int e^{i\boldsymbol{k} \cdot \boldsymbol{x}} \frac{d^n \boldsymbol{k}}{(2\pi)^n} \quad (\text{C.27})$$

mit dem euklidischen Skalarprodukt $\boldsymbol{k} \cdot \boldsymbol{x} = k^1 x^1 + \dots + k^n x^n$.

Die Skalierungseigenschaft lautet jetzt

$$\delta^{(n)}(a\boldsymbol{x}) = \frac{1}{|a|^n} \delta^{(n)}(\boldsymbol{x}) \quad (\text{C.28})$$

für $a \in \mathbb{R}$, und kann erweitert werden: wenn \mathcal{A} eine invertierbare $n \times n$ -Matrix ist, dann ist

$$\delta^{(n)}(\mathcal{A}\boldsymbol{x}) = \frac{1}{|\det \mathcal{A}|} \delta^{(n)}(\boldsymbol{x}). \quad (\text{C.29})$$

Ähnlich gilt im Fall einer Variablenänderung $(x^1, \dots, x^n) \rightarrow (y^1, \dots, y^n)$ bzw. $\boldsymbol{x} \rightarrow \boldsymbol{y}$

$$\delta^{(n)}(\boldsymbol{y}) = \frac{1}{|\det J|} \delta^{(n)}(\boldsymbol{x}) \quad (\text{C.30})$$

mit J der Jacobi-Matrix der Substitution, mit Elementen $\partial y^i / \partial x^j$. Aus dieser letzten Gleichung lässt sich die Verallgemeinerung der Formel (C.22) ableiten.

Literatur zum Anhang C

- Gelfand & Shilov, *Verallgemeinerte Funktionen. Band 1* [36], Kap. I.
- Schwartz, *Mathematische Methoden der Physik. Band 1* [37], Kap. II.

ANHANG D

Legendre-Transformation

D.1 Legendre-Transformation einer Funktion einer Variablen

Sei f eine strikt konvexe reelle Funktion einer reellen Variablen x auf einem Definitionsintervall \mathcal{I} , d.h. für alle $x_1, x_2 \in \mathcal{I}$ mit $x_1 \neq x_2$ und $\lambda \in]0, 1[$ gilt

$$f(\lambda x_1 + (1 - \lambda)x_2) < \lambda f(x_1) + (1 - \lambda)f(x_2). \quad (\text{D.1})$$

Die Ungleichung (D.1) bedeutet, dass der Graph der Funktion unterhalb der Verbindungsstrecke zweier seiner Punkten liegt, wie in Abb. D.1 illustriert wird.

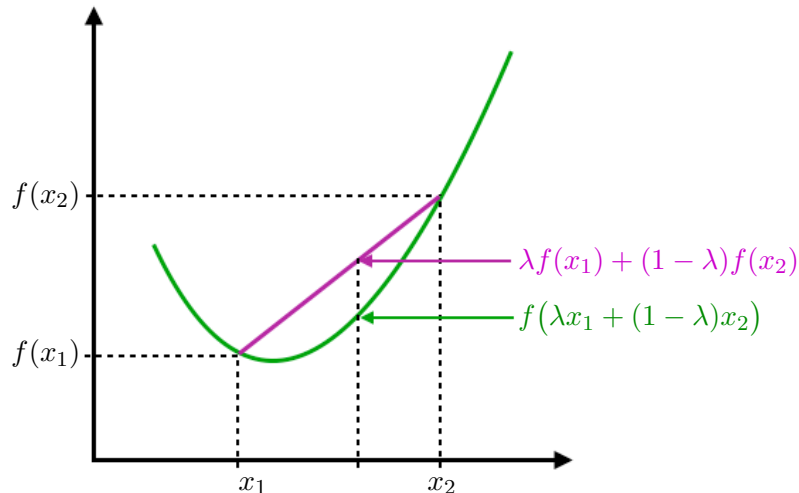


Abbildung D.1 – Beispiel einer strikt konvexen reellen Funktion.

Sei zudem angenommen, dass f zweimal stetig differenzierbar ist. Dann ist die erste Ableitung f' streng monoton wachsend auf \mathcal{I} und demzufolge die zweite Ableitung $f''(x)$ positiv für alle $x \in \mathcal{I}$. Die Funktion y auf \mathcal{I} definiert durch $y(x) \equiv f'(x)$ für alle $x \in \mathcal{I}$ ist eine bijektive Abbildung von \mathcal{I} nach ihrer Wertemenge \mathcal{J} ; daher ist sie invertierbar. Hiernach wird die Umkehrfunktion als $x(y)$ bezeichnet.

Die *Legendre-Transformierte* der Funktion f ist eine Funktion g auf \mathcal{J} , die für alle $y \in \mathcal{J}$ durch

$$g(y) \equiv yx(y) - f(x(y)) \quad (\text{D.2})$$

definiert ist.

Da die Ableitung von g nach der Variablen y gleich $x(y)$ ist, findet man, dass eine zweite Legendre-Transformation von g genau f gibt. Das heißt, die Legendre-Rücktransformation nimmt die gleiche Form wie die direkte Transformation an.

Sei $h(z) \equiv zy(z) - g(y(z))$ mit $z(y) \equiv g'(y)$. Die Anwendung der Kettenregel auf Gl. (D.2) gibt

$$g'(y) = x(y) + yx'(y) - f'(x(y))x'(y) = x(y),$$

d.h. $z(y) = x(y)$. Daher gilt $h(x) = xy(x) - g(y(x)) = f(x)$, wobei Gl. (D.2) benutzt wurde. \square

Zudem gelten Beziehungen zwischen den Ableitungen von f und g nach ihrer jeweiligen Variablen ab der zweiten Ordnung, insbesondere

$$f''(x) g''(y(x)) = f''(x(y)) g''(y) = 1. \quad (\text{D.3})$$

D.2 Legendre-Transformation einer Funktion mehrerer Variablen

Sei jetzt eine zweimal kontinuierlich differenzierbare Funktion $f(x_1, \dots, x_n, t_1, \dots, t_p)$ von $n + p$ reellen Variablen, mit

$$\det \left(\frac{\partial^2 f}{\partial x_j \partial x_k} \right)_{j,k} \neq 0,$$

wobei $()_{j,k}$ die $n \times n$ -Matrix mit Elementen $\partial^2 f / \partial x_j \partial x_k$ bezeichnet. Für jedes $j = 1, \dots, n$ wird eine Funktion y_j durch

$$y_j(\{x_i\}, \{t_k\}) \equiv \frac{\partial f}{\partial x_j}(\{x_i\}, \{t_k\})$$

definiert. Die Umkehrfunktionen, deren Existenz durch die Forderungen an die Matrix der zweiten Ableitungen von f gewährleistet wird, werden mit $x_i(\{y_j\}, \{t_k\})$ bezeichnet.

Die Legendre-Transformierte g von f bezüglich der Variablen $\{x_i\}$ wird durch

$$g(y_1, \dots, y_n, t_1, \dots, t_p) \equiv \sum_{j=1}^n y_j x_j(y_1, \dots, y_n, t_1, \dots, t_p) - f(x_1, \dots, x_n, t_1, \dots, t_p) \quad (\text{D.4})$$

definiert. Hier auch ist die Rücktransformation gleich der direkten Transformation, d.h. eine zweite Legendre-Transformation von g gibt genau f wieder.

Man sieht sofort, dass die ersten Ableitungen von f und g nach einer der Variablen $\{t_k\}$ negativ zueinander sind:

$$\frac{\partial g}{\partial t_j}(\{y_i\}, \{t_k\}) = -\frac{\partial f}{\partial t_j}(\{x_i\}, \{t_k\}). \quad (\text{D.5})$$

Für die zweiten Ableitungen von f und g nach ihren jeweiligen natürlichen Variablen ergeben sich mehrere Beziehungen, wie z.B.

$$\sum_{j=1}^n \frac{\partial^2 g}{\partial y_i \partial y_j} \frac{\partial^2 f}{\partial x_j \partial x_k} = \delta_{ik} \quad (\text{D.6a})$$

analog zur Gl. (D.3) mit dem Kronecker-Symbol δ_{ik} , sowie

$$\frac{\partial^2 g}{\partial y_i \partial t_k} = \sum_{j=1}^n \frac{\partial^2 g}{\partial y_i \partial y_j} \frac{\partial^2 f}{\partial x_j \partial t_k}, \quad (\text{D.6b})$$

$$\frac{\partial^2 g}{\partial t_i \partial t_k} = \frac{\partial^2 f}{\partial t_i \partial t_k} + \sum_{j,l=1}^n \frac{\partial^2 f}{\partial t_i \partial x_j} \frac{\partial^2 g}{\partial y_j \partial y_l} \frac{\partial^2 f}{\partial x_l \partial t_k}. \quad (\text{D.6c})$$

Bemerkung: Manchmal, insbesondere in der Thermodynamik und Statistischen Physik, nennt man $-g$ die Legendre-Transformierte von f anstatt g .

Literatur zum Anhang D

- Arnold, *Mathematical methods of classical mechanics* [1] Kap. 3 § 14.
- Nolting, *Analytische Mechanik* [16] Kap. 2.1.
- Scheck, *Mechanik* [18] Kap. 2.13–2.14.

ANHANG E

Vektoranalysis

In diesem Anhang werden einige nützlichen Formeln der Vektoranalysis gesammelt.
symbolischer *Nabla*-Differentialoperator

$$\vec{\nabla} \equiv \vec{e}_1 \frac{\partial}{\partial x^1} + \vec{e}_2 \frac{\partial}{\partial x^2} + \vec{e}_3 \frac{\partial}{\partial x^3} = \sum_{i=1}^3 \vec{e}_i \frac{\partial}{\partial x^i}. \quad (\text{E.1})$$

E.0.0a

$$\vec{\nabla} \cdot [\vec{\nabla} f(\vec{r})] = \Delta f(\vec{r}), \quad (\text{E.2})$$

d.h. die Divergenz des Gradienten

$$\vec{\nabla} \times [\vec{\nabla} f(\vec{r})] = \vec{0}. \quad (\text{E.3})$$

$$\vec{\nabla} \cdot [\vec{\nabla} \times \vec{V}(\vec{r})] = 0. \quad (\text{E.4})$$

Die Divergenz einer Rotation verschwindet.

$$\vec{\nabla} \times [\vec{\nabla} \times \vec{V}(\vec{r})] = \vec{\nabla} [\vec{\nabla} \cdot \vec{V}(\vec{r})] - \Delta \vec{V}(\vec{r}). \quad (\text{E.5})$$

E.0.0b

Produktregel

$$\vec{\nabla} [f_1(\vec{r}) f_2(\vec{r})] = [\vec{\nabla} f_1(\vec{r})] f_2(\vec{r}) + f_1(\vec{r}) \vec{\nabla} f_2(\vec{r}). \quad (\text{E.6})$$

$$\vec{\nabla} \cdot [f(\vec{r}) \vec{V}(\vec{r})] = [\vec{\nabla} f(\vec{r})] \cdot \vec{V}(\vec{r}) + f(\vec{r}) \vec{\nabla} \cdot \vec{V}(\vec{r}). \quad (\text{E.7})$$

$$\vec{\nabla} \times [f(\vec{r}) \vec{V}(\vec{r})] = [\vec{\nabla} f(\vec{r})] \times \vec{V}(\vec{r}) + f(\vec{r}) \vec{\nabla} \times \vec{V}(\vec{r}). \quad (\text{E.8})$$

$$\vec{\nabla} \cdot [\vec{V}_1(\vec{r}) \times \vec{V}_2(\vec{r})] = \vec{V}_2(\vec{r}) \cdot [\vec{\nabla} \times \vec{V}_1(\vec{r})] - \vec{V}_1(\vec{r}) \cdot [\vec{\nabla} \times \vec{V}_2(\vec{r})]. \quad (\text{E.9})$$

$$\vec{\nabla} \times [\vec{V}_1(\vec{r}) \times \vec{V}_2(\vec{r})] = [\vec{\nabla} \cdot \vec{V}_2(\vec{r})] \vec{V}_1(\vec{r}) - [\vec{\nabla} \cdot \vec{V}_1(\vec{r})] \vec{V}_2(\vec{r}) + [\vec{V}_2(\vec{r}) \cdot \vec{\nabla}] \vec{V}_1(\vec{r}) - [\vec{V}_1(\vec{r}) \cdot \vec{\nabla}] \vec{V}_2(\vec{r}). \quad (\text{E.10})$$

$$\vec{\nabla} \cdot [\vec{V}_1(\vec{r}) \cdot \vec{V}_2(\vec{r})] = \vec{V}_1(\vec{r}) \times [\vec{\nabla} \times \vec{V}_2(\vec{r})] + \vec{V}_2(\vec{r}) \times [\vec{\nabla} \times \vec{V}_1(\vec{r})] + [\vec{V}_1(\vec{r}) \cdot \vec{\nabla}] \vec{V}_2(\vec{r}) + [\vec{V}_2(\vec{r}) \cdot \vec{\nabla}] \vec{V}_1(\vec{r}). \quad (\text{E.11})$$

ANHANG F

Kugelflächenfunktionen

Betrachte man (zweimal kontinuierlich differenzierbare) Funktionen auf der Einheitskugelfläche \mathbb{S}^2 von \mathbb{R}^3 , d.h. Funktionen des Polarwinkels $\theta \in [0, \pi]$ und des Azimutwinkels $\varphi \in [0, 2\pi]$.

Unter diesen Funktionen lassen sich sog. *Kugelflächenfunktionen* $Y_{\ell m}(\theta, \varphi) \in \mathbb{C}$ mit $\ell = 0, 1, \dots$ einer beliebigen ganzen Zahl und $m \in \{-\ell, -\ell+1, \dots, 0, \dots, \ell\}$ mit folgenden Eigenschaften finden:

- Die Kugelflächenfunktionen sind gleichzeitige Eigenfunktionen der zwei hiernach über ihre Operation auf eine zweimal differenzierbare Funktion f definierten Differentialoperatoren $\hat{\mathcal{L}}^2$ und $\hat{\mathcal{L}}_3$

$$\hat{\mathcal{L}}^2 f(\theta, \varphi) \equiv \frac{-1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial f(\theta, \varphi)}{\partial \theta} \right) - \frac{1}{\sin^2 \theta} \frac{\partial^2 f(\theta, \varphi)}{\partial \varphi^2} \quad (\text{F.1a})$$

und

$$\hat{\mathcal{L}}_3 f(\theta, \varphi) \equiv -i \frac{\partial f(\theta, \varphi)}{\partial \varphi}. \quad (\text{F.1b})$$

Genauer gelten

$$\hat{\mathcal{L}}^2 Y_{\ell m}(\theta, \varphi) = \ell(\ell+1) Y_{\ell m}(\theta, \varphi) \quad \text{und} \quad \hat{\mathcal{L}}_3 Y_{\ell m}(\theta, \varphi) = m Y_{\ell m}(\theta, \varphi). \quad (\text{F.2})$$

- Die Kugelflächenfunktionen genügen der Beziehung

$$\int_0^\pi \left[\int_0^{2\pi} Y_{\ell m}^*(\theta, \varphi) Y_{\ell' m'}(\theta, \varphi) d\varphi \right] \sin \theta d\theta = \delta_{\ell \ell'} \delta_{m m'}. \quad (\text{F.3})$$

- Jede zweimal kontinuierlich differenzierbare Funktion f auf \mathbb{S}^2 lässt sich als Summe

$$f(\theta, \varphi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\theta, \varphi) \quad (\text{F.4a})$$

von Kugelflächenfunktionen schreiben, wobei die Koeffizienten $a_{\ell m}$ eindeutig durch

$$a_{\ell m} = \int_0^\pi \left[\int_0^{2\pi} Y_{\ell m}^*(\theta, \varphi) f(\theta, \varphi) d\varphi \right] \sin \theta d\theta \quad (\text{F.4b})$$

gegeben sind.

Diese Eigenschaften werden beispielsweise durch die folgenden Funktionen erfüllt: für $\ell = 0$

$$Y_{00}(\theta, \varphi) = \frac{1}{\sqrt{4\pi}}; \quad (\text{F.5a})$$

für $\ell = 1$

$$Y_{10}(\theta, \varphi) = \sqrt{\frac{3}{4\pi}} \cos \theta \quad (\text{F.5b})$$

$$Y_{1\pm 1}(\theta, \varphi) = \mp \sqrt{\frac{3}{8\pi}} \sin \theta e^{\pm i\varphi}; \quad (\text{F.5c})$$

für $\ell = 2$

$$Y_{20}(\theta, \varphi) = \sqrt{\frac{5}{16\pi}} (\cos^2 \theta - 1) \quad (\text{F.5d})$$

$$Y_{2\pm 1}(\theta, \varphi) = \mp \sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta e^{\pm i\varphi} \quad (\text{F.5e})$$

$$Y_{2\pm 2}(\theta, \varphi) = \sqrt{\frac{15}{32\pi}} \sin^2 \theta e^{\pm 2i\varphi}; \quad (\text{F.5f})$$

usw.

Bemerkungen:

* Der Differentialoperator $\hat{\mathcal{L}}^2$ ist in der Tat der Winkelanteil des Laplace-Operators in Kugelkoordinaten (r, θ, φ) :

$$\Delta f(\vec{r}) = \frac{1}{r} \frac{\partial^2}{\partial r^2} [r f(\vec{r})] - \frac{\hat{\mathcal{L}}^2 f(\vec{r})}{r^2} \quad (\text{F.6})$$

* Die dritte Eigenschaft [Gl. (F.4)] bedeutet, dass die Kugelflächenfunktionen eine (Hilbert-)Basis des Raums der zweimal kontinuierlich differenzierbaren Funktionen auf \mathbb{S}^2 bilden. Die Koeffizienten $a_{\ell m}$ sind dann die Komponenten einer Funktion f auf dieser Basis.

* Führt man eine Abbildung über Funktionen auf \mathbb{S}^2

$$(f, g) \mapsto \langle f, g \rangle \equiv \int_0^\pi \left[\int_0^{2\pi} f^*(\theta, \varphi) g(\theta, \varphi) d\varphi \right] \sin \theta d\theta$$

ein, so ist diese eine hermitesche Sesquilinearform,⁽¹¹⁵⁾ die ein Skalarprodukt auf dem Raum der Funktionen definiert. Dann ist die Basis der $\{Y_{\ell m}\}$ laut Gl. (F.3) eine Orthonormalbasis für dieses Skalarprodukt. Einerseits sind unterschiedliche Kugelflächenfunktionen orthogonal zueinander ($\langle Y_{\ell m}, Y_{\ell' m'} \rangle = 0$ falls $\ell \neq \ell'$ und $m \neq m'$); andererseits ist jede $Y_{\ell m}$ auf 1 normiert ($\langle Y_{\ell m}, Y_{\ell m} \rangle = 1$).

* Manchmal werden reellwertige Kugelflächenfunktionen anstatt komplexwertiger Funktionen definiert. Diese reellen Varianten — wie z.B. $(Y_{11} + Y_{1-1})/\sqrt{2}$ oder $(Y_{11} - Y_{1-1})/\sqrt{2}$ — bilden noch eine Orthonormalbasis und sind Eigenfunktionen zum Operator $\hat{\mathcal{L}}^2$; sie sind aber nicht mehr Eigenfunktionen zu $\hat{\mathcal{L}}_3$.

⁽¹¹⁵⁾D.h. die Abbildung $\langle \cdot, \cdot \rangle$ erfüllt für alle Funktionen f, g, h die Bedingungen i. $\langle f, f \rangle \geq 0$; ii. $\langle f, f \rangle = 0$ nur für $f = 0$; iii. $\langle f, g \rangle = \langle g, f \rangle^*$; iv. $\langle f, \lambda g + \mu h \rangle = \lambda \langle f, g \rangle + \mu \langle f, h \rangle$ für alle $\lambda, \mu \in \mathbb{C}$.

Literaturverzeichnis

- [1] V. I. Arnold, *Mathematical methods of classical mechanics*, 2nd Aufl. (Springer, New York, 1989).
- [2] T. Fließbach, *Lehrbuch zur theoretischen Physik I. Mechanik*, 7. Aufl. (Springer-Verlag, Berlin & Heidelberg, 2015).
- [3] T. Fließbach, *Lehrbuch zur theoretischen Physik II. Elektrodynamik*, 6. Aufl. (Springer-Verlag, Berlin & Heidelberg, 2012).
- [4] H. Goldstein, C. Poole, J. Safko, *Classical mechanics*, 3. Aufl. (Wiley-VCH, Weinheim, 2012).
- [5] H. Goldstein, C. Poole, J. Safko, *Classical mechanics*, 3rd Aufl. (Addison-Wesley, Reading, MA, 2001).
- [6] W. Greiner, *Klassische Mechanik I*, 8. Aufl. (Harri Deutsch, Frankfurt am Main, 2007).
- [7] W. Greiner, *Klassische Mechanik II*, 8. Aufl. (Harri Deutsch, Frankfurt am Main, 2008).
- [8] W. Greiner, *Klassische Elektrodynamik*, 7. Aufl. (Harri Deutsch, Frankfurt am Main, 2008).
- [9] D. J. Griffiths, *Elektrodynamik*, 3. Aufl. (Pearson, München, 2011).
- [10] D. J. Griffiths, *Introduction to Electrodynamics*, 3rd Aufl. (Prentice-Hall, Upper Saddle River, NJ, 1999).
- [11] J. D. Jackson, *Klassische Elektrodynamik*, 4. Aufl. (de Gruyter, Berlin, 2013).
- [12] J. D. Jackson, *Classical electrodynamics*, 3. Aufl. (John Wiley & Sons, New York, 1999).
- [13] L. Landau, E. Lifschitz, *Lehrbuch der theoretischen Physik. Band I: Mechanik*, 14. Aufl. (Harri Deutsch, Frankfurt am Main, 1997).
- [14] L. Landau, E. Lifschitz, *Lehrbuch der theoretischen Physik. Band II: Klassische Feldtheorie*, 12. Aufl. (Harri Deutsch, Frankfurt am Main, 1997).
- [15] W. Nolting, *Grundkurs Theoretische Physik. Band 1: Klassische Mechanik*, 8. Aufl. (Springer, Berlin & Heidelberg, 2006).
- [16] W. Nolting, *Grundkurs Theoretische Physik. Band 2: Analytische Mechanik*, 7. Aufl. (Springer, Berlin & Heidelberg, 2006).
- [17] W. Nolting, *Grundkurs Theoretische Physik. Band 3: Elektrodynamik*, 8. Aufl. (Springer, Berlin & Heidelberg, 2007).
- [18] F. Scheck, *Theoretische Physik 1: Mechanik*, 8. Aufl. (Springer, Berlin & Heidelberg, 2007).
- [19] F. Scheck, *Theoretische Physik 3: Klassische Feldtheorie*, 2. Aufl. (Springer, Berlin & Heidelberg, 2006).
- [20] I. S. Gradshteyn, I. M. Ryzhik, *Table of integrals, series, and products*, 7th Aufl. (Academic Press, Amsterdam, 2007).

- [21] F. W. J. Olver *et al.*, *NIST Handbook of mathematical functions* (University Press, Cambridge, 2010), [available online at <http://dlmf.nist.gov/>].
- [22] R. P. Feynman, R. B. Leighton, M. Sands, *Feynman-Vorlesungen über Physik. Band 2: Elektromagnetismus und Struktur der Materie*, 5. Aufl. (Oldenbourg Wissenschaftsverlag, München, 2007).
- [23] R. P. Feynman, R. B. Leighton, M. Sands, *The Feynman Lectures on Physics. Volume II: Mainly Electromagnetism and Matter*, definitive Aufl. (Addison-Wesley, Reading, MA, 2005).
- [24] R. P. Feynman, R. B. Leighton, M. Sands, *Feynman-Vorlesungen über Physik. Band 1: Mechanik, Strahlung, Wärme*, 5. Aufl. (Oldenbourg Wissenschaftsverlag, München, 2007).
- [25] R. P. Feynman, R. B. Leighton, M. Sands, *The Feynman Lectures on Physics. Volume I: Mainly Mechanics, Radiation, and Heat*, definitive Aufl. (Addison-Wesley, Reading, MA, 2005).
- [26] L. Landau, E. Lifshitz, *Course of Theoretical Physics. Vol. 1: Mechanics*, 3rd Aufl. (Butterworth-Heinemann, Oxford, 1976).
- [27] D. Rim, *An Elementary Proof That Symplectic Matrices Have Determinant One*, [arXiv:1505.04240].
- [28] L. Landau, E. Lifshitz, *Course of Theoretical Physics. Vol. 2: The classical theory of fields*, 4th Aufl. (Butterworth-Heinemann, Oxford, 1975).
- [29] D. Petrascheck, *Eur. J. Phys.* **37** (2016) 015201.
- [30] P. Aichelburg, R. Beig, *Ann. Phys. (NY)* **98** (1976) 264–283.
- [31] J. A. Wheeler, R. P. Feynman, *Rev. Mod. Phys.* **17** (1945) 157–181.
- [32] W. Nolting, *Grundkurs Theoretische Physik. Band 4: Spezielle Relativitätstheorie. Thermodynamik*, 6. Aufl. (Springer, Berlin & Heidelberg, 2005).
- [33] K. Schwarzschild, *Nachr. Ges. Wiss. Göttingen* (1903) 126–131.
- [34] J.-M. Lévy-Leblond, *Am. J. Phys.* **39** (1971) 502–506.
- [35] M. Postnikov, *Lectures in Geometry. Semester II: Linear Algebra and Differential Geometry* (Mir Publishers, Moscow, 1982).
- [36] I. M. Gelfand, G. E. Shilov, *Verallgemeinerte Funktionen (Distributionen) Band 1*, Band 47 von *Hochschulbücher für Mathematik*, 2. Aufl. (Deutscher Verlag der Wissenschaft, Berlin, 1967).
- [37] L. Schwartz, *Mathematische Methoden der Physik. Band 1* (Bibliographisches Institut, Mannheim, 1974).