

# Inhaltsverzeichnis

<b>1</b>	<b>Lagrange Formalismus</b>	<b>5</b>
1.1	Zwangsbedingungen . . . . .	5
1.2	Das D'Alembertsche Prinzip . . . . .	8
1.3	Die Lagrange Bewegungsgleichungen . . . . .	14
1.4	Anwendungsbeispiele zum Lagrange Formalismus . . . . .	19
1.4.1	Sphärisches Pendel . . . . .	19
1.4.2	Starrer Körper . . . . .	21
1.4.3	Geschwindigkeitsabhängige Kraft . . . . .	25
1.5	Euler-Lagrangesche Variationsgleichungen . . . . .	27
1.5.1	Brachistochronen Problem . . . . .	30
1.6	Hamiltonsches Prinzip . . . . .	33
1.7	Erhaltungssätze und Symmetrien I . . . . .	36
1.8	Harmonische Näherung . . . . .	40
1.8.1	Kleine Schwingungen im Lagrange Formalismus . . . . .	40
1.8.2	Beispiel: Modell des Wassermoleküls . . . . .	46
1.8.3	Die Lösung des Beispiels mit Maple . . . . .	49
<b>2</b>	<b>Hamilton Formalismus</b>	<b>53</b>
2.1	Hamiltonsche Bewegungsgleichungen . . . . .	53
2.2	Bedeutung der Hamiltonfunktion und Beispiele . . . . .	58
2.3	Kanonische Transformationen . . . . .	67
2.4	Poisson Klammern . . . . .	74
2.5	Noether Theorem: Symmetrien und Konstanten . . . . .	79
2.5.1	Infinitesimale Drehungen und Drehungen um einen vorgegebenen Winkel . . . . .	84
<b>3</b>	<b>Nichtlineare Systeme, Chaos</b>	<b>87</b>
3.1	Logistische Gleichung . . . . .	87

3.2	Attraktoren in physikalischen Systemen . . . . .	93
3.3	Der Satz von Liouville . . . . .	103
3.4	Überlegungen zur Feldtheorie . . . . .	109
<b>4</b>	<b>Einführung in die Quantenmechanik</b>	<b>111</b>
4.1	Grenzen der klassischen Physik: “Die Physiker spinnen” . . . . .	111
4.2	Experimentelle Hinweise auf die Quantenmechanik . . . . .	114
4.2.1	Der Fotoeffekt . . . . .	114
4.2.2	Schwarzkörperstrahlung . . . . .	116
4.2.3	Compton Effekt . . . . .	119
4.2.4	Doppelspaltexperiment mit Elektronen . . . . .	121
4.2.5	Stationäre Zustände . . . . .	123
4.3	Zusammenfassung . . . . .	124
<b>5</b>	<b>Wellenmechanik eindimensionaler Systeme</b>	<b>129</b>
5.1	Wellenpakete und Erwartungswerte . . . . .	129
5.2	Schrödinger Gleichung und Ehrenfest’sches Theorem . . . . .	139
5.3	Wahrscheinlichkeitsstrom . . . . .	143
5.4	Stationäre Lösungen der Schrödingergleichung . . . . .	147
5.4.1	Tunneleffekt . . . . .	151
5.4.2	Beschreibung von bewegten Teilchen . . . . .	153
5.5	Der Harmonische Oszillator I . . . . .	155
<b>6</b>	<b>Grundlagen der Quantenmechanik</b>	<b>169</b>
6.1	Zustände und Observable . . . . .	169
6.1.1	$\mathcal{H}$ ein linearer Vektorraum über $\mathcal{C}$ . . . . .	170
6.1.2	$\mathcal{H}$ ein unitärer Vektorraum . . . . .	171
6.1.3	Orts- und Impulsdarstellung . . . . .	176
6.1.4	Dynamische Variable . . . . .	177
6.2	Eigenzustände und Matrixdarstellung . . . . .	182
6.2.1	Matrixdarstellung . . . . .	185
6.3	Messung von Observablen . . . . .	188
<b>7</b>	<b>Zentrale Kraftfelder</b>	<b>197</b>
7.1	Drehimpuls in der Ortsdarstellung . . . . .	197
7.2	Zentralfeldprobleme in der Quantenmechanik . . . . .	207

7.3 Das Wasserstoffatom . . . . . 210



# Kapitel 1

## Lagrange Formalismus

### 1.1 Zwangsbedingungen

Behandelt man die Bewegung von  $N$  Teilchen, die sich im 3-dimensionalen Raum bewegen können, so ist die Position dieser Teilchen durch  $3N$  Koordinaten (jeweils 3 für den Ortsvektor  $\vec{r}_i$  eines Teilchens  $i$ ) festgelegt. Die Newtonschen Bewegungsgleichungen für ein solches System bestehen aus  $3N$  Differentialgleichungen zweiter Ordnung für diese Koordinaten. Zur Bestimmung einer eindeutigen Lösungen benötigen wir also 2 mal  $3N$  Anfangsbedingungen, also z.B. die Vektoren für die Positionen und die Geschwindigkeiten dieser  $N$  Teilchen zur Zeit  $t = 0$ . Diese Anfangsbedingungen definieren den Startpunkt der Bewegung im  $6N$  dimensionalen Phasenraum des Problems. Durch die Lösung der  $3N$  Newtonschen Bewegungsgleichungen

$$\frac{d^2}{dt^2} m_i \vec{r}_i = \vec{F}_i(\vec{r}_1, \dots, \vec{r}_N, t) \quad \text{für } i = 1 \dots N, \quad (1.1)$$

erhalten wir die Funktionen für die Positionen der Teilchen der Masse  $m_i$ ,  $\vec{r}_i(t)$ , und ihrer Geschwindigkeiten

$$\vec{v}_i(t) = \frac{d}{dt} \vec{r}_i(t) = \dot{\vec{r}}_i(t).$$

Die Sequenz dieser Punkte  $(\vec{r}_i(t), \vec{v}_i(t))$  im  $6N$  dimensionalen Phasenraum bezeichnet man als die Trajektorie für die zeitliche Entwicklung des Systems.

Man sagt auch: Das System aus  $N$  Teilchen, die sich beliebig im Raum bewegen können, besitzt  $3N$  Freiheitsgrade. Die Zahl  $3N$  bezieht sich auf die Zahl der unabhängigen Koordinaten, die die Position des Teilchen beschreiben.

Diese Zahl der Freiheitsgrade kann eingeschränkt sein. Stellen wir uns als Beispiel ein System aus 2 Massenpunkten vor, die stets einen fest vorgegebenen Abstand  $a$  voneinander haben sollen. Diese Beschränkung der Bewegungsfreiheit der Massenpunkte, man spricht von einer **Zwangsbedingung**, könnte durch eine Verbindungsstange realisiert sein, deren Masse bei der Bewegung vernachlässigbar sein soll. Mathematisch ist diese Zwangsbedingung eindeutig formuliert durch die Gleichung

$$|\vec{r}_2 - \vec{r}_1| - a = 0. \quad (1.2)$$

Allgemein, im Fall von  $N$  Massenpunkten, definiert eine Gleichung vom Typ

$$f(\vec{r}_1, \dots, \vec{r}_N) = 0 \quad (1.3)$$

eine **holonome** Zwangsbedingung. Gilt insbesondere, dass die Zwangsbedingung unabhängig von der Zeit ist, so bezeichnet man die Zwangsbedingung (1.3) als Gleichung für eine **holonome, skleronome** (oder zeitunabhängige) Zwangsbedingung. Wäre die Zwangsbedingung zeitabhängig, wäre also etwa in dem Beispiel der beiden Massenpunkte aus (1.2) der Abstand  $a$  abhängig von der Zeit  $t$ , so hat man eine Zwangsbedingung vom Typ

$$f(\vec{r}_1, \dots, \vec{r}_N, t) = 0. \quad (1.4)$$

In diesem Fall spricht man von einer **holonomen, rheonomen** (oder eben zeitabhängigen) Zwangsbedingung.

Neben diesen holonomen Zwangsbedingungen gibt es aber auch nicht-holonome Zwangsbedingungen. Als Beispiel stellen wir uns zwei Massen vor, die als harte Kugeln mit jeweils einem Durchmesser  $d$  realisiert sind. In diesem Fall müssen die Ortsvektoren der beiden Kugelzentren einen Abstand haben, der größer ist als  $d$ . Wir haben also eine Zwangsbedingung der Form

$$|\vec{r}_2 - \vec{r}_1| - d > 0. \quad (1.5)$$

Allgemein definiert eine Gleichung vom Typ

$$f(\vec{r}_1, \dots, \vec{r}_N) > 0 \quad (1.6)$$

eine **nicht holonome** Zwangsbedingung. Auch in diesem Fall kann man natürlich wieder zwischen skleronomen und rheonomen Bedingungen unterscheiden.

Wir werden uns in der Mechanik vor allen Dingen mit holonomen Zwangsbedingungen beschäftigen. Durch eine holonome Zwangsbedingung wird die Zahl der Freiheitsgrade eines Systems um einen Freiheitsgrad reduziert. Ein Freiheitsgrad ist durch die Zwangsbedingung eingefroren oder eliminiert worden. Gibt es  $k$  Zwangsbedingungen von diesem Typ, also

$$f_\alpha(\vec{r}_1, \dots, \vec{r}_N) = 0, \quad \text{für } \alpha = 1 \dots k, \quad (1.7)$$

so wird die Zahl der Freiheitsgrade auf

$$3N - k$$

reduziert. Wichtig ist allerdings dabei, dass die Zwangsbedingungen wirklich unabhängig voneinander sind. Stellen wir uns als Beispiel 3 Massenpunkte vor, also  $N = 3$ . Die Relativvektoren dieser Massenpunkte seien durch die Zwangsbedingungen

$$\begin{aligned} \vec{r}_1 - \vec{r}_2 - \vec{a} &= 0 \\ \vec{r}_1 - \vec{r}_3 - \vec{b} &= 0 \\ \vec{r}_2 - \vec{r}_3 - \vec{c} &= 0 \end{aligned}$$

eingeschränkt. Jede dieser 3 Gleichung ist eine Vektorgleichung wir haben also insgesamt 9 holonome Zwangsbedingungen. Naiv könnte man also jetzt argumentieren, dass die  $3N$  Freiheitsgrade durch die  $k = 9$  Zwangsbedingungen gerade auf null Freiheitsgrade

reduziert werden. Bei genauerem Hinsehen stellt man aber natürlich fest, dass sich die dritte der 3 oben angegebenen Zwangsbedingungen aus den ersten beiden ergibt mit  $\vec{c} = \vec{b} - \vec{a}$ . In Wirklichkeit gibt es also nur 2 mal 3 also 6 unabhängige Zwangsbedingungen. (Vergleiche dazu auch die Definition eines Starren Körpers und seiner Freiheitsgrade.)

Das Problem ist es nun, die  $3N - k$  Koordinaten zu finden, die diese Freiheitsgrade beschreiben. Solche Koordinaten, für die dann keine Zwangsbedingungen berücksichtigt werden müssen, nennt man **generalisierte Koordinaten**,  $q_i$  für  $i = 1 \dots 3N - k$ . Die Anforderungen an diese generalisierten Koordinaten lauten also:

- Mit den generalisierten Koordinaten werden alle Zustände des Systems, die mit den Zwangsbedingungen verträglich sind, eindeutig beschrieben.
- Es gibt keine holonomen Zwangsbedingungen für die generalisierten Koordinaten.

Als Beispiel betrachten wir die beiden Massenpunkte mit der holonomen Zwangsbedingung (1.2). Es gilt nun, für  $N = 2$  Teilchen und  $k = 1$  Zwangsbedingungen  $3N - k$ , also 5, generalisierte Koordinaten zu finden. In diesem Fall betrachten wir als ersten Schritt die Transformation der Ortsvektoren  $\vec{r}_1$  und  $\vec{r}_2$  auf den Schwerpunktsvektor  $\vec{R}$  des Systems

$$\vec{R} = \frac{m_1 \vec{r}_1 + m_2 \vec{r}_2}{m_1 + m_2}, \quad (1.8)$$

und den Relativvektor

$$\vec{r} = \vec{r}_2 - \vec{r}_1.$$

Die Zwangsbedingung (1.2) besagt, dass der Betrag dieses Relativvektors konstant zu halten ist, nämlich identisch  $a$ . Stellt man den Relativvektor in Kugelkoordinaten dar, so ist er mit den entsprechenden Winkeln  $\vartheta$  und  $\varphi$  eindeutig festgelegt. Wählen wir also als Koordinaten des System 3 Koordinaten zur Beschreibung von  $\vec{R}$  und diese beiden Winkel zur Festlegung von  $\vec{r}$ , so können alle Zustände des Systems durch Angabe dieser 5 Koordinaten eindeutig festgelegt werden. Andererseits gibt es für diese Koordinaten keine Zwangsbedingungen, sie können also als generalisierte Koordinaten genutzt werden.

In diesem Fall war es sehr einfach, generalisierte Koordinaten zu identifizieren. In anderen Fällen ist die Bestimmung von geeigneten generalisierten Koordinaten das zentrale Problem. Hat man generalisierte Koordinaten definiert, gilt es dann, entsprechende Bewegungsgleichungen aufzustellen (wir werden dies später bei der Diskussion der Lagrange Bewegungsgleichungen zweiter Art machen) und zu lösen.

## 1.2 Das D'Alembertsche Prinzip

Wir betrachten wieder ein System bestehend aus  $N$  Massenpunkten, deren Bewegung eventuell durch  $k$  Zwangsbedingungen eingeschränkt sein sollen. Die Ortsvektoren dieser  $N$  Teilchen sind dann durch jeweils  $3N - k$  Koordinaten  $q_1 \dots q_{3N-k}$  eindeutig definiert. Sind die holonomen Zwangsbedingungen für das System explizit zeitabhängig, so wird auch der Zusammenhang zwischen den Ortsvektoren der Teilchen und den generalisierten Koordinaten von der Zeit abhängen:

$$\vec{r}_i(q_1, \dots, q_{3N-k}, t).$$

Wir definieren nun eine **virtuelle Verrückung** als eine infinitesimale Veränderung des Systems, die mit den Zwangsbedingungen zur gegebenen Zeit verträglich ist. Der Ortsvektor des Teilchens  $i$  ändert sich also bei einer solchen virtuellen Verrückung um

$$\delta \vec{r}_i = \sum_{\alpha=1}^{3N-k} \frac{\partial \vec{r}_i}{\partial q_\alpha} \delta q_\alpha. \quad (1.9)$$

Dabei stehen die  $\delta q_\alpha$  für eine infinitesimale Kombination der Änderung der  $3N - k$  generalisierten Koordinaten des Systems.

Zur Verdeutlichung dieser Definition der virtuellen Verrückung betrachten wir einen Käfer, der sich auf dem Boden eines Aufzuges frei in der  $xy$ -Ebene bewegen kann. Der Aufzug selbst bewege sich mit der Geschwindigkeit  $v$  in  $z$ -Richtung aufwärts. Die Bewegung des Käfers unterliegt also einer zeitabhängigen Zwangsbedingung der Form

$$z - vt = 0 \quad (1.10)$$

Eine virtuelle Verrückung für die Bewegung des Käfers muss berücksichtigen, dass der Wert für die  $z$ -Komponente des Käfers durch die Zwangsbedingung (1.10) festgelegt ist, sie hat also die Form

$$\delta \vec{r} = \delta x \hat{e}_x + \delta y \hat{e}_y$$

Bei der Bewegung des Käfers in  $xy$ -Richtung bewegt sich aber gleichzeitig der Aufzug das Stück  $v \delta t$  in  $z$ -Richtung, sodass die reale Bewegung des Käfers die Form

$$d\vec{r} = \delta x \hat{e}_x + \delta y \hat{e}_y + v \delta t \hat{e}_z$$

annimmt.

Im Beispiel des Käfers im Aufzug sorgt die Kraft, die vom Boden des Aufzugs auf den Käfer ausgeübt wird, dafür, dass die Zwangsbedingung erfüllt wird. Allgemein definieren wir all die Kräfte, die die Einhaltung der Zwangsbedingungen herbeiführen, als **Zwangskräfte**.

Die Wirkung und die Bedeutung dieser Zwangskräfte soll an einem weiteren Beispiel verdeutlicht werden. In diesem Fall betrachten wir die Bewegung eines Massenpunktes mit der Masse  $m$  im Gravitationsfeld der Erde, gegeben durch die Kraft (siehe Abb.1.1)

$$\vec{F} = -mg \hat{e}_y. \quad (1.11)$$



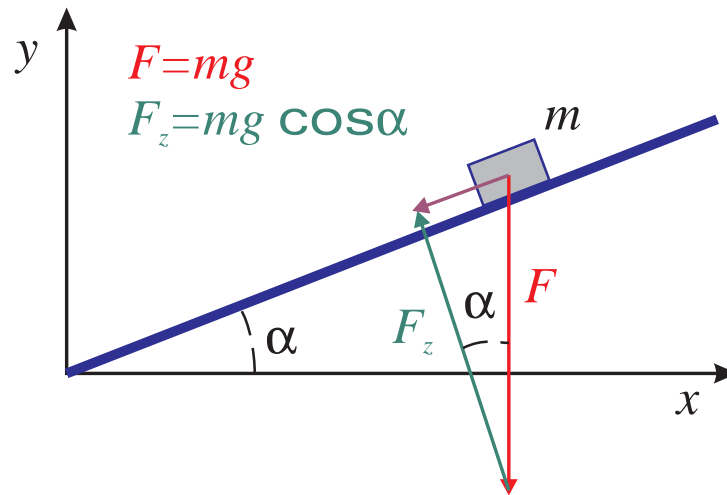


Abbildung 1.1: Kräfte auf einen Massenpunkt der Masse  $m$  auf einer schiefen Ebene

Dieser Massenpunkt soll sich allerdings nicht frei bewegen sondern auf einer schiefen Ebene, die mit der  $x$ -Achse den Winkel  $\alpha$  bildet. Die  $z$ -Komponenten der Vektoren spielen für diese Diskussion keine Rolle und werden deshalb nicht beachtet. Die Beschränkung der Bewegung wird durch die Zwangsbedingung

$$f(x, y) = x \sin \alpha - y \cos \alpha = 0 \quad (1.12)$$

beschrieben. Damit die Zwangskraft gewährleistet, dass die Bewegung nur unter Einhaltung der Zwangsbedingung, also entlang der schiefen Ebene erfolgt, muss sie den Betrag und Richtung haben, wie in Abb. 1.1 skizziert. Das bedeutet: die Zwangskraft ist definiert durch

$$\begin{aligned} |\vec{F}^{\text{Zwang}}| &= mg \cos \alpha \\ \vec{F}^{\text{Zwang}} &= mg \cos \alpha \{ \cos \alpha \hat{e}_y - \sin \alpha \hat{e}_x \} . \end{aligned} \quad (1.13)$$

Die Zwangskräfte sind also so zu konstruieren, dass sich das Teilchen  $i$  bewegt als ob auf dieses Teilchen eine Gesamtkraft wirken würde

$$\vec{F}_i^{\text{Gesamt}} = \vec{F}_i + \vec{F}_i^{\text{Zwang}} \quad (1.14)$$

die neben der explizit wirkenden Kraft  $\vec{F}_i$  auch die Zwangskraft enthält, die für die Beachtung der Zwangsbedingungen sorgt. Wenn man die Zwangskräfte kennt, so kann man die Bewegung der Teilchen des Systems durch die Lösung der Newtonschen Bewegungsgleichungen

$$\vec{F}_i^{\text{Gesamt}} - \dot{\vec{p}}_i = 0 \quad \text{für } i = 1 \dots N, \quad (1.15)$$

berechnen. Natürlich kann man jede dieser  $N$  Bewegungsgleichungen mit einem Vektor  $\Delta \vec{r}_i$  multiplizieren und dann aufaddieren, was zu einer Gleichung

$$\sum_{i=1}^N \left\{ (\vec{F}_i - \dot{\vec{p}}_i) \Delta \vec{r}_i + \vec{F}_i^{\text{Zwang}} \Delta \vec{r}_i \right\} = 0, \quad (1.16)$$

führt. Eine weitere Behandlung erlaubt nun das **D'Alembertsche Prinzip**, welches als ein Grundpostulat für die Beschreibung der Bewegung von Massenpunkten unter Zwangsbedingungen darstellt.

### D'Alembertsche Prinzip:

Die zeitliche Entwicklung des Systems erfolgt in Richtung solcher virtuellen Verrückungen  $\delta\vec{r}_i$ , für die gilt:

$$\sum_{i=1}^N \vec{F}_i^{\text{Zwang}} \delta\vec{r}_i = 0. \quad (1.17)$$

Zu diesem D'Alembertschen Prinzip einige Anmerkungen:

- Das D'Alembertsche Prinzip besagt also, dass nur solche Verrückungen  $\delta\vec{r}_i$  in Betracht kommen, bei denen die Zwangskräfte keine Arbeit verrichten. Anders ausgedrückt: die Zwangskräfte stehen senkrecht zur Ebene der erlaubten virtuellen Verrückungen. Dies ist natürlich auch leicht einzusehen: die Zwangskräfte sollen ja gerade solche Verrückungen zulassen, bei denen alle Bewegungen, bei denen die Zwangsbedingungen eingehalten werden, wie freie Bewegungen ablaufen. Andererseits aber sollen all solche Bewegungen, bei denen die Zwangsbedingungen verletzt werden, vollständig unterdrückt sein.
- Das oben aufgeführte Beispiel des Käfers im Aufzug zeigt uns auf, warum wir die virtuellen Verrückungen benötigen, um das D'Alembertsche Prinzip zu formulieren. Die Zwangskraft in diesem Beispiel steht in Richtung der  $z$ -Achse und damit senkrecht zu den virtuellen Verrückungen in der  $xy$ -Ebene. Die realen Verrückungen besitzen aber eine Komponente in  $z$ -Richtung, sodass die Zwangskräfte entlang der realen Verrückungen Arbeit leisten, eben gerade die Energie, die erforderlich ist, den Käfer mit dem Aufzug hinaufzutransportieren.
- Bei einem System mit mehreren Massenpunkten können wir virtuelle Verrückungen, die mit den Zwangsbedingungen verträglich sind, nicht für jeden Massenpunkt unabhängig identifizieren. Ist z.B. der Abstand zweier Massenpunkt durch eine Zwangsbedingung fixiert, so hängt die erlaubte virtuelle Verrückung  $\delta\vec{r}_2$  des zweiten Punktes von der gleichzeitigen Bewegung des ersten Punktes. Deshalb kann das D'Alembertsche Prinzip natürlich nicht für jeden Massenpunkt isoliert formuliert werden (das wäre also die Gleichung (1.17) für jeden Punkt  $i$  ohne die Summe). Die Forderung, dass die Zwangskräfte senkrecht zu den erlaubten virtuellen Verrückungen steht muss also durch das Skalarprodukt der Zwangskräfte mit den Verrückungen im  $3N$  dimensionalen Raum der Koordinaten aller  $N$  Teilchen formuliert sein, so wie das auch in (1.17) der Fall ist.

Wie kann man nun die erlaubten virtuellen Verrückungen  $\delta\vec{r}_i$ , beziehungsweise die entsprechenden Zwangskräfte identifizieren? Ausgangspunkt dazu sind natürlich die Zwangsbedingungen, die für die Koordinaten der Teilchen vor und nach der Verrückung erfüllt

sein müssen. Es muss also für alle  $l = 1 \dots k$  gelten

$$f_l(\vec{r}_1 \dots \vec{r}_N, t) = f_l(\vec{r}_1 + \delta\vec{r}_1 \dots \vec{r}_N + \delta\vec{r}_N, t) = 0. \quad (1.18)$$

Da es sich bei den virtuellen Verrückungen um infinitesimale Verrückungen handelt, können wir die Zwangsbedingung an der Stelle  $\vec{r}_i + \delta\vec{r}_i$  entwickeln um den Punkt  $\vec{r}_i$  und die Entwicklung nach der ersten Potenz in den  $\delta\vec{r}_i$  abrechnen:

$$f_l(\vec{r}_1 + \delta\vec{r}_1 \dots \vec{r}_N + \delta\vec{r}_N, t) = f_l(\vec{r}_1 \dots \vec{r}_N, t) + \sum_{i=1}^N \frac{\partial f_l}{\partial \vec{r}_i} \delta\vec{r}_i. \quad (1.19)$$

Dabei steht die Ableitung nach dem Vektor für

$$\begin{aligned} \frac{\partial f_l}{\partial \vec{r}_i} \delta\vec{r}_i &= \frac{\partial f_l}{\partial x_i} \delta x_i + \frac{\partial f_l}{\partial y_i} \delta y_i + \frac{\partial f_l}{\partial z_i} \delta z_i \\ &= \vec{\nabla}_i f_l \delta\vec{r}_i. \end{aligned} \quad (1.20)$$

Mit dieser Nomenklatur ergeben sich also aus (1.18) und (1.19) die  $k$  (entspricht der Anzahl der Zwangsbedingungen) Bestimmungsgleichungen für die Verrückungen  $\delta\vec{r}_i$

$$\sum_{i=1}^N \vec{\nabla}_i f_l \delta\vec{r}_i = 0 \quad \text{für } l = 1 \dots k. \quad (1.21)$$

Da die Zwangsbedingungen und damit auch diese  $k$  Gleichungen die erlaubten virtuellen Verrückungen eindeutig festlegen, kann die Gleichung des D'Alembertschen Prinzips (1.17), die ja auch eine Gleichung linear in den  $\delta\vec{r}_i$  ist, keine zusätzliche Bedingung an die virtuellen Verrückungen darstellen. Sie muss vielmehr eine Linearkombination der Bestimmungsgleichungen in (1.21) sein, also vom Typ

$$\sum_{i=1}^N \vec{F}_i^{\text{Zwang}} \delta\vec{r}_i = \sum_{l=1}^k \lambda_l \sum_{i=1}^N \vec{\nabla}_i f_l \delta\vec{r}_i, \quad (1.22)$$

mit Konstanten,  $\lambda_l$ , die noch zu bestimmen sind. Aus dieser Gleichung entnehmen wir eine Darstellung für die Zwangskraft auf das Teilchen  $i$  in der Form

$$\vec{F}_i^{\text{Zwang}} = \sum_{l=1}^k \lambda_l \vec{\nabla}_i f_l. \quad (1.23)$$

Setzen wir diese Darstellung der Zwangskräfte in die Bewegungsgleichung (1.15) ein, so erhalten wir die

### Lagrange Bewegungsgleichungen 1.Art:

$$\dot{\vec{p}}_i = \vec{F}_i + \sum_{l=1}^k \lambda_l \vec{\nabla}_i f_l \quad i = 1 \dots N \quad (1.24)$$

$$f_l(\vec{r}_i, t) = 0 \quad k = 1 \dots l. \quad (1.25)$$

Bei diesen Gleichungen handelt es sich um  $N$  Vektor-Differenzialgleichungen (1.24) und  $k$  algebraische Gleichungen (1.25). Diese Zahl von  $3N + k$  unabhängigen Gleichungen ist notwendig und hinreichend um die Unbekannten,  $N$  Vektorfunktionen  $\vec{r}_i(t)$  und  $k$  Parameter  $\lambda_l(t)$  eindeutig zu bestimmen. Die Gleichungen besitzen also bei vorgegebenen Startwerten für  $\vec{r}_i(t_0)$  und  $\vec{v}_i(t_0)$  eine eindeutige Lösung. Neben dem Ablauf der Bewegung, dargestellt durch  $\vec{r}_i(t)$ , erhält man auch die Koeffizienten  $\lambda_l$  und mit (1.23) die Zwangskräfte, die auf die einzelnen Massenpunkte wirken.

Als ein Beispiel für die Lösung eines Problems mit Zwangsbedingungen unter Benutzung der Lagrange Bewegungsgleichungen erster Art greifen wir das Beispiel des Massenpunktes auf der schiefen Ebene, dargestellt in Abb. 1.1, auf. Für diesen Massenpunkt haben wir zwei Zwangsbedingungen

$$f_1(x, y, z) = x \sin \alpha - y \cos \alpha = 0 \quad (1.26)$$

$$f_2(x, y, z) = z = 0 \quad (1.27)$$

Die Gradienten dieser beiden Zwangsbedingungen berechnen sich zu

$$\vec{\nabla} f_1 = \sin \alpha \hat{e}_x - \cos \alpha \hat{e}_y$$

$$\vec{\nabla} f_2 = \hat{e}_z$$

und damit erhalten wir für die Zwangskraft den Ansatz

$$\vec{F}^{\text{Zwang}} = \lambda_1 (\sin \alpha \hat{e}_x - \cos \alpha \hat{e}_y) + \lambda_2 \hat{e}_z \quad (1.28)$$

Mit der Kraft durch das Gravitationsfeld der Erde (1.11) ergibt sich also die Bewegungsgleichung vom Typ (1.24) zu

$$\begin{pmatrix} m\ddot{x} \\ m\ddot{y} \\ m\ddot{z} \end{pmatrix} = \begin{pmatrix} \lambda_1 \sin \alpha \\ -mg - \lambda_1 \cos \alpha \\ \lambda_2 \end{pmatrix}. \quad (1.29)$$

Diese 3 Gleichungen bilden zusammen mit den 2 Zwangsbedingungen (1.26) und (1.27) die 5 Lagrangegleichungen 1. Art, die wir benötigen um die 5 Unbekannten  $\lambda_1$ ,  $\lambda_2$ ,  $x$ ,  $y$  und  $z$  zu bestimmen. Die Gleichungen für  $z(t)$  und  $\lambda_2$ , das sind die dritte Komponente von (1.29) und (1.27) entkoppeln von den anderen Gleichungen und werden gelöst durch

$$z(t) = \dot{z}(t) = \ddot{z}(t) = 0,$$

und wir können die Beschreibung der Bewegung auf die  $xy$  Ebene beschränken. Multipliziert man die erste Komponente von (1.29) mit einem Faktor  $\sin \alpha$  und subtrahiert davon die  $y$ -Komponente dieser Vektorgleichung multipliziert mit  $\cos \alpha$ , so ergibt sich die resultierende Gleichung

$$m(\ddot{x} \sin \alpha - \ddot{y} \cos \alpha) = \lambda_1 \sin^2 \alpha + mg \cos \alpha + \lambda_1 \cos^2 \alpha = \lambda_1 + mg \cos \alpha. \quad (1.30)$$

Differenziert man die Zwangsbedingung (1.26) zwei mal nach der Zeit und setzt das Ergebnis

$$\ddot{x} \sin \alpha - \ddot{y} \cos \alpha = 0$$

in (1.30) ein, so erhält man

$$\lambda_1 = -mg \cos \alpha .$$

Zusammen mit  $\lambda_2 = 0$  ergibt sich damit nach (1.28) für die Zwangskraft

$$\vec{F}^{\text{Zwang}} = -mg \cos \alpha (\sin \alpha \hat{e}_x - \cos \alpha \hat{e}_y) .$$

Wir erhalten also durch konsequentes Auswerten der Lagrange Bewegungsgleichungen den gleichen Ausdruck für die Zwangskraft wie in (1.13), wo in diesem einfachen Fall der schiefen Ebene die Zwangskraft aus geometrischen Überlegungen bestimmt wurde. Die Bewegung des Massenpunktes ergibt sich, wenn wir nun die erste Komponente von (1.29) mit einem Faktor  $\cos \alpha$  multiplizieren und darauf die zweite Komponente multipliziert mit  $\sin \alpha$  addieren. Dies führt zu

$$m(\ddot{x} \cos \alpha + \ddot{y} \sin \alpha) = -mg \sin \alpha .$$

Benutzt man nun die Abkürzung

$$q = x \cos \alpha + y \sin \alpha \tag{1.31}$$

so erhält man für diese **generalisierte Koordinate** die Bewegungsgleichung

$$m\ddot{q} = -mg \sin \alpha . \tag{1.32}$$

Diese Differenzialgleichung hat die Lösung

$$q(t) = q_0 + \dot{q}_0 t - \frac{1}{2} mg \sin \alpha t^2 \tag{1.33}$$

wobei die Konstanten  $q_0$  und  $\dot{q}_0$ , also der Wert der generalisierten Koordinate und seine Geschwindigkeit zur Zeit  $t = 0$ , durch die Anfangsbedingungen festgelegt sind. Ersetzt man nun in (1.31) z.B. die Koordinate  $x$  durch  $y \cos \alpha / \sin \alpha$ , was ja wegen der Zwangsbedingung (1.26) gilt, so ergibt sich

$$y(t) = \frac{q(t)}{\frac{\cos^2 \alpha}{\sin \alpha} + \sin \alpha} .$$

In analoger Weise ergibt sich auch

$$x(t) = \frac{q(t)}{\frac{\sin^2 \alpha}{\cos \alpha} + \cos \alpha} .$$

Damit sind also über die Bestimmung der Lösung für die generalisierte Koordinate  $q(t)$  in (1.33) auch die kartesischen Koordinaten des Teilchens berechnet.

Die Lösung der Lagrangeschen Bewegungsgleichungen erster Art für ein System mit  $N$  Massenpunkten und  $k$  Zwangsbedingungen ist recht aufwändig, da es auf ein System von  $3N + k$  Gleichungen führt. Durch die Zwangsbedingungen wird also die Zahl der Gleichungen nicht reduziert sondern noch erhöht. Als Lohn für diese Mühe bekommt man aber ein Ergebnis, bei dem auch die Zwangskräfte berechnet sind. Diese Verfahren bietet sich also an, wenn man die Stärke und Richtung dieser Zwangskräfte interessiert ist.

### 1.3 Die Lagrange Bewegungsgleichungen

Auch in diesem Abschnitt behandeln wir wieder das System bestehend aus  $N$  Massenpunkten, deren Bewegungen durch  $k$  Zwangsbedingungen eingeschränkt sein sollen. Hier interessieren wir uns aber nicht für die eventuell auftretenden Zwangskräfte sondern wollen lediglich die zeitliche Entwicklung des Systems, also die Bewegung der Massenpunkte bestimmen. Dazu nehmen wir an, dass es uns gelungen ist einen Satz von  $3N - k$  generalisierten Koordinaten zu bestimmen. Mit diesen generalisierten Koordinaten  $q_1 \dots q_{3N-k}$  können wir dann alle Konfigurationen des Systems, also alle Positionen der Teilchen  $\vec{r}_i$ , definieren, die mit den Zwangsbedingungen verträglich sind. Wir müssen dazu nur angeben wie sich die Ortsvektoren der einzelnen Teilchen als Funktion der generalisierten Koordinaten bestimmen lassen

$$\vec{r}_i(q_1, \dots, q_{3N-k}, t). \quad (1.34)$$

Wir lassen dabei den Fall zu, dass die Zwangsbedingungen rheonom also zeitabhängig sind, was natürlich dazu führt, dass auch der Zusammenhang, wie man die Ortsvektoren  $\vec{r}_i$  bei vorgegebenen Werten für die generalisierten Koordinaten auszurechnen hat, von der Zeit abhängt. Dies wird dadurch dargestellt, dass die Funktion  $\vec{r}_i$  in (1.34) explizit von der Zeit abhängt. Ausserdem werden sich natürlich bei einer Bewegung des Systems die Werte der generalisierten Koordinaten  $q_j$  mit der Zeit ändern, so dass die totale Ableitung eines Ortsvektors  $\vec{r}_i$  nach der Zeit, das ist ja gerade die Geschwindigkeit  $\vec{v}_i$  des Teilchens  $i$  berechnet wird mit

$$\begin{aligned} \dot{\vec{r}}_i = \vec{v}_i &= \frac{d\vec{r}_i}{dt} = \sum_{j=1}^{3n-k} \frac{\partial \vec{r}_i}{\partial q_j} \frac{dq_j}{dt} + \frac{\partial \vec{r}_i}{\partial t} \\ &= \sum_{j=1}^{3n-k} \frac{\partial \vec{r}_i}{\partial q_j} \dot{q}_j + \frac{\partial \vec{r}_i}{\partial t}. \end{aligned} \quad (1.35)$$

Die virtuellen Verrückungen, die mit den Zwangsbedingungen verträglich sind, lassen sich dann darstellen in der Form

$$\delta \vec{r}_i = \sum_{j=1}^{3n-k} \frac{\partial \vec{r}_i}{\partial q_j} \delta q_j. \quad (1.36)$$

Dabei bezeichnen die  $\delta q_j$  beliebige infinitesimale Änderungen der generalisierten Koordinaten  $q_j$ , da ja nach der Definition der generalisierten Koordinaten, diese keinen Beschränkungen unterliegen.

Wir multiplizieren nun die Bewegungsgleichungen der einzelnen Teilchen  $i$  (siehe (1.15)) mit dem zugehörigen Element der virtuellen Verrückung und addieren diese Gleichungen über alle Teilchen  $i$  des Systems. Dies liefert uns:

$$\sum_i m_i \ddot{\vec{r}}_i \delta \vec{r}_i = \sum_i \vec{F}_i \delta \vec{r}_i + \sum_i \vec{F}_i^{\text{Zwang}} \delta \vec{r}_i. \quad (1.37)$$

Wegen des D'Alembertschen Prinzips (1.17) verschwindet der zweite Term auf der rechten Seite dieser Gleichung und wir können mit Hilfe von (1.36) diese Gleichung umschreiben in

$$\sum_{i,j} \vec{F}_i \frac{\partial \vec{r}_i}{\partial q_j} \delta q_j = \sum_{i,j} m_i \ddot{\vec{r}}_i \frac{\partial \vec{r}_i}{\partial q_j} \delta q_j$$

$$= \sum_{i,j} \left\{ \frac{d}{dt} \left( m_i \dot{\vec{r}}_i \frac{\partial \vec{r}_i}{\partial q_j} \right) - m_i \dot{\vec{r}}_i \frac{d}{dt} \frac{\partial \vec{r}_i}{\partial q_j} \right\} \delta q_j. \quad (1.38)$$

Die zweite Gleichung kann man leicht durch Zurückrechnen unter Anwendung der Produktregel für die Zeitableitung verifizieren. Da auch der Vektor  $\partial \vec{r}_i / \partial q_j$  von den generalisierten Koordinaten  $q_k$  und explizit von der Zeit abhängen kann, ist

$$\begin{aligned} \frac{d}{dt} \frac{\partial \vec{r}_i}{\partial q_j} &= \sum_k \frac{\partial^2 \vec{r}_i}{\partial q_j \partial q_k} \frac{dq_k}{dt} + \frac{\partial}{\partial t} \frac{\partial \vec{r}_i}{\partial q_j} \\ &= \frac{\partial}{\partial q_j} \left\{ \sum_k \frac{\partial \vec{r}_i}{\partial q_k} \dot{q}_k + \frac{\partial \vec{r}_i}{\partial t} \right\} \\ &= \frac{\partial}{\partial q_j} \dot{\vec{r}}_i. \end{aligned} \quad (1.39)$$

Bei dem Übergang zur letzten Zeile haben wir die erste Zeile von (1.35) benutzt. Aus der zweiten Zeile dieser Gleichung (1.35) liest man ab, dass gilt:

$$\frac{\partial \dot{\vec{r}}_i}{\partial \dot{q}_j} = \frac{\partial \vec{r}_i}{\partial q_j}. \quad (1.40)$$

Setzt man die Ergebnisse von (1.39) und (1.40) in die Gleichung (1.38) ein, so ergibt sich unter Benutzung von  $\dot{\vec{r}}_i = \vec{v}_i$ :

$$\begin{aligned} \sum_{i,j} \vec{F}_i \frac{\partial \vec{r}_i}{\partial q_j} \delta q_j &= \sum_{i,j} \left\{ \frac{d}{dt} \left( m_i \vec{v}_i \frac{\partial \vec{v}_i}{\partial \dot{q}_j} \right) - m_i \vec{v}_i \frac{\partial \vec{v}_i}{\partial q_j} \right\} \delta q_j \\ &= \sum_{i,j} \left\{ \frac{d}{dt} \frac{\partial}{\partial \dot{q}_j} \left( \frac{1}{2} m_i \vec{v}_i^2 \right) - \frac{\partial}{\partial q_j} \left( \frac{1}{2} m_i \vec{v}_i^2 \right) \right\} \delta q_j. \end{aligned} \quad (1.41)$$

Auch in diesem Fall verifiziert man den Übergang zur zweiten Zeile der Gleichung am einfachsten dadurch, dass man zurück rechnet. Die kinetische Energie des Systems ist ja gegeben durch

$$T = \sum_i \frac{1}{2} m_i \vec{v}_i^2.$$

Ausserdem wissen wir, dass die infinitesimalen Verschiebungen  $\delta q_j$  beliebig sein können. Also gilt die Gleichung insbesondere auch wenn nur ein Element,  $\delta q_{j_0}$  von null verschieden ist. Dies bedeutet, dass wir aus (1.41) folgern können:

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} = \sum_i \vec{F}_i \frac{\partial \vec{r}_i}{\partial q_j} = Q_j. \quad (1.42)$$

Dabei bezeichnen wir  $Q_j$ , also die Kraft, die wirksam wird, wenn die generalisierte Koordinate  $q_j$  verändert wird als “ **Generalisierte Kraft** ”.

Zur weiteren Spezifizierung dieses Begriffes der generalisierten Kraft wollen wir zwei Fälle betrachten:

- Zunächst einmal den Fall eines konservativen Kraftfeldes, bei dem die Kraft auf ein Teilchen  $i$  durch den Gradienten eines Potentials, bezogen auf die Koordinaten dieses Teilchens  $i$  berechnet wird, also:

$$\vec{F}_i = -\vec{\nabla}_i V = -\frac{\partial V}{\partial \vec{r}_i}.$$

In diesem Fall berechnet sich die generalisierte Kraft

$$Q_j = \sum_i \vec{F}_i \frac{\partial \vec{r}_i}{\partial q_j} = - \sum_i \frac{\partial V}{\partial \vec{r}_i} \frac{\partial \vec{r}_i}{\partial q_j} = - \frac{\partial V}{\partial q_j}. \quad (1.43)$$

Wenn man also das Potenzial für die mit den Zwangsbedingungen verträglichen Positionen der Teilchen als Funktion der generalisierten Koordinaten  $q_j$  darstellt, so ergibt sich die generalisierte Kraft aus der Ableitung des Potentials nach der entsprechenden Koordinate  $q_j$ .

- Man kann dieses Konzept von Kräften, die durch Potentiale definiert sind aber auch noch erweitern auf bestimmte Kräfte, die von Geschwindigkeiten abhängen. In diesem Fall nehmen wir an, dass eine Potenzialfunktion  $V$  als Funktion der generalisierten Koordinaten  $q_j$  und der zugehörigen Geschwindigkeiten  $\dot{q}_j$  gegeben ist. Die geschwindigkeitsabhängige Kraft sei dann definiert durch

$$Q_j = \frac{d}{dt} \frac{\partial V}{\partial \dot{q}_j} - \frac{\partial V}{\partial q_j} \quad (1.44)$$

Es ist klar, dass dieser Fall eine Verallgemeinerung des konservativen Kraftfeldes darstellt. Für den Fall, dass  $V$  nicht von  $\dot{q}_j$  abhängt reduziert sich ja (1.44) auf den Fall (1.43). Wir werden aber am Ende dieses Abschnittes sehen, dass mit der Verallgemeinerung von (1.44) auf geschwindigkeitsabhängige Kräfte auch der wichtige Fall der Lorentz Kraft, also der Kraft eines Magnetfeldes auf eine bewegte Ladung, erfasst wird.

Setzt man nun die Darstellung der generalisierten Kraft für den allgemeinen Fall von (1.44) in (1.42) ein, so erhält man

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} = \frac{d}{dt} \frac{\partial V}{\partial \dot{q}_j} - \frac{\partial V}{\partial q_j},$$

oder auch

$$\frac{d}{dt} \frac{\partial (T - V)}{\partial \dot{q}_j} - \frac{\partial (T - V)}{\partial q_j} = 0. \quad (1.45)$$

Definieren wir also die **Lagrange Funktion** als Differenz aus kinetischer Energie des Systems  $T$  und der potenziellen Energie  $V$ :

$$L := T - V, \quad (1.46)$$

so können wir die Differenzialgleichungen (1.45) kompakt schreiben in der Form der

### Lagrange Bewegungsgleichungen 2.Art:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = 0. \quad (1.47)$$

Diese Gleichung gilt für alle  $3N - k$  generalisierten Koordinaten. Die Lösung dieser Gleichungen liefern zusammen mit den Anfangsbedingungen für die Bewegung des Systems



ein Ergebnis für die Werte der generalisierten Koordinaten  $q_j(t)$  als Funktion der Zeit. Da wir über (1.34) aus den aktuellen Werten für die generalisierten Koordinaten die Ortsvektoren für die einzelnen Teilchen  $\vec{r}_i$  berechnen können, ist damit das Problem, die zeitliche Entwicklung des Systems zu beschreiben gelöst.

Wir fassen also noch einmal die Schritte zur Beschreibung des Systems mit Hilfe der Lagrange Bewegungsgleichungen zusammen:

- Bestimme für das System aus  $N$  Teilchen mit  $k$  holonomen Zwangsbedingungen einen Satz von  $3N - k$  generalisierten Koordinaten.
- Berechne die kinetische Energie  $T$  und das Potenzial  $V$  als Funktion dieser generalisierten Koordinaten  $q_j$  und der zugehörigen Geschwindigkeiten  $\dot{q}_j$ . Dabei entspricht im Fall von konservativen Kräften das Potenzial  $V$  der potenziellen Energie des Systems. Bei geschwindigkeitsabhängigen Kräften, hängt auch  $V$  von den Geschwindigkeiten  $\dot{q}_i$  ab und die generalisierten Kräfte ergeben sich aus  $V$  entsprechend der Gleichung (1.44).
- Aus der Differenz  $T - V = L$  erhält man die Lagrange Funktion und kann die Lagrange Bewegungsgleichung zweiter Art (1.47) aufstellen.
- Die Lösung dieser Bewegungsgleichungen liefern zusammen mit Anfangsbedingungen ( $q_i(t = 0), \dot{q}_i(t = 0)$ ) die Werte für die generalisierten Koordinaten für alle Zeiten.
- Über (1.34) sind damit auch die Positionen der Teilchen des Systems bestimmt.

Wir haben in diesem Abschnitt die Lagrange Bewegungsgleichungen zweiter Art aus den Newtonschen Bewegungsgleichungen unter Benutzung des D'Alembertschen Prinzips hergeleitet. Wir wollen uns nun davon überzeugen, dass man andererseits auch aus den Lagrangeschen Bewegungsgleichungen die Newtonschen Bewegungsgleichungen herleiten kann. Als einfaches Beispiel betrachten wir dazu ein einzelnes Teilchen der Masse  $m$ , das sich ohne Zwangsbedingungen in einem konservativen Kraftfeld, das durch ein Potenzial  $V$  definiert ist, bewegt. Da keine Zwangsbedingungen vorliegen können wir die kartesischen Koordinaten des Teilchens als generalisierte Koordinaten benutzen. Das bedeutet

$$q_i = x_i = \{x, y, z\} \quad \text{für } i = \{1, 2, 3\} .$$

Die kinetische Energie ist dann gegeben durch

$$T = \frac{1}{2}m \sum_i \dot{x}_i^2$$

und

$$L = T - V = \frac{1}{2}m \sum_i \dot{x}_i^2 - V(x_i) .$$

Daraus ergibt sich für

$$\frac{\partial L}{\partial \dot{x}_i} = \frac{\partial T}{\partial \dot{x}_i} = m\dot{x}_i ,$$

und

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}_i} = m\ddot{x}_i. \quad (1.48)$$

Andererseits gilt

$$\frac{\partial L}{\partial x_i} = -\frac{\partial V}{\partial x_i}. \quad (1.49)$$

Bringt man diese Ergebnisse von (1.48) und (1.49) in die Lagrangsche Bewegungsgleichung (1.47) so ergibt sich

$$m\ddot{x}_i = -\frac{\partial V}{\partial x_i},$$

beziehungsweise in Vektorschreibweise

$$m\ddot{\vec{r}} = -\vec{\nabla}V,$$

also gerade die Newtonsche Bewegungsgleichung für ein Teilchen im konservativen Kraftfeld.

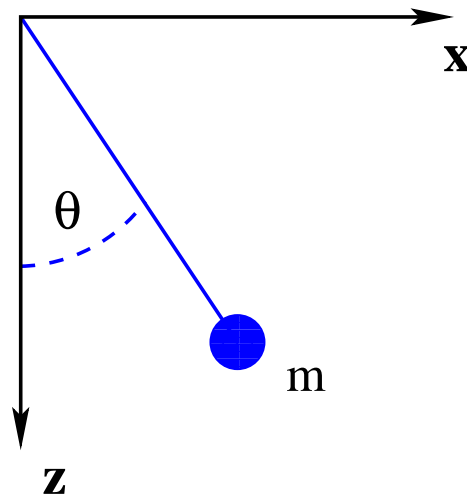


Abbildung 1.2: Sphärisches Pendel

## 1.4 Anwendungsbeispiele zum Lagrange Formalismus

### 1.4.1 Sphärisches Pendel

Als ein erstes Beispiel betrachten wir die Bewegung eines sphärischen Pendels im Schwerfeld der Erde. Dabei handelt es sich um eine Stange der Länge  $l$ , deren Masse wir vernachlässigen können und eine Punktmasse der Masse  $m$ . Das eine Ende der Stange ist im Koordinatenursprung befestigt, am anderen Ende befindet sich die Masse  $m$ . Die Stange sorgt also dafür, dass sich der Massenpunkt auf einer Kugelschale mit dem Radius  $l$  um den Koordinatenursprung bewegen kann (siehe auch Abb 1.2).

Es liegt also nahe, den Ortsvektor des Massenpunktes durch Kugelkoordinaten  $(r, \theta, \varphi)$  zu beschreiben. Die Zwangsbedingung, die durch die Stange realisiert wird, legt dann die Koordinate  $r = l$  fest, während die Winkelkoordinaten als freie generalisierte Koordinaten des Systems betrachtet werden können. Um den Lagrange Formalismus anwenden zu können müssen wir zunächst einmal die kinetische Energie des Systems als Funktion der generalisierten Koordinaten  $(\theta, \varphi)$  und der zugehörigen Geschwindigkeiten  $(\dot{\theta}, \dot{\varphi})$  berechnen. Dazu benötigen wir die Darstellung der Geschwindigkeit in Kugelkoordinaten

$$\begin{aligned}
 \vec{v} &= \frac{d}{dt} \begin{pmatrix} l \sin \theta \cos \varphi \\ l \sin \theta \sin \varphi \\ l \cos \theta \end{pmatrix} \\
 &= l \dot{\theta} \begin{pmatrix} \cos \theta \cos \varphi \\ \cos \theta \sin \varphi \\ -\sin \theta \end{pmatrix} + l \dot{\varphi} \begin{pmatrix} -\sin \theta \sin \varphi \\ \sin \theta \cos \varphi \\ 0 \end{pmatrix} \\
 &= l \dot{\theta} \hat{e}_\theta + l \sin \theta \dot{\varphi} \hat{e}_\varphi.
 \end{aligned} \tag{1.50}$$

Dabei bezeichnen  $\hat{e}_\theta$  und  $\hat{e}_\varphi$  die Einheitsvektoren der Kugelkoordinaten. Da diese Vektoren orthonormal sind, ergibt sich für die kinetische Energie

$$T = \frac{1}{2} m \vec{v}^2 = \frac{1}{2} m l^2 (\dot{\theta}^2 + \dot{\varphi}^2 \sin^2 \theta). \tag{1.51}$$

Wir sehen also, dass die kinetische Energie sowohl von den Geschwindigkeiten  $\dot{\theta}$  und  $\dot{\varphi}$  abhängt aber darüber hinaus auch von der generalisierten Koordinate  $\theta$ .

Auf die Masse des sphärischen Pendels  $m$  soll die Schwerkraft wirken und das Koordinatensystem sei so ausgerichtet, dass diese Schwerkraft parallel zur  $z$ -Achse wirkt. Dies bedeutet, dass wir das zugehörige Potenzial mit der Konstanten  $g$  für die Erdbeschleunigung schreiben können

$$V = -mgz = -mgl \cos \theta.$$

Damit ergibt sich die Lagrangefunktion zu

$$L = T - V = \frac{ml^2}{2} (\dot{\theta}^2 + \dot{\varphi}^2 \sin^2 \theta) + mgl \cos \theta. \quad (1.52)$$

Wir betrachten zunächst die Lagrangesche Bewegungsgleichung für die generalisierte Koordinate  $\varphi$  und berechnen dazu:

$$\begin{aligned} \frac{d}{dt} \frac{\partial L}{\partial \dot{\varphi}} &= \frac{d}{dt} ml^2 \sin^2 \theta \dot{\varphi} \\ &= \frac{\partial L}{\partial \varphi} = 0. \end{aligned} \quad (1.53)$$

Da die Lagrangefunktion des sphärischen Pendels nicht von der generalisierten Koordinate  $\varphi$  abhängt, ist die zeitliche Ableitung von  $\partial L / \partial \dot{\varphi}$  identisch 0, was bedeutet, dass diese Ableitung  $\partial L / \partial \dot{\varphi}$  eine Konstante der Bewegung ist. Es gilt also:

$$ml^2 \sin^2 \theta \dot{\varphi} = C. \quad (1.54)$$

Bei vorgegebenen Anfangsbedingungen, also Werten für die generalisierten Koordinaten  $\theta, \varphi$  und den zugehörigen Geschwindigkeiten  $\dot{\theta}, \dot{\varphi}$  zur Startzeit  $t = 0$  kann man die Konstante  $C$  bestimmen. Die Lagrangesche Bewegungsgleichung für die generalisierte Koordinate  $\theta$  liefert

$$\begin{aligned} \frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}} &= \frac{d}{dt} ml^2 \dot{\theta} \\ &= ml^2 \ddot{\theta} \\ &= \frac{\partial L}{\partial \theta} \\ &= ml^2 \dot{\varphi}^2 \sin \theta \cos \theta - mgl \sin \theta. \end{aligned} \quad (1.55)$$

Ersetzt man in dieser Gleichung  $\dot{\varphi}$  entsprechend (1.54) so ergibt sich daraus die Differentialgleichung für die generalisierte Koordinate

$$l^2 \ddot{\theta} - \frac{C^2 \cos \theta}{m^2 l^4 \sin^3 \theta} + gl \sin \theta = 0.$$

Diese einfache Differenzialgleichung zweiter Ordnung kann nun z.B. numerisch gelöst werden und man erhält für die vorgegebene Anfangsbedingungen das Ergebnis  $\theta(t)$ . Durch Benutzen von (1.54) ergibt sich daraus die Geschwindigkeit  $\dot{\varphi}(t)$ , was uns schliesslich auch die Funktion  $\varphi(t)$  liefert. Die generalisierten Koordinaten sind damit für alle Zeiten  $t$  bestimmt, was natürlich auch bedeutet, dass der Ortsvektor  $\vec{r}(t)$  der Pendelmasse berechnet werden kann.

## 1.4.2 Starrer Körper

Als zweites Anwendungsbeispiel betrachten wir die Bewegungsgleichung eines starren Körpers, der ja schon im ersten Teil, der Vorlesung *Physik I*, behandelt wurde. Es sei daran erinnert, dass ein starrer Körper sechs Freiheitsgrade besitzt. Wir können seine Bewegung durch sechs generalisierte Koordinaten beschreiben. Zur Definition dieser generalisierten Koordinaten legen wir einen Punkt des Körpers,  $P_0$  fest (dies kann z.B. der Schwerpunkt des Körpers sein, muss es aber nicht) und benutzen den Ortsvektor  $\vec{R}_0$ , beziehungsweise die kartesischen Koordinaten dieses Punktes  $(X_0, Y_0, Z_0)$  um die Position des Körpers zu beschreiben. Dadurch sind also bereits 3 generalisierte Koordinaten festgelegt.

Die weiteren 3 Koordinaten werden benötigt um die Orientierung des starren Körpers zu beschreiben. Dazu legen wir ein körperfestes Koordinatensystem fest, das seinen Koordinatenursprung im Punkt  $P_0$  besitzt und für welches die Richtung der kartesischen Achsen durch 3 weitere Punkte des Körpers festgelegt sind. Zum Vergleich ziehen wir ein raumfestes Koordinatensystem heran, dessen Ursprung ebenfalls im Punkt  $P_0$  fixiert ist, dessen Koordinatenachsen aber unabhängig vom Körper an raumfesten Orientierungspunkten festgemacht sind.

Die Orientierung des körperfesten Koordinatensystems relativ zu dem raumfesten System in  $P_0$  ist durch 3 Eulerwinkel eindeutig definiert. Diese 3 Eulerwinkel,  $\phi, \theta$  und  $\psi$ , definieren auch die Orientierung des Körpers im Raum und sollen als weitere generalisierte Koordinaten dienen.

Zur Anwendung des Lagrange Formalismus auf die Bewegung eines starren Körpers müssen wir in jedem Fall die kinetische Energie dieses starren Körpers als Funktion der generalisierten Koordinaten und Geschwindigkeiten berechnen. Dazu beginnen wir mit der Definition der kinetischen Energie als Summe über die kinetischen Energien aller Massenpunkte des Körpers mit Massen  $m_\alpha$  und Ortsvektoren  $\vec{r}_\alpha$

$$\begin{aligned}
 T &= \sum_{\alpha} \frac{1}{2} m_{\alpha} \dot{\vec{r}}_{\alpha}^2 \\
 &= \sum_{\alpha} \frac{1}{2} m_{\alpha} \left( \dot{\vec{R}}_0 + \dot{\vec{\rho}}_{\alpha} \right)^2 \\
 &= \frac{1}{2} \dot{\vec{R}}_0^2 \sum_{\alpha} m_{\alpha} + \dot{\vec{R}}_0 \sum_{\alpha} m_{\alpha} \dot{\vec{\rho}}_{\alpha} + \sum_{\alpha} \frac{1}{2} m_{\alpha} \dot{\vec{\rho}}_{\alpha}^2 \quad (1.56)
 \end{aligned}$$

Bei dem Übergang zur zweiten Zeile dieser Gleichung wurde der Ortsvektor jedes Massenpunktes  $\vec{r}_\alpha$  zerlegt in die Vektorsumme aus dem Ortsvektor des Punktes  $P_0$ ,  $\vec{R}_0$ , und den Relativvektor  $\vec{\rho}_\alpha$  des Massenpunktes  $\alpha$  relativ zum Punkt  $P_0$ . Dadurch erhalten wir in der dritten Zeile die kinetische Energie als Summe von drei Summanden. Der erste Summand entspricht der kinetischen Energie des Referenzpunktes  $P_0$  mit einer Masse  $M = \sum_{\alpha} m_{\alpha}$ , die der Gesamtmasse des Körpers entspricht. Der zweite Term ist ein Mischterm, der sowohl die Geschwindigkeit des Referenzpunktes  $P_0$  enthält als auch die Relativgeschwindigkeiten  $\dot{\vec{\rho}}_{\alpha}$ . Dieser Term verschwindet wenn

- Der Bezugspunkt  $P_0$  gleich dem Schwerpunkt des Körpers ist. Denn der Vektor

$$\sum_{\alpha} m_{\alpha} \vec{\rho}_{\alpha}, \quad (1.57)$$

ist der Vektor, der auf den Schwerpunkt des Systems zeigt multipliziert mit der Masse  $M$  im Koordinatensystem mit dem Ursprung  $P_0$ . Ist aber  $P_0$  selbst der Schwerpunkt, so ist dieser Vektor gleich dem Nullvektor. Dann ist aber die zeitliche Ableitung des Vektors (1.57) identisch Null. Da diese Ableitung von (1.57) in dem Mischterm auftritt, verschwindet dieser.

Der Mischterm ist ausserdem Null

- wenn der Punkt  $P_0$  sich nicht bewegt. In diesem Fall ist ja  $\dot{\vec{R}}_0 = 0$  und es verschwinden sowohl der erste als auch der zweite Summand in der dritten Zeile von (1.56).

Wir nehmen also im Folgenden an, dass der Punkt  $P_0$  festgehalten sei, so dass sich die Bewegung des Körpers auf eine Drehung reduziert, bei der die Drehachse durch den Punkt  $P_0$  weist. Erfolgt diese Drehung mit einer Winkelgeschwindigkeit  $\vec{\omega}$ , so berechnet sich

$$\dot{\vec{\rho}}_\alpha = \omega \times \rho_\alpha,$$

und die kinetische Energie ergibt sich als

$$T = \frac{1}{2} \vec{\omega} \cdot (\mathcal{I} \vec{\omega}). \quad (1.58)$$

Dabei ist  $\mathcal{I}$  der Trägheitstensor des starren Körpers bezogen auf den Referenzpunkt  $P_0$ . Dieser Trägheitstensor hat die besonders einfache Diagonalgestalt

$$\mathcal{I} = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} \quad (1.59)$$

wenn wir ihn in der Basis der Eigenvektoren darstellen. In diesem Fall ist ja

$$T = \frac{1}{2} (I_1 \omega_1^2 + I_2 \omega_2^2 + I_3 \omega_3^2), \quad (1.60)$$

wobei die  $\omega_i$  die Komponenten des Vektors der Winkelgeschwindigkeit  $\vec{\omega}$  bezogen auf die Basis der Eigenvektoren von  $\mathcal{I}$  sind. Diese Eigenvektoren bilden ein kartesisches Koordinatensystem, das fest mit dem Körper verbunden ist (ein körperfestes Koordinatensystem), mit dem Koordinatenursprung  $P_0$ . Es liegt also nahe, dieses körperfeste Koordinatensystem auch als körperfestes Referenzsystem zur Charakterisierung der Orientierung des Körpers heranzuziehen.

Was noch zu tun bleibt, ist die kinetische Energie in der Form von (1.60) umzurechnen auf die oben eingeführten generalisierten Koordinaten,  $\phi, \theta$  und  $\psi$ , und die zugehörigen Geschwindigkeiten. Dazu betrachten wir als erstes, wie sich eine Geschwindigkeit des dritten Eulerwinkels  $\psi$ ,  $\dot{\psi}$  in den Koordinaten des körperfesten Systems,  $\omega_i$  darstellt. Der dritte Eulerwinkel  $\psi$  beschreibt ja gerade eine Drehung um die dritte Achse ( $z$ -Achse) des körperfesten Koordinatensystems. Es gilt also

$$\vec{\omega}(\dot{\psi}) = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \dot{\psi} \end{pmatrix}. \quad (1.61)$$

Wie stellt sich nun eine Bewegung mit der Geschwindigkeit  $\dot{\theta}$  in der körperfesten Basis dar? Eine Bewegung mit der Geschwindigkeit  $\dot{\theta}$  entspricht einer Drehung um die  $x$ -Achse, des Koordinatensystems, das sich nach 2 Schritten der Eulertransformation vom raumfesten in das körperfeste Koordinatensystem ergibt. In diesem System ist der Vektor der Winkelgeschwindigkeit also gegeben durch die Komponenten  $(\dot{\theta}, 0, 0)$ . Um diesen Vektor in der körperfesten Basis der  $\omega_i$  darzustellen müssen wir die dritte Transformation anwenden

$$\begin{aligned}\vec{\omega}(\dot{\theta}) &= \begin{pmatrix} \cos \psi & \sin \psi & 0 \\ -\sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \dot{\theta} \\ 0 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} \dot{\theta} \cos \psi \\ -\dot{\theta} \sin \psi \\ 0 \end{pmatrix}.\end{aligned}\quad (1.62)$$

Weiter benötigen wir noch die Darstellung einer Bewegung mit der Geschwindigkeit  $\dot{\phi}$  in der körperfesten Basis der Eigenvektoren des Trägheitstensors. In der raumfesten Basis oder auch in der Basis nach der ersten Transformation mit dem Euler-Winkel  $\phi$  entspricht eine Drehung um den Euler Winkel  $\phi$ , einer Drehung um die  $z$ -Achse. Die Winkelgeschwindigkeit  $\dot{\phi}$  wird also in diesem Koordinatensystem durch die Komponenten  $(0, 0, \dot{\phi})$  dargestellt. Die Transformation auf das körperfeste Koordinatensystem der Eigenvektoren liefert

$$\begin{aligned}\vec{\omega}(\dot{\phi}) &= \begin{pmatrix} \cos \psi & \sin \psi & 0 \\ -\sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ \dot{\phi} \end{pmatrix} \\ &= \begin{pmatrix} \dot{\phi} \sin \theta \sin \psi \\ \dot{\phi} \sin \theta \cos \psi \\ \dot{\phi} \cos \theta \end{pmatrix}.\end{aligned}\quad (1.63)$$

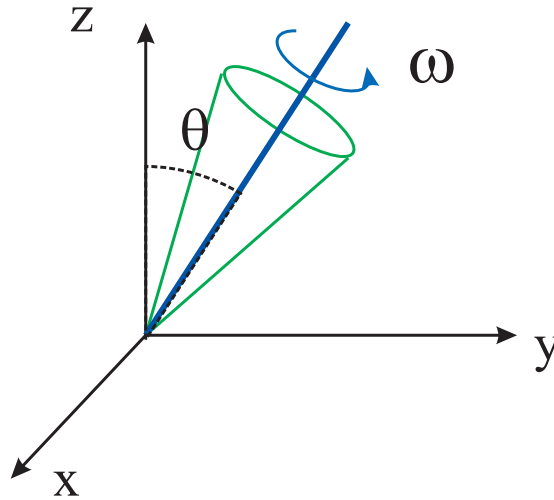
Zusammengenommen ergeben (1.61), (1.62) und (1.63) also eine allgemeine Winkelgeschwindigkeit in den Variablen  $\dot{\psi}$ ,  $\dot{\theta}$  und  $\dot{\phi}$  dargestellt im Koordinatensystem der Eigenvektoren zum Trägheitstensor

$$\vec{\omega} = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} = \begin{pmatrix} \dot{\phi} \sin \theta \sin \psi + \dot{\theta} \cos \psi \\ \dot{\phi} \sin \theta \cos \psi - \dot{\theta} \sin \psi \\ \dot{\phi} \cos \theta + \dot{\psi} \end{pmatrix}.\quad (1.64)$$

Mit (1.60) erhalten wir also für die kinetische Energie der Rotationsbewegung eines starren Körpers

$$T = \frac{I_1}{2} (\dot{\phi} \sin \theta \sin \psi + \dot{\theta} \cos \psi)^2 + \frac{I_2}{2} (\dot{\phi} \sin \theta \cos \psi - \dot{\theta} \sin \psi)^2 + \frac{I_3}{2} (\dot{\phi} \cos \theta + \dot{\psi})^2.\quad (1.65)$$

Wir wollen den Fall eines rotierenden Kreisels betrachten, bei dem der starre Körper aus einem Kegel besteht, der, wie in Abb. 1.3 dargestellt, auf seiner Spitze, die im Koordinatenursprung fixiert sein soll, rotiert. Für den ruhende Referenzpunkt,  $P_0$ , wählen wir natürlich in diesem Fall diese Spitze des Kegels und es gilt deshalb den Trägheitstensor  $\mathcal{I}$  bezogen auf diesen Referenzpunkt zu bestimmen.

Abbildung 1.3: *Symmetrischer Kreisel*

Für einen solch symmetrischen Kreisel, der kreissymmetrisch um die Zentrumsachse ist, gilt, dass diese Symmetrieachse gleichzeitig eine Hauptträgheitsachse ist. Der zugehörige Eigenvektor des Trägheitstensors liegt also parallel zu dieser Hauptträgheitsachse. Wir bezeichnen das entsprechende Hauptträgheitsmoment mit  $I_3$ . Wegen dieser hohen Symmetrie des Kreisels (man spricht auch von einem symmetrischen Kreisel), gilt, dass die beiden anderen Hauptträgheitsmomente identisch sind

$$I_1 = I_2, \quad (1.66)$$

und die zugehörigen Eigenvektoren beliebig senkrecht zur Symmetrieachse gewählt werden können. Wegen (1.66) vereinfacht sich der Ausdruck (1.65) zu

$$T = \frac{I_1}{2} (\dot{\phi}^2 \sin^2 \theta + \dot{\theta}^2) + \frac{I_3}{2} (\dot{\phi}^2 \cos^2 \theta + 2\dot{\phi}\dot{\psi} \cos \theta + \dot{\psi}^2). \quad (1.67)$$

Betrachten wir nun die potenzielle Energie dieses Kreisels, die die Wirkung der Erdanziehung (in Richtung von  $\hat{e}_z$ ) beschreiben soll. Diese Gravitationskraft wirkt als ob die Kraft auf die Gesamtmasse lokalisiert im Schwerpunkt des Kreisels wirken würde. Bezeichnen wir also den Abstand des Schwerpunktes des Kegels von der Spitze mit  $l$  und die Gesamtmasse mit  $M$ , so ist das Gravitationspotenzial gegeben durch

$$V = Mgl \cos \theta. \quad (1.68)$$

Wir sehen also aus den expliziten Ausdrücken für die kinetische und die potenzielle Energie in (1.67) bzw. (1.68), dass die Lagrangefunktion

$$L = T - V = L(\dot{\phi}, \dot{\theta}, \dot{\psi}, \theta)$$

zwar von allen drei generalisierten Geschwindigkeiten abhängt, aber nicht von den generalisierten Koordinaten  $\phi$  und  $\psi$ . Man bezeichnet solche Koordinaten auch als **zyklische Koordinaten** (allgemeine Diskussion erfolgt später). Wegen der Lagrangeschen Bewegungsgleichungen gilt für dies zyklischen Koordinaten

$$\begin{aligned} \frac{\partial L}{\partial \phi} &= 0 = \frac{d}{dt} \frac{\partial L}{\partial \dot{\phi}}, \\ \frac{\partial L}{\partial \psi} &= 0 = \frac{d}{dt} \frac{\partial L}{\partial \dot{\psi}}. \end{aligned} \quad (1.69)$$



Die Ausdrücke

$$\begin{aligned}\frac{\partial L}{\partial \dot{\phi}} &= I_1 \dot{\phi} \sin^2 \theta + I_3 (\dot{\phi} \cos^2 \theta + \dot{\psi} \cos \theta) = c_\phi, \\ \frac{\partial L}{\partial \dot{\psi}} &= I_3 (\dot{\psi} + \dot{\phi} \cos \theta) = I_3 \omega_3 = c_\psi.\end{aligned}\quad (1.70)$$

sind also Konstanten der Bewegung. Aus der zweiten dieser beiden Gleichungen entnehmen wir, dass die Winkelgeschwindigkeit  $\omega_3$  also die Rotationsgeschwindigkeit des Kreisels um die Symmetrieachse konstant bleibt. Ausserdem können wir natürlich diese zweite Gleichung nach  $\dot{\psi}$  auflösen

$$\dot{\psi} = \frac{c_\psi - I_3 \dot{\phi} \cos \theta}{I_3},$$

dieses Ergebnis in die erste der beiden Gleichungen von (1.70) einsetzen und dies auflösen nach

$$\dot{\phi} = \frac{c_\phi - c_\psi \cos \theta}{I_1 \sin^2 \theta}.\quad (1.71)$$

Die Winkelgeschwindigkeit  $\dot{\phi}$  ist die Geschwindigkeit mit der die Figurenachse um die raumfeste  $z$ -Achse präzediert. Auch diese Geschwindigkeit der Präzessionsbewegung bleibt also konstant, wenn der Winkel zwischen Symmetrieachse des Kreisels und der laborfesten  $z$ -Achse,  $\theta$  konstant bleibt. Ändert sich allerdings  $\theta$  so muss sich nach (1.71) auch  $\dot{\phi}$  ändern.

Die Bewegungsgleichung für die Winkelkoordinate  $\theta$  ist auch leicht zu aufzustellen, die Lösung dieser Gleichung ist jedoch nicht sehr transparent und deshalb soll hier darauf verzichtet werden.

### 1.4.3 Geschwindigkeitsabhängige Kraft

Als drittes Beispiel wollen wir zeigen, dass die Definition einer geschwindigkeitsabhängigen Kraft nach (1.44) insbesondere die Wirkung eines elektromagnetischen Feldes auf ein geladenes Teilchen erfasst. Wir betrachten dazu die Bewegung eines Teilchens mit der Ladung  $e$  in einem elektromagnetischen Feld, das durch die Potentiale  $\Phi(\vec{r})$  und  $\vec{A}(\vec{r})$  definiert ist. Aus diesen Potentialen berechnen sich die elektrischen und magnetischen Felder durch die Beziehungen

$$\vec{E} = -\vec{\nabla}\Phi \quad \text{und} \quad \vec{B} = \vec{\nabla} \times \vec{A}.\quad (1.72)$$

Das geladene Teilchen bewege sich ohne weitere Zwangsbedingungen, so dass wir wieder die kartesischen Koordinaten als generalisierte Koordinaten des Systems heranziehen können. Wir betrachten die Potenzialfunktion

$$V(x, y, z, \dot{x}, \dot{y}, \dot{z}) = e\Phi(\vec{r}) - e\vec{A}(\vec{r}) \cdot \dot{\vec{r}},\quad (1.73)$$

die offensichtlich von den generalisierten Koordinaten  $(x, y, z)$  und den zugehörigen Geschwindigkeiten abhängt. Mit der allgemeinen Beziehung (1.44) können wir nun die generalisierte Kraft  $Q_i$  in Richtung der Koordinate  $q_i = x$  berechnen zu

$$\begin{aligned}F_x &= -\frac{\partial V}{\partial x} + \frac{d}{dt} \frac{\partial V}{\partial \dot{x}} \\ &= -e \frac{\partial \Phi}{\partial x} + e \dot{\vec{r}} \cdot \frac{\partial \vec{A}}{\partial x} - e \frac{d}{dt} A_x\end{aligned}\quad (1.74)$$

Es soll nun gezeigt werden, dass die beiden letzten Terme in dieser Gleichung gilt

$$e\dot{\vec{r}} \cdot \frac{\partial \vec{A}}{\partial x} - e \frac{d}{dt} A_x,$$

mit der  $x$ -Komponente von

$$e \left( \dot{\vec{r}} \times (\vec{\nabla} \times \vec{A}) \right)_x = e \left( \dot{\vec{r}} \times \vec{B} \right)_x,$$

identifiziert werden kann. Dazu schreiben wir diese  $x$ -Komponente explizit aus

$$\begin{aligned} \left( \dot{\vec{r}} \times \vec{B} \right)_x &= \dot{y} (\vec{\nabla} \times \vec{A})_z - \dot{z} (\vec{\nabla} \times \vec{A})_y \\ &= \dot{y} \left( \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) - \dot{z} \left( \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) \\ &= \dot{y} \frac{\partial A_y}{\partial x} + \dot{z} \frac{\partial A_z}{\partial x} + \dot{x} \frac{\partial A_x}{\partial x} - \dot{x} \frac{\partial A_x}{\partial x} - \dot{y} \frac{\partial A_x}{\partial y} - \dot{z} \frac{\partial A_x}{\partial z} \\ &= \dot{\vec{r}} \cdot \frac{\partial \vec{A}}{\partial x} - \frac{d}{dt} A_x \end{aligned}$$

womit die Behauptung bewiesen wäre. Damit können wir also (1.74) umschreiben in

$$\begin{aligned} F_x &= -e \left( \vec{\nabla} \Phi \right)_x + e \left( \dot{\vec{r}} \times \vec{B} \right)_x \\ &= e E_x + e \left( \dot{\vec{r}} \times \vec{B} \right)_x. \end{aligned} \tag{1.75}$$

Die generalisierte Kraft  $F_x$  für das geschwindigkeitsabhängige Potenzial aus (1.73) ist also gerade die  $x$ -Komponente der Coulombkraft plus der Lorentzkraft auf die bewegte Ladung  $e$ . Natürlich verläuft der Nachweis für die anderen kartesischen Komponenten  $F_y$  und  $F_z$  entsprechend.

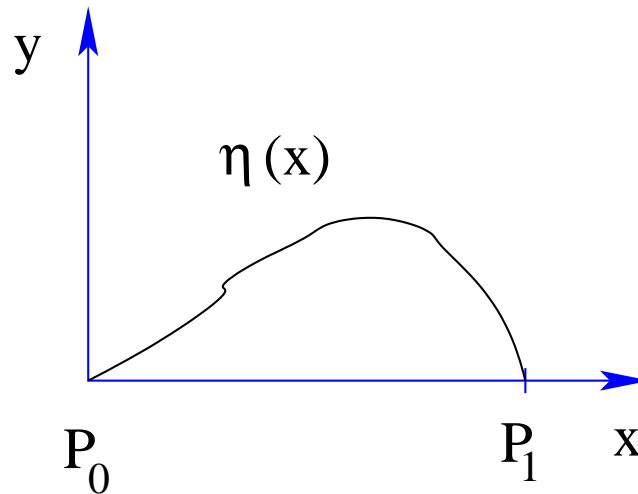


Abbildung 1.4: Beispiel für einen Weg im Variationsverfahren

## 1.5 Euler-Lagrangesche Variationsgleichungen

In diesem Abschnitt wollen wir uns aus einem Grund, der erst etwas später deutlich werden wird, mit Variationsproblemen befassen. Wir werden dazu zunächst ein einfaches Beispiel betrachten, an dem wir die Problemstellung verdeutlichen wollen und das prinzipielle Vorgehen. In diesem Beispiel wollen wir verifizieren, dass die kürzeste Verbindungslinie zwischen zwei Punkten durch die Verbindungsgerade zwischen diesen Punkten gegeben ist. Ohne Beschränkung der Allgemeinheit können wir annehmen, dass diese beiden Punkte in der  $x-y$  Ebene liegen und zwar, wie in der Abb. 1.4 skizziert  $P_1$  am Koordinatenursprung, also  $x_1 = y_1 = 0$ , und  $P_2$  mit den Koordinaten  $x_2 = 1$ ,  $y_2 = 0$ . Wir wollen nun zeigen, dass der kürzeste Verbindungsweg von  $P_1$  nach  $P_2$  durch die Verbindungsgerade gegeben ist, die ja durch die Funktion

$$y_0(x) = 0, \quad (1.76)$$

beschrieben wird. Dazu betrachten wir die Konkurrenz aller möglichen Funktionen, die durch die Punkte  $P_1$  und  $P_2$  gehen (siehe auch Abb. 1.4)

$$y_\alpha(x) = y_0(x) + \alpha\eta(x). \quad (1.77)$$

Dabei ist  $\alpha$  eine reelle Konstante und  $\eta(x)$  eine beliebige Funktion mit

$$\eta(0) = \eta(1) = 0. \quad (1.78)$$

Wir werden dann zu zeigen haben, dass für jede dieser Funktionen  $\eta(x)$  der Ansatz (1.77) den kürzesten Weg liefert, wenn  $\alpha = 0$  ist. Nehmen wir also mal als Beispiel die Funktion

$$\eta(x) = x^2 - x, \quad (1.79)$$

die natürlich die Bedingung (1.78) erfüllt. Die Länge des Weges, der durch eine Funktion  $y(x)$  definiert ist kann man berechnen durch das Integral

$$S = \int_{P_1}^{P_2} ds = \int_{P_1}^{P_2} \sqrt{dx^2 + dy^2} = \int_0^1 \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx. \quad (1.80)$$

Setzen wir nun in die Definition von  $y_\alpha(x)$  in (1.77) die Funktion  $\eta(x)$  aus (1.79) ein, so ergibt sich mit

$$\frac{dy}{dx} = 2\alpha x - \alpha$$

für die Länge des Weges nach (1.80)

$$\begin{aligned} S(\alpha) &= \int_0^1 dx \sqrt{1 + (2\alpha x - \alpha)^2} \\ &= \frac{\sqrt{1 + \alpha^2}}{2} + \frac{1}{4\alpha} \ln \left( \frac{\sqrt{1 + \alpha^2} + \alpha}{\sqrt{1 + \alpha^2} - \alpha} \right). \end{aligned} \quad (1.81)$$

Die Berechnung dieses Integrals ist nicht trivial. Man kann sich dann davon überzeugen, dass diese Länge des Weges in der Tat minimal wird, wenn der Parameter  $\alpha$  den Wert 0 annimmt. Damit haben wir also auf doch recht mühsame Weise gezeigt, dass alle Modifikationen des geraden Verbindungsweges für die spezielle Form  $\eta(x)$  aus (1.79) nur zu einer Verlängerung des Weges führen. Wir können uns also jetzt eine andere Funktion  $\eta(x)$ , die (1.78) erfüllt ausdenken und die gleiche Prozedur wiederholen.

Man sieht natürlich schnell ein, dass dieses Verfahren, den kürzesten Verbindungsweg zu finden nicht sehr geeignet ist: man muss ja schliesslich alle möglichen Funktionen  $\eta(x)$  untersuchen. Die Absicht dieses Beispiels war es auch nicht einen effizienten Algorithmus für ein solches Optimierungsverfahren vorzuführen, es sollte vielmehr dazu dienen, die folgenden Überlegungen durch dieses konkrete Beispiel zu veranschaulichen.

Wir werden deshalb eine andere Methode entwickeln um Variationsprobleme zu lösen vom Typ, dass die Funktion  $y_\alpha(x)$  zu finden ist für die das Integral

$$J(y_\alpha) = \int_{x_0}^{x_1} f(y_\alpha, y'_\alpha, x) dx, \quad (1.82)$$

ein Extremum ist. Wir sehen, dass das bereit diskutierte Beispiel von (1.80) einen Spezialfall darstellt, bei dem die Funktion unter dem Integral

$$f(y_\alpha, y'_\alpha, x) = \sqrt{1 + \left( \frac{dy_\alpha}{dx} \right)^2}$$

nur von der Ableitung

$$y'_\alpha = \frac{dy_\alpha}{dx}$$

nicht aber von der Funktion selbst oder der Integrationsvariablen abhängt. Bei diesem Variationsverfahren sollen alle Funktionen  $y_\alpha(x)$  berücksichtigt werden, die von der Form (1.77) sind mit

$$\eta(x_0) = \eta(x_1) = 0, \quad (1.83)$$

die Werte der Funktion am Anfangs- und Endpunkt der Integration sollen also fest vorgegeben sein. Eine notwendige Bedingung dafür, dass  $J(y_\alpha)$  ein Extremum ist, ist gegeben durch

$$\begin{aligned} \frac{dJ}{d\alpha} = 0 &= \int_{x_0}^{x_1} \frac{d}{d\alpha} f(y_\alpha, y'_\alpha, x) dx \\ &= \int_{x_0}^{x_1} \left( \frac{\partial f}{\partial y_\alpha} \frac{\partial y_\alpha}{\partial \alpha} + \frac{\partial f}{\partial y'_\alpha} \frac{\partial y'_\alpha}{\partial \alpha} \right) dx. \end{aligned} \quad (1.84)$$

Aus dem Ansatz (1.77) für die Funktionen  $y_\alpha(x)$  ergibt sich

$$\begin{aligned}\frac{\partial y_\alpha}{\partial \alpha} &= \eta(x) \\ y'_\alpha(x) &= \frac{\partial y_0}{\partial x} + \alpha \frac{\partial \eta}{\partial x}, \\ \frac{\partial y'_\alpha}{\partial \alpha} &= \frac{\partial \eta}{\partial x} = \eta'(x).\end{aligned}\tag{1.85}$$

Setzt man diese Ergebnisse in (1.84) ein, ergibt sich

$$0 = \int_{x_0}^{x_1} \left( \frac{\partial f}{\partial y_\alpha} \eta + \frac{\partial f}{\partial y'_\alpha} \eta' \right) dx.\tag{1.86}$$

Den zweiten Summanden unter diesem Integral bearbeiten wir nun mit Hilfe der partiellen Integration

$$\int_{x_0}^{x_1} \frac{\partial f}{\partial y'_\alpha} \frac{d\eta}{dx} dx = \underbrace{\frac{\partial f}{\partial y'_\alpha} \eta}_{=0} \Big|_{x_0}^{x_1} - \int_{x_0}^{x_1} \frac{d}{dx} \frac{\partial f}{\partial y'_\alpha} \eta dx.$$

Dabei ist der erste Term auf der rechten Seite dieser Gleichung identisch null, da ja wegen (1.83) die Funktion  $\eta$  an den Grenzpunkten des Integrals verschwindet. Mit diesem Ergebnis ergibt sich aus (1.86)

$$0 = \int_{x_0}^{x_1} \left( \frac{\partial f}{\partial y_\alpha} - \frac{d}{dx} \frac{\partial f}{\partial y'_\alpha} \right) \eta(x) dx.\tag{1.87}$$

Dies muss für alle Funktionen  $\eta(x)$  gelten, also insbesondere auch für Funktionen, die überall identisch null sind bis auf einen beliebigen Wert  $\tilde{x}$  aus dem Integrationsintervall  $[x_0, x_1]$ . Das ist aber nur dann möglich, wenn für die gesuchte Funktion  $y_\alpha = y$  gilt

die Euler-Lagrangesche Variationsgleichung

$$\frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial y'} = 0 \quad \text{für alle } x \in [x_0, x_1].\tag{1.88}$$

Als erste Anwendung betrachten wir den bereits oben diskutierten Fall mit

$$f(y, y', x) = \sqrt{1 + (y')^2}$$

Die Anwendung der Eulerschen Variationsgleichung liefert

$$\frac{\partial f}{\partial y} = 0 = \frac{d}{dx} \frac{\partial f}{\partial y'} = \frac{d}{dx} \frac{y'}{\sqrt{1 + y'^2}}.$$

Dies bedeutet, dass  $y'$  als Funktion von  $x$  eine Konstante ist, die wir mit  $a$  bezeichnen. Daraus ergibt sich

$$y' = a \quad \Rightarrow \quad y(x) = ax + b.$$

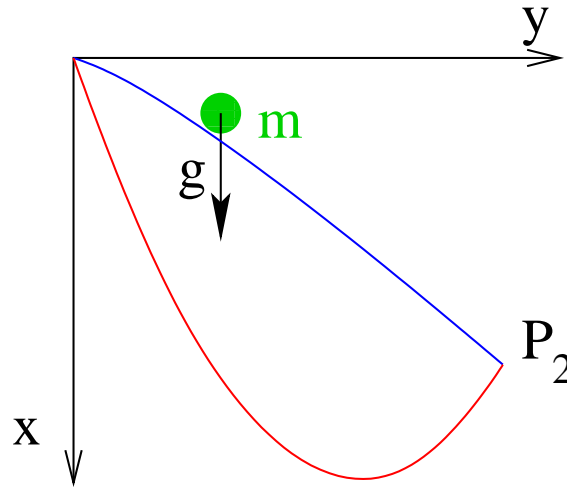


Abbildung 1.5: Skizze zum Brachistochronen Problem

Die Konstanten  $a$  und  $b$  müssen so gewählt werden, dass die Randbedingungen  $y(0) = y(1) = 0$  erfüllt werden, was natürlich zu der erwarteten Lösung (1.76)  $y = 0$  führt.

Man kann sich leicht davon überzeugen, dass entsprechende Gleichungen auch gelten, wenn wir ein Variationsverfahren betrachten, in dem mehrere Funktionen  $y_i(x)$  für  $i = 1 \dots n$  optimiert werden sollen. In diesem Fall haben wir also das Variationsproblem, dass wir die Funktionen  $y_i(x)$  bestimmen wollen, die an den Grenzwerten fest vorgeben sind

$$y_i(x_0) = y_{i0} \quad \text{und} \quad y_i(x_1) = y_{i1} \quad \text{für } i = 1 \dots n, \quad (1.89)$$

und für die gilt, dass

$$J = \int_{x_0}^{x_1} f(y_1 \dots y_n, y'_1 \dots y'_n, x) dx, \quad (1.90)$$

ein Extremum ist. Dies führt zu  $n$  Euler-Lagrange Gleichungen vom Typ (1.88) der Form

$$\frac{\partial f}{\partial y_i} - \frac{d}{dx} \frac{\partial f}{\partial y'_i} = 0. \quad (1.91)$$

Der Beweis erfolgt analog zu dem eindimensionalen Fall, wobei man unabhängige Variationen  $\eta_i(x)$  für die einzelnen Funktionen betrachtet.

### 1.5.1 Brachistochronen Problem

Als weiteres Anwendungsbeispiel betrachten wir das Problem der Brachistochrone.<sup>1</sup> Dabei stellt man die Frage: Auf welcher Bahn  $y(x)$  gelangt ein reibungslos gleitender Massenpunkt unter dem Einfluss der Schwerkraft am schnellsten von Punkt  $P_1$  mit  $(x_1, y_1) = (0, 0)$  nach  $P_2$  mit den Koordinaten  $(x_2, y_2)$ ? (siehe auch Abb. 1.5). Dabei soll die Masse  $m$  ihren Weg mit der Startgeschwindigkeit  $v = 0$  beginnen. Der obere der beiden in Abb. 1.5 skizzierten Wege ist offensichtlich kürzer als der untere. Dafür wird die Masse aber auf

<sup>1</sup>Der Name Brachistochrone hat seinen Ursprung in der griechischen Sprache und bedeutet so viel wie in kürzester Zeit

dem unteren Weg durch die Beschleunigung  $g$  eine größere Geschwindigkeit erreichen. Die Frage ist also auf welchem Weg sind Weglänge und Geschwindigkeit so optimiert, dass die Zeit minimal wird.

Zur Beantwortung dieser Fragen berechnen wir die Zeit, die Masse  $m$  auf einem bestimmten Weg benötigt zu

$$J = \int_{P_1}^{P_2} dt = \int_{P_1}^{P_2} \frac{ds}{v} = \int_{x_1}^{x_2} \frac{\sqrt{1+y'^2}}{v(x)} dx. \quad (1.92)$$

Dabei haben wir die Differenzialform für die Zeit,  $dt$ , ersetzt durch den Quotient aus dem Wegelement  $ds$  und der aktuellen Geschwindigkeit  $v$  der Masse  $m$ . Bei der letzten Gleichung wurde  $ds$  entsprechend (1.80) substituiert.

Zur Bestimmung der jeweiligen Geschwindigkeit  $v$  betrachten wir die Energiebilanz der Masse  $m$  als Summe der jeweiligen kinetischen Energie und der potenziellen Energie im Schwerfeld der Erde

$$E = \frac{1}{2}mv^2 - mgx = 0.$$

Wegen der Anfangsbedingung  $v = 0$  und  $x = 0$  ist diese Energie auf  $E = 0$  normiert. Diese Gleichung kann nun aufgelöst werden nach

$$v(x) = \sqrt{2gx}.$$

Setzt man das Ergebnis in (1.92) ein so ergibt sich das zu minimierende Integral in der Form

$$J = \int_0^{x_2} \frac{\sqrt{1+y'^2}}{\sqrt{2gx}} dx,$$

Der Integrand hängt also in diesem Fall von der Ableitung  $y'$  und der Variablen  $x$  ab, nicht aber von  $y$

$$f(y', x) = \frac{\sqrt{1+y'^2}}{\sqrt{2gx}}. \quad (1.93)$$

Die Euler-Lagrange Gleichung (1.88) liefert also

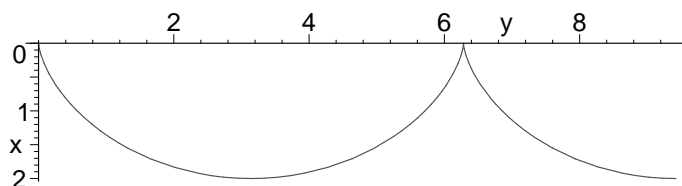
$$\frac{\partial f}{\partial y} = 0 = \frac{d}{dx} \frac{\partial f}{\partial y'},$$

was bedeutet, dass die Ableitung  $\partial f / \partial y'$  eine Konstante (unabhängig von  $x$ ) ergibt

$$\frac{\partial f}{\partial y'} = \frac{1}{\sqrt{2gx}} \frac{y'}{\sqrt{1+y'^2}} = \frac{1}{\sqrt{2g}\sqrt{2a}}, \quad (1.94)$$

wobei der letzte Ausdruck die Konstante ist, die wir hier etwas kompliziert definiert haben. Diese Gleichung können wir umformen zu

$$\begin{aligned} \frac{y'^2}{(1+y'^2)x} &= \frac{1}{2a} \\ y'^2 &= \frac{x}{2a\left(1-\frac{x}{2a}\right)} = \frac{x^2}{2ax-x^2} \\ y' &= \frac{dy}{dx} = \frac{x}{\sqrt{2ax-x^2}} \end{aligned} \quad (1.95)$$

Abbildung 1.6: *Zykloide als Lösung des Brachistochronen Problems*

Diese letzte Gleichung können wir integrieren

$$\int_{P_1}^{P_2} dy = y(x_2) - y(0) = \int_0^{x_2} \frac{x}{\sqrt{2ax - x^2}} dx .$$

Zur Berechnung des Integrals benutzen wir die folgende Substitution

$$\begin{aligned} x &= a(1 - \cos \theta) \\ dx &= \frac{dx}{d\theta} d\theta = a \sin \theta d\theta , \end{aligned} \tag{1.96}$$

was uns führt auf

$$\begin{aligned} y &= \int_0^{\theta_2} \frac{a^2(1 - \cos \theta) \sin \theta}{\sqrt{(2a^2 - 2a^2 \cos \theta - a^2(1 - \cos \theta)^2)}} d\theta \\ &= \int_0^{\theta_2} a(1 - \cos \theta) d\theta \\ &= a(\theta - \sin \theta) . \end{aligned} \tag{1.97}$$

Die mit (1.96) und (1.97) parametrisierte Bahnkurve, die sogenannte Zyklode, ist in Abb. 1.6 für den Fall  $a = 1$  dargestellt. Die Konstante  $a$  muss nun so gewählt werden, dass diese Bahnkurve den Endpunkt  $P_2$  erreicht.



## 1.6 Hamiltonsches Prinzip

Wir kehren wieder zurück zu der Beschreibung von mechanischen Systemen, die, unter Berücksichtigung von möglichen Zwangsbedingungen, durch  $n = 3N - k$  generalisierte Koordinaten  $q_i$  beschrieben werden. Dabei bezeichnet  $n$  die Zahl der Freiheitsgrade. Für den Fall  $n = 2$  wird die zeitliche Entwicklung des Systems durch die Werte der Funktionen  $q_1(t)$  und  $q_2(t)$  als Funktion der Zeit beschrieben. Diese zeitliche Entwicklung wird durch eine Bahnkurve im zweidimensionalen Raum  $R^2$ , also einer Trajektorie in der Ebene, die durch die Koordinaten  $q_1$  und  $q_2$  aufgespannt ist, beschrieben.

Das Hamiltonsche Prinzip besagt: Das System bewegt sich von einem Startpunkt zum Zeitpunkt  $t_1$  zum Endpunkt zur Zeit  $t_2$  auf der Trajektorie für die die Wirkung  $S$ , definiert durch

$$S := \int_{t_1}^{t_2} dt L(q_i(t), \dot{q}_i(t), t), \quad (1.98)$$

ein Extremum ist. Dabei sind die Start- und Endpunkte  $q_i(t_1), q_i(t_2)$  vorgegeben.

Anders ausgedrückt gilt nach dem Hamiltonschen Prinzip für die realisierte Trajektorie das Variationsprinzip

$$\delta \left( \int_{t_1}^{t_2} dt L \right) = 0, \quad (1.99)$$

für Variationen des Weges  $\delta q_i(t)$ , bei denen der Start- und der Zielpunkt festgehalten sind. Da die Variationen des Weges durch die Variationen der generalisierten Koordinaten beschrieben werden, erfüllen die zur Variation zugelassenen Wege natürlich auch die Zwangsbedingungen des Systems.

Wir wollen in diesem Abschnitt zeigen, dass das Hamiltonsche Prinzip für eine konservative Kraft oder aber auch für eine geschwindigkeitsabhängig Kraft direkt aus den Newtonschen Bewegungsgleichungen plus dem D'Alembertschen Prinzip folgt (siehe (a) in 1.7). Mit dem im vorhergehenden Abschnitt geleisteten Vorarbeiten ist es dann sehr einfach zu sehen, dass das Hamiltonsche Prinzip auf die Bewegungsgleichungen II. Art des Lagrange Formalismus führt ((b) in 1.7). Darüber hinaus haben wir natürlich bereits im Abschnitt 1.3 gezeigt, dass die Newtonschen Bewegungsgleichungen mit dem D'Alembertschen Prinzip auch direkt auf die Lagrangeschen Bewegungsgleichungen führen. Andererseits haben wir in dem Abschnitt natürlich auch gezeigt, dass aus den Lagrangeschen Bewegungsgleichungen für ein freies Teilchen im konservativen Kraftfeld die Newtonsche Bewegungsgleichung hergeleitet werden kann.

Bevor uns dem Beweis zuwenden, dass aus dem Newtonschen Bewegungsgleichungen das Hamiltonsche Prinzip folgt, sei auf wichtige Unterschiede in diesen Formulierungen hingewiesen:

- Die Newtonschen Bewegungsgleichungen sind Differenzialgleichungen zweiter Ordnung in der Zeit. Dies bedeutet, dass sie eine Aussage enthalten über die Bewegung zu jeweils einem bestimmten Zeitpunkt. Relevant in diesen Gleichungen sind nur die generalisierten Koordinaten des Systems und ihre Ableitungen nach der Zeit zu jeweils einem Zeitpunkt. Das Hamiltonsche Prinzip hingegen ist eine Aussage über

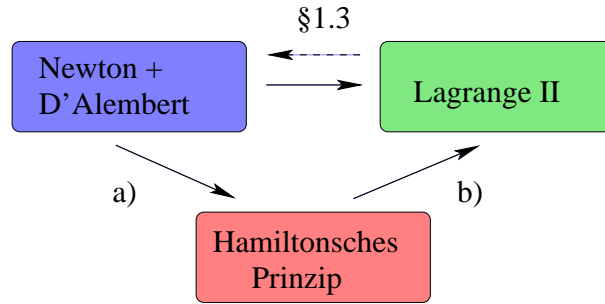


Abbildung 1.7: Bewiesene Zusammenhänge zwischen verschiedenen Formulierungen der Mechanik

das Wirkungsintegral, also über den gesamten Weg, den das System zurücklegt von einer Startzeit  $t_1$  bis zu Endzeit  $t_2$ .

- Zur Bestimmung einer eindeutigen Lösung werden bei den Newtonschen Bewegungsgleichungen Randbedingungen in der Form angegeben, dass man für jeden Freiheitsgrad, d.h. für jede generalisierte Koordinate, den Wert dieser Koordinate und ihre Ableitung nach der Zeit zur Startzeit  $t_1$  angibt, also 2 Randbedingungen pro Freiheitsgrad. Auch beim Hamiltonschen Prinzip werden pro Freiheitsgrad zwei Randbedingungen vorgegeben. In diesem Fall sind das aber jeweils der Wert der generalisierten Koordinate zur Startzeit und zur Zielzeit.

Multipliziert man die Newtonschen Bewegungsgleichungen für das Teilchen  $i$ , auf das die reale Kraft  $\vec{F}_i$  und die Kraft zur Einhaltung der Zwangsbedingung,  $\vec{F}_i^{\text{Zwang}}$ , mit dem entsprechenden Vektor einer virtuellen Verrückung  $\delta\vec{r}_i$ , und summiert über alle Teilchen, so ergibt sich

$$0 = \sum_i \left( \vec{F}_i + \vec{F}_i^{\text{Zwang}} - m_i \ddot{\vec{r}}_i \right) \delta\vec{r}_i. \quad (1.100)$$

Wegen des D'Alembertschen Prinzips verschwindet der zweite Term in dieser Gleichung, der Beitrag der Zwangskräfte. Die resultierende Gleichung gilt für alle Zeiten der Bewegung und wir können sie deshalb integrieren zu

$$\begin{aligned} 0 &= \int_{t_1}^{t_2} dt \sum_i \left( \vec{F}_i - m_i \ddot{\vec{r}}_i \right) \delta\vec{r}_i \\ &= - \int_{t_1}^{t_2} dt \sum_i \left( \frac{\partial V}{\partial \vec{r}_i} \right) \delta\vec{r}_i + \int_{t_1}^{t_2} dt \sum_i \left( \frac{d}{dt} \frac{\partial V}{\partial \dot{\vec{r}}_i} \right) \delta\vec{r}_i - \int_{t_1}^{t_2} dt \sum_i m_i \ddot{\vec{r}}_i \delta\vec{r}_i. \end{aligned} \quad (1.101)$$

Bei dem Übergang zur zweiten Gleichung haben wir die Kraft auf das Teilche  $i$ ,  $\vec{F}_i$  ersetzt durch die Definition einer konservativen oder geschwindigkeitsabhängigen Kraft aus einem Potenzial  $V$ . Den zweiten Term auf der rechten Seite dieser Gleichung können wir nun partiell integrieren

$$\int_{t_1}^{t_2} dt \sum_i \frac{d}{dt} \frac{\partial V}{\partial \dot{\vec{r}}_i} \delta\vec{r}_i = \underbrace{\sum_i \frac{\partial V}{\partial \dot{\vec{r}}_i} \delta\vec{r}_i \Big|_{t_1}^{t_2}}_{=0} - \int_{t_1}^{t_2} dt \sum_i \frac{\partial V}{\partial \vec{r}_i} \delta\vec{r}_i. \quad (1.102)$$

Dabei verschwindet der erste Term auf der linken Seite, da wir uns auf virtuelle Verrückungen beschränken wollen, die am Start- und Zielpunkt verschwinden, also vorgegebene Randbedingungen für die Koordinaten zur Zeit  $t_1$  und  $t_2$  respektieren. Damit können wir die ersten beiden Terme auf der rechten Seite der zweiten Zeile von (1.101) zusammenfassen zu

$$-\int_{t_1}^{t_2} dt \sum_i \left( \frac{\partial V}{\partial \vec{r}_i} \right) \delta \vec{r}_i - \int_{t_1}^{t_2} dt \sum_i \frac{\partial V}{\partial \dot{\vec{r}}_i} \delta \dot{\vec{r}}_i = -\delta \int_{t_1}^{t_2} dt V. \quad (1.103)$$

Dies bezeichnet die Variation des Potentials  $V$  integriert über einen Weg, der wie die virtuellen Verrückungen sowohl die Zwangsbedingungen als auch die Randbedingungen zu den Zeiten  $t_1$  und  $t_2$  respektiert.

Wir betrachten nun den dritten Term auf der rechten Seite der zweiten Zeile von Gleichung (1.101) und formen diesen ebenfalls um mit Hilfe der partiellen Integration

$$\begin{aligned} \int_{t_1}^{t_2} dt \sum_i m_i \ddot{\vec{r}}_i \delta \vec{r}_i &= \underbrace{\sum_i m_i \dot{\vec{r}}_i \delta \vec{r}_i}_{=0} \Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} dt \sum_i m_i \dot{\vec{r}}_i \delta \dot{\vec{r}}_i \\ &= - \int_{t_1}^{t_2} dt \sum_i \frac{\partial}{\partial \dot{\vec{r}}_i} \left( \frac{m_i}{2} \dot{\vec{r}}_i^2 \right) \delta \dot{\vec{r}}_i \\ &= -\delta \int_{t_1}^{t_2} dt T, \end{aligned} \quad (1.104)$$

wobei der Term in der letzten Zeile die Variation des Integrals der kinetischen Energie  $T$  bezeichnet, integriert über einen Weg mit den gleichen Bedingungen wie in (1.103). Die Ergebnisse von (1.103) und (1.104) eingesetzt in (1.102) liefern das Hamiltonsche Prinzip

$$0 = \delta \int_{t_1}^{t_2} dt (T - V) = \delta \int_{t_1}^{t_2} dt L$$

was ja zu beweisen war.

Der Nachweis, dass aus dem Hamiltonschen Prinzip die Lagrangeschen Bewegungsgleichung der zweiten Art folgen (Schritt b in Abb. 1.7) ist nach den geleisteten Vorarbeiten sehr einfach. Wir müssen dazu nur verifizieren, dass das Hamiltonsche Prinzip (1.99) den Forderungen des Euler-Lagrangeschen Variationsprinzips von (1.90) entspricht wenn wir identifizieren

$$\begin{aligned} x &\rightarrow t \\ y_i(x) &\rightarrow q_i(t) \\ y'_i &\rightarrow \dot{q}_i \\ f(y_i, y'_i, t) &\rightarrow L(q_i, \dot{q}_i, t). \end{aligned}$$

Die entsprechende Umschreibung der Euler-Lagrangeschen Variationsgleichung (1.91) liefert dann die Lagrangeschen Bewegungsgleichungen

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0.$$

## 1.7 Erhaltungssätze und Symmetrien I

Die Ziffer I in der Überschrift zu diesem Abschnitt weist darauf hin, dass wir den Zusammenhang zwischen Symmetrien des Systems und Erhaltungsgrößen der Bewegung nicht nur in diesem Abschnitt behandeln wollen, sondern auf dieses wichtige Thema auch im weiteren Verlauf der Vorlesung noch zurückkommen werden.

Für eine erste Beschreibung definieren wir einen **generalisierten Impuls** eines Systems als die partielle Ableitung der Lagrange Funktion  $L$  nach einer generalisierten Geschwindigkeit  $\dot{q}_j$ :

$$p_j := \frac{\partial L}{\partial \dot{q}_j}. \quad (1.105)$$

Diese Größe  $p_j$  bezeichnet man auch häufig als **kanonischen Impuls** oder als zur generalisierten Koordinate  $q_j$  **konjugierten Impuls**. Zur Charakterisierung dieses Begriffes *generalisierter Impuls* zunächst einige Anmerkungen:

- Im Fall eines freien Teilchens der Masse  $m$  in einem konservativen Kraftfeld beschrieben durch ein Potenzial  $V$  können wir die kartesischen Koordinaten des Teilchens als generalisierte Koordinaten annehmen und die Lagrange Funktion in der Form schreiben

$$L = \frac{m}{2} (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - V(x, y, z).$$

Damit ergibt sich für den generalisierten Impuls, den wir auch als zur generalisierten Koordinate  $x$  konjugierten Impuls bezeichnen können

$$p_x = \frac{\partial L}{\partial \dot{x}} = m\dot{x}.$$

Dieser Impuls  $p_x$  ergibt sich also wie die gewöhnliche Impulskomponente  $p_x$  des Newtonschen Impulses als das Produkt der Masse mit der Geschwindigkeit des Teilchens.

- Die generalisierten Impulse entsprechen aber natürlich nicht immer der Definition des konventionellen Newtonschen Impulses, sie haben häufig sogar eine andere Dimension. Als Beispiel betrachten wir hier die Bewegung eines Massenpunktes in einem konservativen Kraftfeld, dessen Bewegungsfreiheit durch Zwangsbedingungen auf den Mantel eines Zylinders mit dem Radius  $R$  und der zentralen Achse entlang der  $z$ -Achse eingeschränkt ist. In diesem Fall bieten sich natürlich die Zylinderkoordinaten  $\varphi$  und  $z$  als generalisierte Koordinaten an. Die Lagrangefunktion ergibt sich dann als

$$L = \frac{m}{2} (R^2 \dot{\varphi}^2 + \dot{z}^2) - V(\varphi, z). \quad (1.106)$$

Der zum Winkel  $\varphi$  zugehörige kanonische Impuls ist dann

$$p_\varphi = \frac{\partial L}{\partial \dot{\varphi}} = mR^2 \dot{\varphi}. \quad (1.107)$$

Dies ist gerade die  $z$ -Komponente des Drehimpulses des Teilchens bezogen auf den Koordinatenursprung.

- Ganz andere Definitionen für den generalisierten Impuls ergeben sich, wenn man geschwindigkeitsabhängige Kräfte und entsprechende Potenziale betrachtet. Als Beispiel sei die Lagrange Funktion eines geladenen Teilchens in einem elektromagnetischen Feld herangezogen (siehe Abschnitt 1.3, (1.73)). In diesem Fall ist die Lagrangefunktion gegeben durch

$$L = \frac{m}{2} (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - e\Phi(x, y, z) + e \begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix} \cdot \begin{pmatrix} A_x(x, y, z) \\ A_y(x, y, z) \\ A_z(x, y, z) \end{pmatrix}. \quad (1.108)$$

Der generalisierte Impuls  $p_x$  ergibt sich also in diesem Fall als

$$p_x = m\dot{x} + eA_x(x, y, z). \quad (1.109)$$

Ausserdem soll hier der Begriff **zyklische Koordinate** oder auch **ignorable Koordinate** eingeführt werden. Ist ein System durch die generalisierten Koordinaten  $q_1 \dots q_n$  definiert, so sagt man, dass die generalisierte Koordinate  $q_j$  zyklische ist genau dann, wenn die Lagrange Funktion nicht von dieser Koordinate abhängt. Es gilt also:

$$q_j \text{ ist zyklisch} \iff \frac{\partial L}{\partial q_j} = 0. \quad (1.110)$$

Damit können wir sofort den folgenden Satz beweisen:

**Satz:** Der zu einer zyklischen Koordinate  $q_j$  zugehörige generalisierte Impuls  $p_j$  ist eine Konstante der Bewegung.

Mit den Lagrangeschen Bewegungsgleichungen gilt ja

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} = \frac{d}{dt} p_j = \frac{\partial L}{\partial q_j} \stackrel{\text{siehe (1.110)}}{=} 0.$$

Ist also etwa im oben diskutierten Beispiel (1.106) des Teilchens auf dem Zylindermantel das Potenzial unabhängig von der Richtung die durch den Winkel  $\varphi$  definiert ist, so ist diese generalisierte Koordinate zyklisch. Es gilt also in diesem Fall, dass der zugehörige Impuls  $p_\varphi$ , der ja nach (1.107) die  $z$ -Komponente des Drehimpulses ist, eine Konstante der Bewegung ist. Ist also das Problem symmetrisch in Bezug auf eine Drehung um die  $z$ -Achse, so ist der zugehörige Bahndrehimpuls des Teilchens,  $l_z$ , eine Konstante der Bewegung. Dies gilt auch, wie man leicht sieht, wenn die Beschränkung der Bewegung auf die Oberfläche des Zylinders entfällt.

Eine andere Symmetrie liegt vor, wenn das System unabhängig ist gegenüber einer Verschiebung der Zeit, man sagt auch invariant gegenüber einer Translation der Zeit. Dies bedeutet, dass die Lagrangefunktion nicht explizit von der Zeit abhängt und auch die Zwangsbedingungen nicht zeitabhängig sind. Ist in diesem Fall das Potenzial  $V$  unabhängig von den Geschwindigkeiten der Teilchen, so ist die Gesamtenergie  $E$ , definiert als die Summe von kinetischer Energie der Teilchen und dem Wert des Potenzials  $V$

$$E = T + V = \text{konst.} \quad (1.111)$$

eine Konstante der Bewegung. In diesem Fall übernimmt also die Energie die Rolle der zur Koordinate Zeit  $t$  konjugierten Impulsvariable, eine Interpretation auf die wir noch zurückkommen werden.

Zum Beweis dieser Behauptung betrachten wir die totale Ableitung der Lagrangefunktion  $L$  nach der Zeit:

$$\begin{aligned} \frac{d}{dt}L(q_j\dot{q}_j, t) &= \sum_j \left( \frac{\partial L}{\partial q_j} \frac{dq_j}{dt} + \frac{\partial L}{\partial \dot{q}_j} \frac{d\dot{q}_j}{dt} \right) + \underbrace{\frac{\partial L}{\partial t}}_{=0 \text{ nach Voraus.}} \\ &= \sum_j \left( \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} \dot{q}_j + \frac{\partial L}{\partial \dot{q}_j} \frac{d\dot{q}_j}{dt} \right) \\ &= \sum_j \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_j} \dot{q}_j \right) \\ &= \frac{d}{dt} \sum_j p_j \dot{q}_j. \end{aligned}$$

Bei dem Übergang zur zweiten Zeile wurden die Lagrange Bewegungsgleichungen benutzt. Man kann also das Ergebnis dieser Rechnung zusammenfassen zu

$$\frac{d}{dt} \left( \sum_j p_j \dot{q}_j - L \right) = 0,$$

was ja bedeutet, dass

$$\sum_j p_j \dot{q}_j - L = \text{konstant}, \quad (1.112)$$

eine Konstante der Bewegung ist. Zur Berechnung dieser Konstanten wollen wir zunächst die kinetische Energie als Funktion der  $q_i$  und  $\dot{q}_i$  darstellen. Da die Zwangsbedingungen nach Voraussetzung nicht zeitabhängig sein sollen, ist auch die Darstellung der Ortsvektoren der einzelnen Massenpunkte als Funktion der generalisierten Koordinaten nicht explizit abhängig von der Zeit. Für die Geschwindigkeiten des Massenpunkte  $i$  ergibt sich dann

$$\frac{d\vec{r}_i}{dt} = \dot{\vec{r}}_i = \sum_j \frac{\partial \vec{r}_i}{\partial q_j} \dot{q}_j. \quad (1.113)$$

Dies führt zu einer kinetischen Energie der Form

$$T = \sum_i \frac{m_i}{2} \left( \sum_j \frac{\partial \vec{r}_i}{\partial q_j} \dot{q}_j \right)^2 = \sum_{j,k} \frac{\alpha_{jk}}{2} \dot{q}_j \dot{q}_k, \quad (1.114)$$

also eine quadratische Form in den Geschwindigkeiten  $\dot{q}_j$  mit Koeffizienten

$$\alpha_{jk} = \sum_i m_i \frac{\partial \vec{r}_i}{\partial q_j} \frac{\partial \vec{r}_i}{\partial q_k} = \alpha_{kj}, \quad (1.115)$$

die von den generalisierten Koordinaten  $q_l$  abhängen können, nicht aber von den zugehörigen Geschwindigkeiten  $\dot{q}_l$ . Damit berechnet sich

$$\begin{aligned} \frac{\partial T}{\partial \dot{q}_j} &= \alpha_{jj} \dot{q}_j + \sum_{j \neq k} \left( \frac{\alpha_{jk}}{2} \dot{q}_k + \frac{\alpha_{kj}}{2} \dot{q}_k \right) \\ &= \sum_k \alpha_{kj} \dot{q}_k. \end{aligned} \quad (1.116)$$

Für den Fall, dass das Potenzial nicht von den Geschwindigkeiten abhängt gilt dann für die generalisierten Impulse

$$p_j = \frac{\partial L}{\partial \dot{q}_j} = \frac{\partial T}{\partial \dot{q}_j} = \sum_k \alpha_{kj} \dot{q}_k,$$

und

$$\sum_j p_j \dot{q}_j = \sum_{j,k} \alpha_{kj} \dot{q}_k \dot{q}_j = 2T, \quad (1.117)$$

wie ein Vergleich mit (1.114) zeigt. Damit ist aber die Erhaltungsgröße aus (1.112)

$$\sum_j p_j \dot{q}_j - L = 2T - (T - V) = T + V = E$$

mit der Energie identifiziert. Die Energie ist also unter den gegebenen Voraussetzungen eine Konstante der Bewegung.

## 1.8 Harmonische Näherung

In diesem Abschnitt werden wir uns mit dem Lagrange Formalismus befassen für den Fall, dass das System nur leichte Auslenkungen aus seiner Ruhelage erfährt. Ein Beispiel hierfür sind Anregungsmoden für das Wassermolekül, so wie es in Abb. 1.8 dargestellt ist. Natürlich müssen wir uns darüber im Klaren sein, dass die Behandlung eines solchen atomaren Systems die Behandlung mit den Werkzeugen der Quantenmechanik erfordert. Die Behandlung im Rahmen der klassischen Mechanik, die wir hier anstreben kann also nicht zu quantitativen Ergebnissen führen. Allerdings werden wir bei dieser Diskussion Begriffe entwickeln, wie z.B. die Frequenzen der Normalschwingungen, die so auch für die Diskussion in der Quantenmechanik relevant sind. Diese Schwingungsfrequenzen der Normalmoden sind gerade auch die Frequenzen, bei denen das Wasser elektromagnetische Wellen absorbiert und damit, wie wir aus der Diskussion im Teil Optik der Vorlesung wissen, die interessanten Parameter zur Beschreibung des Frequenzverhaltens des Brechungsindex.

Das Wassermolekül kann auf verschiedene Arten angeregt werden. Es kann rotieren, es kann zu Schwingungen angeregt werden, bei denen sich der Winkel zwischen den beiden Wasserstoffatomen ändert, oder aber auch zu Schwingungen, bei denen sich der Abstand zwischen den Wasserstoffatomen und dem Sauerstoffatom ändert. Wir wollen uns auf diese Schwingungen der Abstände beschränken und werden deshalb als Beispiel im zweiten Teil dieses Paragraphen die einfache lineare Kette behandeln, wie sie als vereinfachtes Modell des Wassermoleküls im rechten Teil der Abb 1.8 dargestellt ist.

Dies ist ein sehr einfaches Modell für die Schwingungsmoden eines relativ einfachen Moleküls. Wir werden dabei sehen, wie wir aus den Schwingungsmoden etwas lernen können über die Masse der beteiligten Atome und den Kräften, die zwischen diesen wirken. Für das Wassermolekül ist die Frage des atomaren Aufbaus natürlich längst beantwortet. Aber in der aktuellen Forschung ist die Frage nach der Struktur von komplexen Molekülen im Bereich der Biologie von großem Interesse. Ein Zugang zur Untersuchung dieser Strukturen besteht darin, im Experiment die charakteristischen Frequenzen zu bestimmen, bei denen solche Moleküle elektromagnetische Wellen absorbieren. Dann versucht man Modelle für die Struktur dieser Moleküle zu entwickeln, die genau diese charakteristischen Frequenzen reproduzieren.

Charakteristische Frequenzen von komplexen Systemen sind aber nicht für die Physik auf atomarer Ebene interessant, sie spielen auch bei makroskopischen Objekten (Bauwesen: Stabilität von Bauten bei Erschütterungen, Astronomie: Schwingungsmoden von Sternen ...) eine große Rolle. Im ersten Teilabschnitt werden wir allgemein die Begriffe und die Technik für die Behandlung und Berechnung solcher Schwingungsmoden entwickeln, im zweiten Teil wird dann das Beispiel des linearisierten Wassermoleküls behandelt. Zum besseren Verständnis ist es vielleicht hilfreich für den Leser bei der Bearbeitung des allgemeinen Teils stets das konkrete Beispiel vor Augen zu behalten.

### 1.8.1 Kleine Schwingungen im Lagrange Formalismus

In diesem Abschnitt wird eine Methode entwickelt, mit der man ein System beschreibt, das durch  $n$  generalisierte Koordinaten  $q_i$  beschrieben wird und das lediglich kleine Ab-



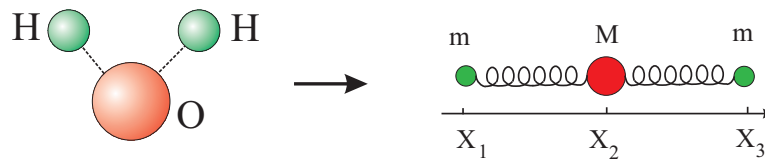


Abbildung 1.8: Modell des Wassermoleküls

weichungen von einem Punkt im Konfigurationsraum erfährt, bei dem das Potenzial (wir nehmen hier ein lokales, geschwindigkeitsunabhängiges Potenzial an) ein Minimum hat. Bezeichnen wir die Werte der generalisierten Koordinaten, bei denen  $V$  minimal ist, mit  $q_{i0}$ , so können wir das Potenzial in der Nähe dieses Minimums gut durch eine Taylorentwicklung um diesen Punkt beschreiben. Es gilt also

$$V(q_i) = V(q_{i0}) + \sum_j \frac{\partial V}{\partial q_j} (q_j - q_{j0}) + \sum_{j,k} \frac{1}{2} \frac{\partial^2 V}{\partial q_j \partial q_k} (q_j - q_{j0})(q_k - q_{k0}) + \dots \quad (1.118)$$

Für kleine Abweichungen der generalisierten Koordinaten

$$\eta_i = q_i - q_{i0}, \quad (1.119)$$

können wir die Terme dritter und höherer Ordnung in dieser Entwicklung vernachlässigen. Ausserdem gilt am Minimum, dass

$$\frac{\partial V}{\partial q_j} (q_{i0}) = 0,$$

so dass wir bei einer Normierung bei der  $V(q_{i0}) = 0$  gesetzt wird, das Potenzial aus (1.118) annähern können durch

$$V(\eta_i) = \sum_{j,k} \frac{1}{2} K_{jk} \eta_j \eta_k \quad \text{mit} \quad K_{jk} = \frac{\partial^2 V}{\partial q_j \partial q_k}. \quad (1.120)$$

Die Zahlen  $K_{jk}$  können wir als Elemente einer symmetrischen Matrix mit

$$K_{kj} = K_{jk}, \quad (1.121)$$

der Dimension  $n$  (das ist die Zahl der Freiheitsgrade) anordnen. Man bezeichnet diese  $K_{jk}$  als Elemente des **Kraftkonstantentensors**.

Als nächstes werden wir nun eine entsprechende Entwicklung bis zu quadratischen Gliedern in den Abweichungen  $\eta_i$  für die kinetische Energie vornehmen. Im Abschnitt 1.7 (1.114) haben wir gezeigt, dass man im Fall zeitunabhängiger Zwangsbedingungen die kinetische Energie eines Systems schreiben kann in der Form:

$$T = \sum_{j,k=1}^n \frac{\alpha_{jk}}{2} \dot{q}_j \dot{q}_k = \sum_{j,k=1}^n \frac{\alpha_{jk}}{2} \dot{\eta}_j \dot{\eta}_k, \quad (1.122)$$

mit

$$\begin{aligned} \alpha_{jk} &= \sum_{a=1}^N m_a \frac{\partial \vec{r}_a}{\partial q_j} \frac{\partial \vec{r}_a}{\partial q_k} \\ &= \alpha_{jk}(q_{i0}) + \sum_{l=1}^n \frac{\partial \alpha_{jk}}{\partial q_l} \underbrace{(q_l - q_{l0})}_{=\eta_l} + \dots \dots \end{aligned}$$

In der zweiten Zeile wurden die Funktionen  $\alpha_{jk}$ , die ja im allgemeinen Fall von den generalisierten Koordinaten  $q_i$  abhängen um den Punkt  $q_i = q_{i0}$  in einer Taylorreihe entwickelt. Setzt man aber diese Entwicklung in (1.122) ein, so sieht man, dass bereits das zweite Glied in dieser Entwicklung Beiträge zu  $T$  liefert, die von dritter Ordnung in den kleinen Größen  $\eta_i$  (bzw.  $\dot{\eta}_i$ ) sind, so dass wir unter konsequenter Vernachlässigung solcher Terme dritter Ordnung die kinetische Energie approximieren müssen durch

$$T = \sum_{j,k=1}^n \frac{M_{jk}}{2} \dot{\eta}_j \dot{\eta}_k \quad \text{mit} \quad M_{jk} := \alpha_{jk}(q_{i0}). \quad (1.123)$$

Auch die Zahlen  $M_{jk}$ , das sind die Werte der Funktionen  $\alpha_{jk}$  an der Stelle  $q_{i0}$  bilden eine symmetrische Matrix der Dimension  $n$ , den sogenannten **Massentensor**. Aus den Darstellungen (1.120) und (1.123) erhalten wir also die harmonische Näherung für die Lagrangefunktion

$$L(\eta_i, \dot{\eta}_i) = \sum_{j,k=1}^n \frac{M_{jk}}{2} \dot{\eta}_j \dot{\eta}_k - \frac{K_{jk}}{2} \eta_j \eta_k. \quad (1.124)$$

Dabei haben die Auslenkungen aus dem Minimum des Potentials,  $\eta_i$ , und die zugehörigen Geschwindigkeiten die Rolle der generalisierten Koordinaten und Geschwindigkeiten übernommen. Die Lagrange Bewegungsgleichungen lauten dann

$$\begin{aligned} \frac{d}{dt} \frac{\partial L}{\partial \dot{\eta}_i} &= \frac{d}{dt} \sum_j M_{ij} \dot{\eta}_j = \sum_j M_{ij} \ddot{\eta}_j \\ &= \frac{\partial L}{\partial \eta_i} = - \sum_j K_{ij} \eta_j. \end{aligned} \quad (1.125)$$

Fassen wir die einzelnen Auslenkungen zu einem Spaltenvektor der Dimension  $n$  zusammen (dabei soll der Doppelpfeil daran erinnern, dass wir hier Vektoren der Dimension  $n$  vorliegen haben)

$$\vec{\eta} = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix} \quad \text{bzw.} \quad \dot{\vec{\eta}} = \begin{pmatrix} \dot{\eta}_1 \\ \vdots \\ \dot{\eta}_n \end{pmatrix}, \quad (1.126)$$

so können wir die Bewegungsgleichung (1.125) auch umschreiben in die Vektorgleichung

$$M \ddot{\vec{\eta}} = -K \vec{\eta}, \quad (1.127)$$

wobei  $M$  ( $K$ ) hier für die Matrix des Massentensors (Kraftkonstantentensors) steht. Zur Lösung dieser Gleichung (die ja eigentlich für  $n$  gekoppelte Differenzialgleichungen steht) machen wir den Ansatz einer harmonischen Schwingung

$$\vec{\eta}(t) = \vec{\eta}_0 \exp(i\omega t), \quad (1.128)$$

wobei wir wie üblich natürlich nur an den Realteil dieser Gleichung interessiert sind. Setzt man diesen Ansatz in (1.127) ein, so ergibt sich

$$\underbrace{(K - \omega^2 M)}_{=B(\omega^2)} \vec{\eta}_0 = \vec{0}. \quad (1.129)$$

Diese Gleichung besitzt nur dann eine nichttriviale Lösung, wenn die Determinante der Matrix

$$\det B(\omega^2) = \det (K - \omega^2 M) = 0. \quad (1.130)$$

Wäre nämlich die Determinante der hier definierten Matrix  $B$  ungleich Null, so würde eine inverse Matrix  $B^{-1}$  existieren. Multipliziert man (1.129) von links mit  $B^{-1}$ , so führt das auf

$$B^{-1} B \vec{\eta}_0 = \vec{\eta}_0 = \vec{0},$$

was aber bedeutet, dass wir nur die triviale Lösung erhalten. Die Determinante der Matrix  $B$  ist ein Polynom  $n$ ten Grades in den gesuchten Frequenzen  $\omega^2$ . Gleichung (1.130) ist also eine Bestimmungsgleichung für diese Winkelgeschwindigkeiten  $\omega^2$ .<sup>2</sup>

Man wird also  $n$  Lösungen  $\omega_\alpha^2$  für (1.130) erhalten. Mit den Mitteln der Linearen Algebra kann man nachweisen, dass alle diese Lösungen reell sind und positiv

$$\omega_\alpha^2 \geq 0. \quad (1.131)$$

Aus der Sicht der Mathematik liegt dies daran, dass die Matrizen  $K$  und  $M$  symmetrisch sind und positiv definit. Dies bedeutet z.B. für  $K$ , dass  $V(\eta_i)$  in (1.120) für alle Werte der  $\eta_i$  positiv ist, was ja direkt damit verbunden ist, dass  $V$  an der Stelle  $\eta_i = 0$  (d.h.  $q_i = q_{i0}$ ) ein echtes Minimum aufweist.

Wir wollen aber an dieser Stelle auf den mathematischen Nachweis von (1.131) verzichten und dieses Ergebnis aus physikalischer Sicht erläutern. Wäre nämlich eine der Lösungen  $\omega_\alpha^2$  negativ oder komplex, so gäbe es für  $\omega_\alpha$  zwei komplexe Lösungen mit

$$\omega_{\alpha,1} = w_\alpha + iu_\alpha \quad \text{und} \quad \omega_{\alpha,2} = -w_\alpha - iu_\alpha.$$

Eine der beiden Lösung besäße also einen negativen Imaginärteil  $-u_\alpha$ . Dies führt aber nach (1.128) zu einer Zeitabhängigkeit der generalisierten Koordinaten von der Form

$$\eta_i(t) = \eta_{i0} \exp(-i\omega_\alpha t) \exp(u_\alpha t)$$

also eine Schwingung, bei der die Amplitude exponentiell mit der Zeit anwächst. Eine solche Bewegung ist aber für eine Schwingung um ein Minimum nicht zu erwarten.

Für positive Werte von  $\omega_\alpha^2$  liefert der Ansatz (1.128) eine harmonische Schwingung wie ein Oszillator mit einer reellen Winkelgeschwindigkeit  $\omega_\alpha$ . Man bezeichnet diese harmonische Schwingungen als **Normalschwingungen** oder auch **Normalmoden** des Systems.

Wir können also nun eine dieser Lösungen  $\omega_\alpha^2$  herausgreifen und bestimmen dafür durch die Lösung des linearen Gleichungssystems (siehe (1.129))

$$(K - \omega_\alpha^2 M) \vec{\eta}_{\alpha 0} = \vec{0}.$$

Bezeichnet man die Elemente dieses Vektors der Normalschwingung zur Frequenz  $\omega_\alpha$  mit  $A_{i\alpha}$ :

$$\vec{\eta}_{\alpha 0} = \begin{pmatrix} A_{1\alpha} \\ \vdots \\ A_{n\alpha} \end{pmatrix},$$

---

<sup>2</sup>Nimmt man an, dass die Matrix  $M$  der Einheitsmatrix entspricht, so ist (1.129) ein gewöhnliches Eigenwertproblem. Die Gleichung (1.130) entspricht dann der Aufgabe die Nullstellen des charakteristischen Polynoms von  $K$  zu finden. Deshalb bezeichnet man Gleichungen vom Typ (1.129) auch als verallgemeinertes Eigenwertproblem.

so schreibt sich dieses Gleichungssystem in der ausführlichen Schreibweise

$$\sum_{j=1}^n (K_{ij} - \omega_\alpha^2 M_{ij}) A_{j\alpha} = 0, \quad \text{für } i = 1 \dots n. \quad (1.132)$$

Wie wir (hoffentlich) alle aus der Linearen Algebra wissen, kann man diese Koeffizienten z.B. berechnen durch das Rezept

$$A_{j\alpha} = (-1)^{1+j} \det \left( [K - \omega_\alpha^2 M]^{1j} \right). \quad (1.133)$$

Dabei bezeichnet  $[K - \omega_\alpha^2 M]^{1j}$  die Matrix der Dimension  $n - 1$ , die sich aus  $(K - \omega_\alpha^2 M)$  ergibt, wenn man die erste Zeile streicht und die  $j$ -te Spalte.

Zum Beweis dieses Rechenrezeptes zeigen wir zunächst, dass (1.133) die erste Gleichung von (1.132 für  $i = 1$ ) erfüllt, da

$$\sum_{j=1}^n (K_{1j} - \omega_\alpha^2 M_{1j}) A_{j\alpha} = \det(K - \omega_\alpha^2 M) = 0,$$

gilt. Die erste Gleichung in dieser Zeile verifiziert man dadurch, dass man die Determinante von  $(K - \omega_\alpha^2 M)$  nach den Elementen der ersten Zeile entwickelt. Die zweite Gleichung ist dann gerade die Bestimmungsgleichung für die  $\omega_\alpha$  (1.130). Damit ist also die erste Zeile von (1.132) erfüllt.

Für  $i$  ungleich 1 liefert das Einsetzen von (1.133) in die  $i$ -te Zeile des Gleichungssystems (1.132) die Determinante der Matrix  $(K - \omega_\alpha^2 M)$ , wo allerdings die  $i$ -te Zeile durch die erste Zeile ersetzt ist. Die Determinante einer Matrix, in der 2 Zeilen identisch sind ist ebenfalls gleich 0, wodurch also die Gültigkeit des Rechenrezeptes (1.133) bestätigt wäre.

Nachdem wir nun wissen, wie man diese Eigenvektoren der Normalschwingungen berechnen kann, wollen wir noch einige Eigenschaften herausarbeiten. Wir betrachten dazu zwei Lösungen zu unterschiedlichen Winkelgeschwindigkeiten  $\omega_\alpha^2 \neq \omega_\beta^2$ . Die entsprechenden Gleichungssysteme liefern also

$$\begin{aligned} \sum_j K_{ij} A_{j\alpha} &= \omega_\alpha^2 \sum_j M_{ij} A_{j\alpha} \\ \sum_j K_{ij} A_{j\beta} &= \omega_\beta^2 \sum_j M_{ij} A_{j\beta} \end{aligned}$$

Multipliziert man die erste dieser Gleichungen mit  $A_{i\beta}$  und summiert diese über alle  $i$ , sowie die zweite mit  $A_{i\alpha}$  und summiert ebenfalls über  $i$ , ergibt sich

$$\begin{aligned} \sum_{i,j} K_{ij} A_{j\alpha} A_{i\beta} &= \omega_\alpha^2 \sum_{i,j} M_{ij} A_{j\alpha} A_{i\beta} \\ \sum_{i,j} K_{ij} A_{i\alpha} A_{j\beta} &= \omega_\beta^2 \sum_{i,j} M_{ij} A_{i\alpha} A_{j\beta}. \end{aligned} \quad (1.134)$$

Subtrahiert man diese Gleichungen voneinander und berücksichtigt, dass  $M_{ji} = M_{ij}$ , sowie  $K_{ji} = K_{ij}$ , so erhält man

$$0 = (\omega_\alpha^2 - \omega_\beta^2) \sum_{i,j} M_{ij} A_{i\alpha} A_{j\beta}.$$

Da der Faktor  $(\omega_\alpha^2 - \omega_\beta^2)$  nach Voraussetzung ungleich 0 ist, bedeutet dies, dass

$$\sum_{i,j} M_{ij} A_{i\alpha} A_{j\beta} = \vec{\eta}_{\alpha 0}^t M \vec{\eta}_{\beta 0} = 0, \quad \text{für } \alpha \neq \beta.$$

Die Vektoren  $\vec{\eta}_{\alpha 0}$  können nun noch so normiert werden, dass diese Multiplikation den Wert 1 ergibt für den Fall  $\alpha = \beta$ . Dies bedeutet also zusammengefasst

$$\sum_{i,j} M_{ij} A_{i\alpha} A_{j\beta} = \vec{\eta}_{\alpha 0}^t M \vec{\eta}_{\beta 0} = \delta_{\alpha\beta} \quad (1.135)$$

mit dem Kronecker Symbol  $\delta_{\alpha\beta}$ . Man kann diese Gleichung so interpretieren, dass die Vektoren  $\vec{\eta}_{\alpha 0}$  orthogonal sind bezüglich der Metrik, die durch den Massentensor  $M$  definiert ist<sup>3</sup>.

Man kann die Gleichung (1.135) aber natürlich auch als Gleichung für Matrixelemente mit den Zeilen- und Spaltenindizes  $\alpha$  und  $\beta$  ansehen. Dann ist diese Gleichung gleichbedeutend mit

$$A^t M A = 1, \quad (1.136)$$

was ja bedeutet, dass der Massentensor  $M$  durch die Transformation, die durch die Matrix  $A$  beschrieben wird diagonalisiert wird. Dabei ist  $A$  im Allgemeinen aber keine orthogonale Transformation. Die zu  $A$  inverse Matrix ist ja nicht  $A^t$  sondern

$$A^{-1} = A^t M.$$

Mit diesen Ergebnissen (speziell (1.135)) können wir nun z.B. die zweite Gleichung von (1.134) umschreiben auf

$$\sum_{i,j} K_{ij} A_{i\alpha} A_{j\beta} = \omega_\beta^2 \delta_{\alpha\beta}$$

was in Matrixschreibweise ja bedeutet

$$A^t K A = \begin{pmatrix} \omega_1^2 & & 0 \\ & \ddots & \\ 0 & & \omega_n^2 \end{pmatrix}. \quad (1.137)$$

Durch die Transformation mit  $A$  wird also auch der Kraftkonstantentensor  $K$  auf eine Matrix der Diagonalgestalt transformiert. In der Diagonale stehen die Quadrate der Frequenzen der Normalschwingungen  $\omega_\alpha$ . Durch die Transformation werden also gleichzeitig  $M$  und  $K$  diagonalisiert.

Mit diesen Normalschwingungen haben wir spezielle Lösungen der Bewegungsgleichungen gefunden. Die Bewegungsgleichungen sind lineare Gleichungen und deshalb ist auch jede Linearkombination dieser Normalschwingungen ein Lösung. Die Entwicklungskoeffizienten dieser Linearkombination müssen noch aus den Startbedingungen hergeleitet werden. Sind die Startbedingungen so, dass nur eine Normalschwingung des Systems angeregt wird, so wird die Bewegung des Systems als periodische Schwingung mit der entsprechenden Winkelgeschwindigkeit ablaufen. Im allgemeinen Fall wird man aber Startbedingungen haben,

<sup>3</sup>Für den einfacheren Fall des normalen Eigenwertproblems, bei dem  $M$  die Eins-Matrix ist, bedeutet dies natürlich, dass die Eigenvektoren orthogonal sind.

die nicht einer Normalschwingung entsprechen. Vielmehr definieren diese Startbedingungen eine spezielle Überlagerung von Normalschwingungen. Da diese Normalschwingungen verschiedenen Frequenzen besitzen, wird die Bewegung im Allgemeinen nicht periodisch ablaufen.

Wir müssen also Koeffizienten  $x_\alpha$  finden, so dass die Starbedingungen für die Auslenkungen

$$\begin{aligned}\eta_j(t=0) &= \Re \left\{ \sum_{\alpha} \chi_{\alpha} A_{j\alpha} \exp(i\omega_{\alpha} t) \right\} \\ &= \sum_{\alpha} \Re\{\chi_{\alpha}\} A_{j\alpha} \quad \text{und} \\ \dot{\eta}_j(t=0) &= \Re \left\{ \sum_{\alpha} i\omega_{\alpha} \chi_{\alpha} A_{j\alpha} \exp(i\omega_{\alpha} t) \right\} \\ &= - \sum_{\alpha} \omega_{\alpha} \Im\{\chi_{\alpha}\} A_{j\alpha} \quad \text{für } j = 1 \dots n .\end{aligned}\tag{1.138}$$

Dabei steht  $\Re$  für den Realteil und  $\Im$  für den Imaginärteil des Ausdruckes im Argument der Funktion. Nehmen wir z.B. an, dass das System zur Startzeit in Ruhe sein soll, also  $\dot{\eta}_j(t=0) = 0$  so können wir reelle Werte für die Koeffizienten  $\chi_{\alpha}$  annehmen und müssen lediglich das erste Gleichungssystem lösen. Sind dies  $\chi_{\alpha}$  bestimmt, so ergibt sich dann in diesem Fall die Zeitabhängigkeit der Auslenkungen durch

$$\eta_j(t) = \sum_{\alpha} \Re\{\chi_{\alpha}\} A_{j\alpha} \cos(\omega_{\alpha} t) .\tag{1.139}$$

### 1.8.2 Beispiel: Modell des Wassermoleküls

Als Beispiel für die Anwendung der Techniken, die wir gerade entwickelt haben, betrachten wir das eindimensionale Modell des Wassermoleküls, das im rechten Teil der Abb. 1.8 skizziert ist. Als generalisierte Koordinaten bieten sich natürlich die  $x$  Koordinaten des Systems an. Die potenzielle Energie des Kraftfeldes, das durch die beiden Federn dargestellt ist ergibt sich dann als

$$V(x_i) = \frac{k}{2} (x_2 - x_1 - r_0)^2 + \frac{k}{2} (x_3 - x_2 - r_0)^2 ,$$

wobei  $r_0$  den Abstand der Atompaaire bezeichnet bei denen das Potenzial minimal ist. Bezeichnen wir wie vorher die Abweichung der generalisierten Koordinaten  $x_i$  von der Ruhelage bei der das Potenzial verschwindet mit

$$\eta_i = x_i - x_{i0} ,$$

so ergibt sich für die potenzielle Energie

$$\begin{aligned}V(\eta_j) &= \frac{k}{2} (\eta_2 - \eta_1)^2 + \frac{k}{2} (\eta_3 - \eta_2)^2 \\ &= \frac{k}{2} (\eta_1^2 + 2\eta_2^2 + \eta_3^2 - \eta_1\eta_2 - \eta_2\eta_1 - \eta_2\eta_3 - \eta_3\eta_2) \\ &= \sum_{i,j} \frac{1}{2} K_{ij} \eta_i \eta_j\end{aligned}$$

mit einem Kraftkonstantentensor

$$K = \begin{pmatrix} k & -k & 0 \\ -k & 2k & -k \\ 0 & -k & k \end{pmatrix}. \quad (1.140)$$

In diesem Fall müssen wir also den Kraftkonstantentensor nicht als Näherung (vergl. (1.120)) bestimmen sondern erhalten ihn wegen der Einfachheit des Potentials direkt.

Die kinetische Energie berechnet sich zu

$$\begin{aligned} T &= \frac{m}{2} \dot{x}_1^2 + \frac{M}{2} \dot{x}_2^2 + \frac{m}{2} \dot{x}_3^2 \\ &= \frac{m}{2} \dot{\eta}_1^2 + \frac{M}{2} \dot{\eta}_2^2 + \frac{m}{2} \dot{\eta}_3^2 \\ &= \sum_{i,j} \frac{1}{2} M_{ij} \dot{\eta}_i \dot{\eta}_j, \end{aligned}$$

mit einem Massentensor

$$M = \begin{pmatrix} m & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & m \end{pmatrix}. \quad (1.141)$$

Auch die Darstellung der kinetischen Energie über den Massentensor ergibt sich also in diesem Fall ohne weitere Näherung. Daraus erhalten wir dann das charakteristische Polynom (vergleiche (1.130) und die Berechnung mit MAPLE im nächsten Abschnitt)

$$\begin{aligned} \det(K - \omega^2 M) &= \det \begin{pmatrix} k - \omega^2 m & -k & 0 \\ -k & 2k - \omega^2 M & -k \\ 0 & -k & k - \omega^2 m \end{pmatrix} \\ &= -\omega^6 m^2 M + 2k\omega^4(m^2 + Mm) - \omega^2 k^2(m + M) \\ &= 0, \end{aligned} \quad (1.142)$$

mit den Nullstellen

$$\begin{aligned} \omega_1^2 &= 0 \\ \omega_2^2 &= \frac{k}{m} \\ \omega_3^2 &= \frac{k(2m + M)}{Mm}. \end{aligned} \quad (1.143)$$

Betrachten wir zunächst einmal die Lösung  $\omega_1^2 = 0$ . Entsprechend (1.132) ergibt sich der entsprechende Eigenvektor aus

$$K \vec{\eta}_{10} = \begin{pmatrix} k & -k & 0 \\ -k & 2k & -k \\ 0 & -k & k \end{pmatrix} \begin{pmatrix} A_{11} \\ A_{21} \\ A_{31} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Man findet dass alle Vektoren

$$\vec{\eta}_{10} = \begin{pmatrix} A_{11} \\ A_{21} \\ A_{31} \end{pmatrix} = \chi \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad (1.144)$$

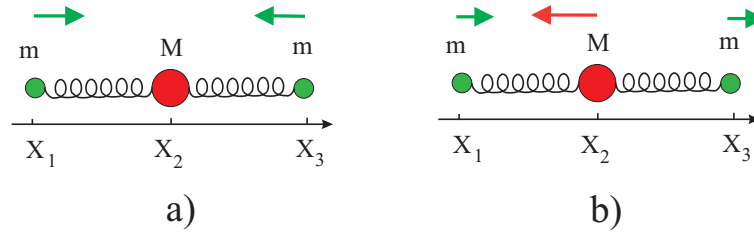


Abbildung 1.9: Normalschwingungen des eindimensionalen Wassermoleküls

mit beliebigem Wert für  $\chi$  eine Lösung darstellen. Bei dieser Lösung werden also alle 3 Massenpunkte des Systems immer um den gleichen Betrag verschoben. Es handelt sich also um eine Verschiebung des Gesamtsystems, des gesamten Moleküls. In unserem Modell gibt es keine Kraft, die versucht das Molekül auf seinen Platz zu halten. Die Rückstellkraft ist also identisch Null und deshalb ist auch die Frequenz dieser Bewegung des Gesamtsystems gleich Null.

Für die zweite Frequenz einer Normalschwingung,  $\omega_2^2 = k/m$  in (1.143), ergibt sich der Lösungsvektor aus

$$\left(K - \frac{k}{m}M\right) \vec{\eta}_{20} = 0,$$

zu

$$\vec{\eta}_{20} = \begin{pmatrix} A_{12} \\ A_{22} \\ A_{32} \end{pmatrix} = \chi \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}. \quad (1.145)$$

Bei dieser Schwingung bleibt also das mittlere Atom der Masse  $M$  in Ruhe und die beiden äusseren Atome schwingen mit entgegengesetzter Richtung, so wie das in dem linken Teil der Abb. 1.9 dargestellt ist. Der Schwerpunkt des Systems bleibt dabei unverändert. Diese Normalschwingung ist also vollkommen entkoppelt von der Bewegung des Schwerpunktes, die wir bei der ersten Normalschwingung vorliegen hatten.

Für die dritte Normalschwingung mit der Frequenz  $\omega_3$  aus (1.143) ergibt sich der Lösungsvektor zu

$$\vec{\eta}_{30} = \begin{pmatrix} A_{13} \\ A_{23} \\ A_{33} \end{pmatrix} = \chi \begin{pmatrix} 1 \\ -\frac{2m}{M} \\ 1 \end{pmatrix}. \quad (1.146)$$

Die äusseren Atome werden in die gleiche Richtung ausgelenkt, das mittlere Atome in die Gegenrichtung (siehe rechter Teil der Abb. 1.9).

Zur weiteren Konkretisierung wird in den Rechnungen zum dem Beispiel mit Maple angenommen, dass

$$\frac{k}{m} = 1 \quad \text{und} \quad M = 2m. \quad (1.147)$$

Insbesondere wird für diesen Fall die explizite Lösung bestimmt für Startbedingungen, bei denen die drei Teilchen zur Zeit  $t = 0$  ruhen und das Teilchen 1 um eine Längeneinheit aus der optimalen Lagen verschoben ist. Daraus ergeben sich entsprechend der Nomenklatur von (1.138) Koeffizienten  $x_\alpha = (1/4, -1/2, 1/4)$  für  $\alpha = 1, 2, 3$ . Die Auslenkungen der 3 Teilchen als Funktion der Zeit ist dann in der Abbildung am Ende dieses gesamten Abschnittes dargestellt.



### 1.8.3 Die Lösung des Beispiels mit Maple

```
> with(linalg):
```

```
Warning, new definition for norm
```

```
Warning, new definition for trace
```

```
> assume(k>0);
```

```
> assume(m>0);
```

```
> assume(M>0);
```

**Definition der Tensoren K und M (siehe (1.140) und (1.140))**

```
> m1:=matrix([[k,-k,0],[-k,2*k,-k],[0,-k,k]]);
```

$$m1 := \begin{bmatrix} k^{\sim} & -k^{\sim} & 0 \\ -k^{\sim} & 2k^{\sim} & -k^{\sim} \\ 0 & -k^{\sim} & k^{\sim} \end{bmatrix}$$

```
> m2:=matrix([[m,0,0],[0,M,0],[0,0,m]]);
```

$$m2 := \begin{bmatrix} m^{\sim} & 0 & 0 \\ 0 & M^{\sim} & 0 \\ 0 & 0 & m^{\sim} \end{bmatrix}$$

```
> m3:=m1-x*m2;
```

$$m3 := m1 - x m2$$

**Berechnung der Nullstellen des charakteristischen Polynoms (1.142)**

```
> det(m3);
```

$$-2k^{\sim 2} x m^{\sim} - x M^{\sim} k^{\sim 2} + 2k^{\sim} x^2 M^{\sim} m^{\sim} + 2x^2 m^{\sim 2} k^{\sim} - x^3 m^{\sim 2} M^{\sim}$$

```
> solve(det(m3)=0,x);
```

$$0, \frac{k^{\sim}}{m^{\sim}}, \frac{k^{\sim} (2m^{\sim} + M^{\sim})}{M^{\sim} m^{\sim}}$$

**Bestimmung der Normalschwingungen nach (1.132)**

```
> b:=vector([0,0,0]);
```

$$b := [0, 0, 0]$$

```
> n1:=linsolve(m1,b);
```

$$n1 := [-t_1, -t_1, -t_1]$$

```
> n2:=linsolve(m1-k/m*m2,b);
```

$$n2 := [-t_1, 0, -t_1]$$

```
> n3:=linsolve(m1-k*(2*m+M)/(M*m)*m2,b);
```

$$n3 := \left[ t_1, -2 \frac{-t_1 m^{\sim}}{M^{\sim}}, -t_1 \right]$$

```
> m4:=transpose(matrix([n1,n2,n3]));
```

$$m4 := \begin{bmatrix} -t_1 & -t_1 & -t_1 \\ -t_1 & 0 & -2 \frac{-t_1 m^{\sim}}{M^{\sim}} \\ -t_1 & -t_1 & -t_1 \end{bmatrix}$$

```
> start:=vector([1,0,0]);
```

$start := [1, 0, 0]$

>  $s1 := \text{linsolve}(m4, start);$

$$s1 := \left[ \frac{m\tilde{}}{-t_1(2m\tilde{}} + M\tilde{)}, -\frac{1}{2} \frac{1}{-t_1}, \frac{1}{2} \frac{M\tilde{}}{-t_1(2m\tilde{}} + M\tilde{)} \right]$$

### Konkretisierung des Beispiels nach (1.147)

>  $m:=1;$

$m := 1$

>  $M:=2;$

$M := 2$

>  $k:=1;$

$k := 1$

>  $m1 := \text{matrix}([[k, -k, 0], [-k, 2*k, -k], [0, -k, k]]);$

$$m1 := \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

>  $m2 := \text{matrix}([[m, 0, 0], [0, M, 0], [0, 0, m]]);$

$$m2 := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

>  $m3 := m1 - x*m2;$

$m3 := m1 - x m2$

>  $\text{det}(m3);$

$-4x + 6x^2 - 2x^3$

>  $\text{solve}(\text{det}(m3)=0, x);$

$0, 1, 2$

>  $b := \text{vector}([0, 0, 0]);$

$b := [0, 0, 0]$

>  $n1 := \text{linsolve}(m1, b);$

$n1 := [-t_1, -t_1, -t_1]$

>  $n2 := \text{linsolve}(m1 - k/m*m2, b);$

$n2 := [-t_1, 0, -t_1]$

>  $n3 := \text{linsolve}(m1 - k*(2*m+M)/(M*m)*m2, b);$

$n3 := [t_1, -t_1, -t_1]$

>  $m4 := \text{transpose}(\text{matrix}([n1, n2, n3]));$

$$m4 := \begin{bmatrix} -t_1 & -t_1 & -t_1 \\ -t_1 & 0 & -t_1 \\ -t_1 & -t_1 & -t_1 \end{bmatrix}$$

>  $start := \text{vector}([1, 0, 0]);$

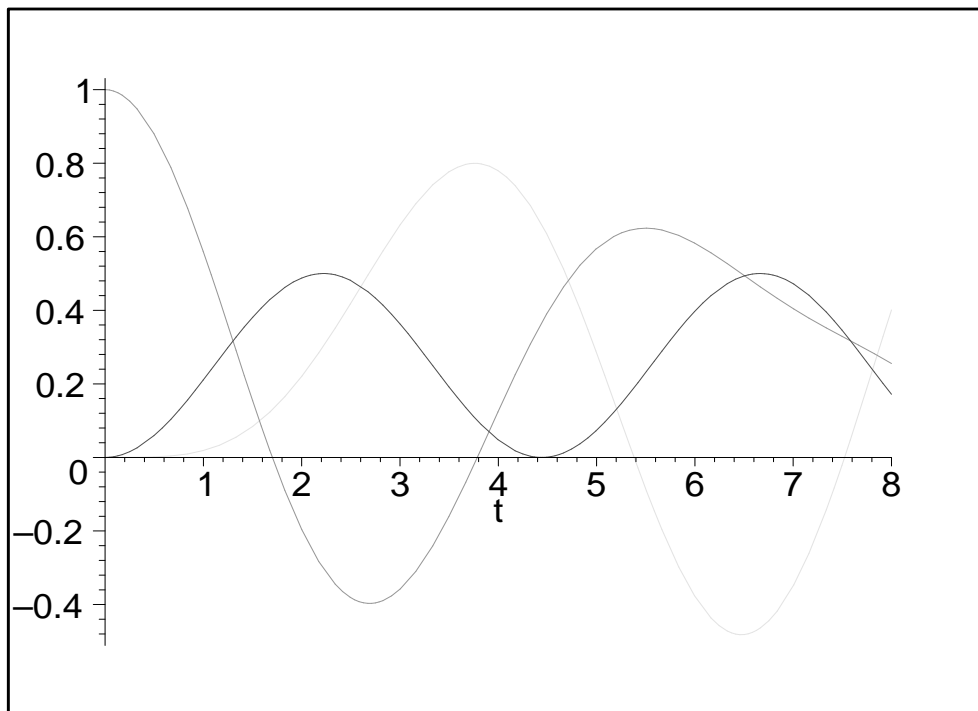
$start := [1, 0, 0]$

>  $s1 := \text{linsolve}(m4, start);$

$$s1 := \left[ \frac{1}{4} \frac{1}{t_1}, -\frac{1}{2} \frac{1}{t_1}, \frac{1}{4} \frac{1}{t_1} \right]$$

Darstellung der Auslenkungen der einzelnen Teilchen aus der Ruhelage als Funktion der Zeit

```
> plot(
> {s1[1]*n1[1]+s1[2]*n2[1]*cos(t)+s1[3]*n3[1]*cos(sqrt(2)*t),s1[1]*n1[2]
> ]+s1[2]*n2[2]*cos(t)+s1[3]*n3[2]*cos(sqrt(2)*t),s1[1]*n1[3]+s1[2]*n2[3]
> ]*cos(t)+s1[3]*n3[3]*cos(sqrt(2)*t)},t=0..8);
```





# Kapitel 2

## Hamilton Formalismus

### 2.1 Hamiltonsche Bewegungsgleichungen

Den stationären Zustand eines Systems mit  $n$  Freiheitsgraden, also z.B. die aktuellen Positionen der Massenpunkte eines Vielteilchensystems zu einem festen Zeitpunkt, können wir durch  $n$  generalisierte Koordinaten  $q_i$  beschreiben. Wenn wir die zeitliche Entwicklung etwa mit Hilfe des Lagrange Formalismus berechnen wollen, treten in der Lagrange Funktion neben den Koordinaten auch die zugehörigen Geschwindigkeiten  $\dot{q}_i$  auf. Die Lagrange Funktion  $L$  hängt also von  $2 * n$  Variablen ab, sie ist in einem Raum der Dimension  $2n$  definiert, den wir als **Phasenraum** bezeichnen.

Die Lagrangeschen Bewegungsgleichungen führen auf ein System von  $n$  gekoppelten Differenzialgleichungen zweiter Ordnung in der Zeit für die Funktionen  $q_i(t)$ . Zur Bestimmung einer eindeutigen Lösung müssen wir noch  $2n$  Randbedingungen vorgeben. Dies geschieht im Allgemeinen dadurch, dass wir für einen Startzeitpunkt  $t_0$  die Werte für die Koordinaten  $q_i(t_0)$  und die Geschwindigkeiten  $\dot{q}_i(t_0)$  vorgeben. Wir legen also einen Startpunkt in diesem  $2n$  dimensionalen Phasenraum fest.

Man kann jetzt versuchen, die Basis des Phasenraumes so zu optimieren, dass die Bewegungsgleichungen besonders einfach werden also z.B. voneinander entkoppeln. Dazu kann man geeignetere generalisierte Koordinaten,  $Q_j$ , suchen die als Funktionen der  $q_i$  gegeben sind. Man spricht in diesem Fall von einer Punkttransformation, da man den aktuellen Zustandspunkt eines Systems durch neue Koordinaten  $Q_j$  beschreibt. Durch diese Punkttransformation der generalisierten Koordinaten sind auch die zugehörigen Geschwindigkeiten  $\dot{Q}_j$  definiert und wir müssen die Lagrange Funktion als Funktion von  $Q_j$  und  $\dot{Q}_j$  bestimmen um die Lagrangeschen Bewegungsgleichungen auszuwerten. Die möglichen Transformationen für die Darstellung des  $2n$  dimensionalen Phasenraums sind also auf die  $n$  generalisierten Koordinaten beschränkt.

Die gekoppelten Bewegungsgleichungen werden natürlich besonders effizient vereinfacht, wenn einzelne Variable der Bewegung als Konstanten der Bewegung einen festen Wert zugewiesen bekommen. Wir haben im Abschnitt 1.7 über Konstanten der Bewegung gesehen, dass in dem Fall wo die Lagrange Funktion unabhängig von einer der generalisierten Koordinaten ist, der zugehörige kanonische Impuls eine Konstante der Bewegung ist. Ist also z.B. die Koordinate  $q_\alpha$  eine zyklische Koordinate - das bedeutet ja gerade, dass  $L$

nicht von  $q_\alpha$  abhängt, so gilt für den zugehörigen Impuls

$$p_\alpha = \frac{\partial L}{\partial \dot{q}_\alpha}, \quad (2.1)$$

dass er eine Konstante der Bewegung ist. Es wäre also hilfreich, wenn wir die Bewegungsgleichungen auf Gleichungen für die generalisierten Impulse transformieren könnten. Dann könnte man sofort die Gleichungen für die Impulse zu zyklischen Koordinaten abkoppeln.

Das Ziel ist es an Stelle der Lagrangefunktion  $L(\dot{q}_i, q_i, t)$  eine neue Funktion  $H$  zu finden, die auch im  $2n$  dimensionalen Phasenraum definiert ist, allerdings von den generalisierten Koordinaten  $q_i$  und den zugehörigen Impulsen  $p_i$  abhängt:  $H(p_i, q_i, t)$ . Wir können uns dann in dem Fall, dass  $p_\alpha$  eine Konstante der Bewegung ist, sofort auf den Teil des Phasenraumes beschränken, der durch einen konstanten Wert für  $p_\alpha$  definiert ist. Dadurch wird die Anzahl der gekoppelten Gleichungen um eine reduziert.

Man erreicht eine solche Transformation, bei der eine Variable  $q_\alpha$  durch den zugehörigen Impuls ersetzt wird, durch eine **Legendre Transformation**. Zur allgemeinen Formulierung einer Legendre Transformation nehmen wir eine Funktion  $f(x, y)$  an. Man erhält aus dieser Funktion eine neue Funktion, die ausschliesslich von den Variablen

$$x \quad \text{und} \quad \psi := \frac{\partial f}{\partial y}, \quad (2.2)$$

abhängt, in dem man definiert

$$g := \psi y - f(x, y). \quad (2.3)$$

Zum Beweis, dass diese Funktion  $g$  nur von  $x$  und  $\psi$  abhängt, betrachten wir Differenzialformen. Eine Differenzialform  $df$  einer Funktion  $f$  ist definiert als die Summe über alle Variable von der  $f$  abhängt. Jeder Summand enthält die partielle Ableitung dieser Funktion nach der jeweiligen Variablen, multipliziert mit der Differenzialform für die Variable. Es gilt also für die Funktion  $f$ , die nur von  $x$  und  $y$  abhängen soll, dass die Differenzialform  $df$  gegeben ist durch

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy. \quad (2.4)$$

Berechnen wir also z.B. die Differenzialform für die Funktion  $h = \psi y$ , so ergibt sich

$$\begin{aligned} dh &= \frac{\partial h}{\partial \psi} d\psi + \frac{\partial h}{\partial y} dy \\ &= y d\psi + \psi dy. \end{aligned}$$

Damit berechnen wir

$$\begin{aligned} dg &= d(\psi y) - df \\ &= y d\psi + \psi dy - \frac{\partial f}{\partial x} dx - \underbrace{\frac{\partial f}{\partial y}}_{=\psi} dy \\ &= y d\psi - \frac{\partial f}{\partial x} dx. \end{aligned} \quad (2.5)$$

Aus dieser Gleichung können wir ablesen, dass  $g$  nur von den unabhängigen Variablen  $\psi$  und  $x$  abhängt mit

$$\frac{\partial g}{\partial \psi} = y \quad \text{und} \quad \frac{\partial g}{\partial x} = -\frac{\partial f}{\partial x}.$$

Diese Technik der Legendre Transformation wenden wir nun an, um in der Darstellung des Phasenraumes von den  $2n$  Koordinaten  $q_i$  und  $\dot{q}_i$  auf die generalisierten Koordinaten  $q_i$  und  $p_i$  zu wechseln. Dazu wenden wir die Legendre Transformation für alle  $n$  Freiheitsgrade an und definieren die **Hamilton Funktion** des Systems durch

$$H := \sum_{i=1}^n p_i \dot{q}_i - L(q_i, \dot{q}_i, t), \quad (2.6)$$

mit den generalisierten Impulsen definiert in (2.1). Zum Beweis, dass  $H$  nur von den  $q_i$  und  $p_i$  abhängt, berechnen wir wieder die Differenzialform

$$\begin{aligned} dH &= \sum_{i=1}^n \left( p_i d\dot{q}_i + \dot{q}_i dp_i - \underbrace{\frac{\partial L}{\partial \dot{q}_i}}_{=p_i} d\dot{q}_i - \frac{\partial L}{\partial q_i} dq_i \right) - \frac{\partial L}{\partial t} dt \\ &= \sum_{i=1}^n \left( \dot{q}_i dp_i - \frac{\partial L}{\partial q_i} dq_i \right) - \frac{\partial L}{\partial t} dt \\ &= \sum_{i=1}^n \left( \frac{\partial H}{\partial p_i} dp_i + \frac{\partial H}{\partial q_i} dq_i \right) + \frac{\partial H}{\partial t} dt. \end{aligned} \quad (2.7)$$

Aus der zweiten Zeile können wir ablesen, dass die Hamiltonfunktion  $H$  lediglich von  $p_i$ ,  $q_i$  und  $t$  abhängt, da ja andere Terme nicht in der Darstellung von  $dH$  auftauchen. Die dritte Gleichung von (2.7) ist dann die allgemeine Darstellung der Differenzialform von  $H$  und so können wir aus dem Vergleich der zweiten mit der dritten Zeile entnehmen, dass

$$\begin{aligned} \frac{\partial H}{\partial p_i} &= \dot{q}_i \\ \frac{\partial H}{\partial q_i} &= -\frac{\partial L}{\partial q_i} = -\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = -\frac{d}{dt} p_i \\ \frac{\partial H}{\partial t} &= -\frac{\partial L}{\partial t} \end{aligned} \quad (2.8)$$

In der zweiten Zeile haben wir die Lagrange Bewegungsgleichung und die Definition des generalisierten Impulses benutzt. Die ersten beiden Zeilen fast man zusammen zu den **Hamiltonschen Bewegungsgleichungen** oder auch **Kanonischen Bewegungsgleichungen** in der Form

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \quad \text{und} \quad \dot{p}_i = -\frac{\partial H}{\partial q_i} \quad \text{für } i = 1 \dots n. \quad (2.9)$$

Dabei handelt es sich also um  $2n$  Differenzialgleichungen erster Ordnung in der Zeit. Dies ist direkt vergleichbar mit den  $n$  Differenzialgleichungen zweiter Ordnung in der Zeit, die wir im Fall der Lagrange Bewegungsgleichungen zu lösen haben. Wir wissen ja, dass man solche Differenzialgleichungen zweiter Ordnung umschreiben kann in jeweils 2 Differenzialgleichungen erster Ordnung, so dass der Aufwand zur Lösung der Bewegungsgleichung im Fall des Hamilton Formalismus im Allgemeinen vom gleichen Umfang sein wird wie der im Fall des Lagrange Formalismus. Vorteile bietet der Hamilton Formalismus nur in so fern, dass durch die  $q_i$  und  $p_i$ , wie wir im übernächsten Abschnitt sehen werden, eine flexiblere Darstellung des Phasenraumes ermöglicht wird.

Die Bestimmung der zeitlichen Entwicklung eines Systems mit Hilfe des Hamilton Formalismus lässt sich in die folgenden Schritte gliedern:

- Bestimme die Lagrange Funktion  $L(q_i, \dot{q}_i, t)$ .
- Diese erlaubt die Definition der generalisierten Impulse

$$p_i = \frac{\partial L}{\partial \dot{q}_i}.$$

- Damit wird die Hamiltonfunktion definiert durch

$$H(p_i, q_i, t) = \sum_{i=1}^n p_i \dot{q}_i - L.$$

Dabei ist darauf zu achten, dass  $H$  als Funktion der  $p_i$  und  $q_i$  bestimmt wird.

- Dies erlaubt die Aufstellung der Hamiltonschen Bewegungsgleichung (2.9) und man ist aufgefordert dies  $2n$  gekoppelten Differenzialgleichungen erster Ordnung zu lösen.
- Die spezielle Lösung dieser Differenzialgleichungen muss dann durch  $2n$  Rand- oder Startbedingungen (z.B. die Werte der  $q_i$  und  $p_i$  zur Startzeit  $t = 0$ ) eindeutig festgelegt werden.

Im verbleibenden Teil dieses Abschnittes sollen die Hamiltonschen Bewegungsgleichungen direkt aus dem Hamiltonschen Prinzip hergeleitet werden. Das Hamiltonsche Prinzip besagt ja, dass die Natur die Trajektorie von einem gegebenen Startpunkt ( $q_i(t_1)$  sind vorgegeben) zu einem gegebenen Endpunkt (auch die  $q_i$  zur Endzeit  $t_2$  sind vorgegeben) verfolgt, für die das Wirkungsintegral

$$S = \int_{t_1}^{t_2} L(q_i, \dot{q}_i, t) dt$$

ein Extremum ist. Wenn man also die Lagrange Funktion  $L$  gemäß (2.6) umschreibt, ergibt sich das **modifizierte Hamiltonsche Prinzip** durch die Forderung

$$\delta S = \delta \int_{t_1}^{t_2} \left[ \sum_{i=1}^n p_i \dot{q}_i - H(p_i, q_i, t) \right] dt = 0. \quad (2.10)$$

Dabei werden die Wege durch die Angaben der generalisierten Koordinaten und der Impulskoordinaten parameterisiert. Variationen um den "richtigen Weg ( $q_{i0}, p_{i0}$ ) wollen wir parameterisieren in der Form

$$\begin{aligned} q_{i\alpha}(t) &= q_{i0}(t) + \alpha \eta_i(t) \\ p_{i\alpha}(t) &= p_{i0}(t) + \alpha \epsilon_i(t) \end{aligned} \quad (2.11)$$

wobei die Randbedingungen durch die Forderung

$$\eta_i(t_1) = \eta_i(t_2) = 0 \quad \text{für } i = 1 \dots n, \quad (2.12)$$

vorgegeben sind. Für die Variationen der Impulskoordinaten, dargestellt durch die  $\epsilon_i(t)$  gibt es keine weiteren Randbedingungen, da ja durch die  $2n$  Bedingungen von (2.12) bereits alle Freiheiten ausgeschöpft sind. Ansonsten sind aber die Orts- und Impulskoordinaten bei diesem modifizierten Hamiltonschen Prinzip vollkommen gleichberechtigte Variable.



Mit der Parameterisierung der Trajektorien aus (2.11) liefert das modifizierte Hamiltonsche Prinzip (2.10)

$$\begin{aligned}
\delta S = 0 &= \frac{\partial}{\partial \alpha} \int_{t_1}^{t_2} \left[ \sum_{i=1}^n p_{i\alpha} \dot{q}_{i\alpha} - H(p_{i\alpha}, q_{i,t\alpha}) \right] dt \\
&= \int_{t_1}^{t_2} \sum_{i=1}^n \left[ \frac{\partial p_{i\alpha}}{\partial \alpha} \dot{q}_{i\alpha} + \frac{\partial \dot{q}_{i\alpha}}{\partial \alpha} p_{i\alpha} - \frac{\partial H}{\partial q_{i\alpha}} \frac{\partial q_{i\alpha}}{\partial \alpha} - \frac{\partial H}{\partial p_{i\alpha}} \frac{\partial p_{i\alpha}}{\partial \alpha} \right] dt \\
&= \int_{t_1}^{t_2} \sum_{i=1}^n \left[ \epsilon_i \dot{q}_{i\alpha} + \frac{d\eta_i}{dt} p_{i\alpha} - \frac{\partial H}{\partial q_{i\alpha}} \eta_i - \frac{\partial H}{\partial p_{i\alpha}} \epsilon_i \right] dt. \tag{2.13}
\end{aligned}$$

Bei dem Übergang zur letzten Zeile wurde die Parameterisierung (2.11) benutzt um die Ableitungen nach  $\alpha$  zu berechnen. Zur weiteren Bearbeitung betrachten wir den 2. Term in dieser Gleichung und formen ihn mit Hilfe der partiellen Integration um in

$$\int_{t_1}^{t_2} \frac{d\eta_i}{dt} p_{i\alpha} dt = \underbrace{\eta_i p_{i\alpha} \Big|_{t_1}^{t_2}}_{=0} - \int_{t_1}^{t_2} \frac{dp_{i\alpha}}{dt} \eta_i dt,$$

(die Null im ersten Term der linken Seite folgt aus (2.12)). Setzt man dieses Ergebnis in (2.13) ein und berücksichtigt, dass die Ableitung nach  $\alpha$  an der Stelle  $\alpha = 0$  erfolgt, so dass wir die Indices  $i\alpha$  in  $i0$  umwandeln können, beziehungsweise einfach  $i$  schreiben, ergibt sich

$$0 = \int_{t_1}^{t_2} \sum_{i=1}^n \left[ \epsilon_i \left( \dot{q}_i - \frac{\partial H}{\partial p_i} \right) - \eta_i \left( \dot{p}_i + \frac{\partial H}{\partial q_i} \right) \right] dt. \tag{2.14}$$

Die Funktionen  $\eta_i$  und  $\epsilon_i$  im Integranden sind beliebige Funktionen der Zeit. Wir können also z.B. alle Funktionen identisch null wählen, es sei lediglich

$$\eta_j(t) = \delta(t - t_0)$$

die Diracsche  $\delta$  Funktion mit einer beliebigen Zeit  $t_0$  aus dem Intervall  $[t_1, t_2]$ . Setzt man diese Variationsfunktionen  $\eta_i$  und  $\epsilon_i$  in (2.14) ein, so ergibt sich für den Zeitpunkt  $t = t_0$

$$0 = \dot{p}_j + \frac{\partial H}{\partial q_j}.$$

Entsprechend gilt für Variationsfunktionen, die alle identisch 0 sind bis auf  $\epsilon_j$ , die wieder einer Funktion  $\delta(t - t_0)$  entspricht. Eingesetzt in (2.14) ergibt sich für für den beliebigen Zeitpunkt  $t = t_0$

$$0 = \dot{q}_j - \frac{\partial H}{\partial p_j}.$$

Wir haben also so direkt aus dem modifizierten Hamiltonschen Prinzip die Hamiltonschen Bewegungsgleichungen (2.9) hergeleitet.

## 2.2 Bedeutung der Hamiltonfunktion und Beispiele

Nachdem wir im letzten Abschnitt die Grundgleichungen des Hamilton Formalismus hergeleitet haben, sollen hier einige Eigenschaften der Hamiltonfunktion erläutert und an Beispielen veranschaulicht werden.

Zunächst soll gezeigt werden, dass die Hamiltonfunktion eine Konstante der Bewegung ist, wenn die Lagrangefunktion (und damit auch die Hamiltonfunktion) nicht explizit von der Zeit abhängen. Dies bedeutet, dass der Zahlenwert der Hamiltonfunktion sich in diesem Fall während der zeitlichen Entwicklung des Systems nicht ändert, auch wenn natürlich die aktuellen Werte für die generalisierten Koordinaten und Impulse von der Zeit abhängen.

Zum Beweis dieser Behauptung betrachten wir die totale Ableitung nach der Zeit

$$\frac{dH(p_i, q_i, t)}{dt} = \sum_{i=1}^n \left( \frac{\partial H}{\partial p_i} \dot{p}_i + \frac{\partial H}{\partial q_i} \dot{q}_i \right) + \frac{\partial H}{\partial t} .$$

Berücksichtigen wir nun, dass die partielle Ableitung von  $H$  nach der Zeit gleich dem negativen der partiellen Ableitung der Lagrangefunktion ist (2.8) und damit nach Voraussetzung identisch Null ist, und beachten ausserdem die Hamiltonschen Bewegungsgleichungen (2.9), so ergibt sich

$$\frac{dH(p_i, q_i, t)}{dt} = \sum_{i=1}^n \left[ \frac{\partial H}{\partial p_i} \left( -\frac{\partial H}{\partial q_i} \right) + \frac{\partial H}{\partial q_i} \frac{\partial H}{\partial p_i} \right] = 0 , \quad (2.15)$$

was wir ja beweisen wollten.

Sind die Zwangsbedingungen nicht zeitabhängig und ist das Potenzial unabhängig von den Geschwindigkeiten, so haben wir ja in (1.117) gezeigt, dass

$$\sum_j p_j \dot{q}_j = 2T ,$$

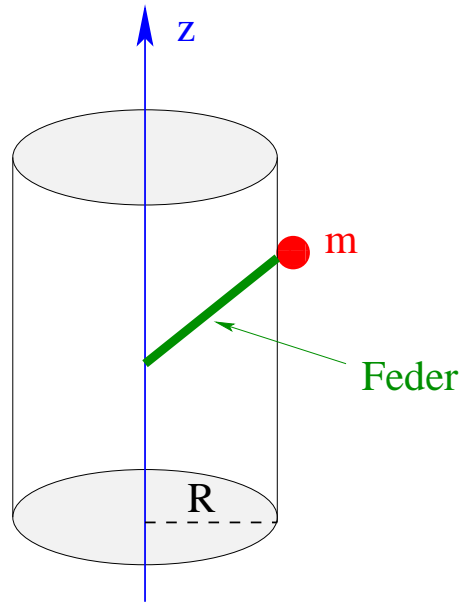
also das doppelte der kinetischen Energie  $T$  ist. In diesem Fall ist der Wert der Hamiltonfunktion

$$H = \sum_j p_j \dot{q}_j - L = 2T - (T - V) = T + V , \quad (2.16)$$

gleich der Energie des Systems. Man muss also in diesem Fall die Hamiltonfunktion nicht über die Definitionsgleichung (2.6) bestimmen sondern man kann sie direkt als Summe von kinetischer Energie plus potenzieller Energie, dargestellt als Funktion der generalisierten Koordinaten  $q_i$  und der kanonischen Impulse,  $p_i$ , berechnen.

Als ein erstes Anwendungsbeispiel betrachten wir die Bewegung eines Massenpunktes  $m$ , die auf dem Mantel eines Zylinders (Radius  $R$ , Zylinderachse parallel zur  $z$ -Achse, siehe Abb. 2.1) beschränkt sein soll. Ausserdem sei dieser Massenpunkt über eine Feder mit der Federkonstanten  $k$  mit dem Koordinatenursprung (auf der Zylinderachse) verbunden. Zur Behandlung dieses Problems bieten sich die Zylinderkoordinaten  $(r, \varphi, z)$  an. Die kinetische Energie für ein Teilchen der Masse  $m$  schreibt sich in diesen Zylinderkoordinaten

$$T = \frac{m}{2} \left( \dot{r}^2 + r^2 \dot{\varphi}^2 + \dot{z}^2 \right) .$$

Abbildung 2.1: Massenpunkt  $m$  auf dem Mantel eines Zylinders

Ist die Bewegung der Masse  $m$  auf den Mantel des Zylinders mit Radius  $r = R$  beschränkt, reduziert sich dieser Ausdruck auf ( $\dot{r} = 0$ , siehe auch (1.106))

$$T = \frac{m}{2} (R^2 \dot{\varphi}^2 + \dot{z}^2) .$$

Mit der potenziellen Energie  $V = k/2(R^2 + z^2)$  für die Kraft der Feder ergibt sich also die Lagrangefunktion

$$L = \frac{m}{2} (R^2 \dot{\varphi}^2 + \dot{z}^2) - \frac{k}{2} (R^2 + z^2) . \quad (2.17)$$

Da die Lagrangefunktion nicht von der generalisierten Koordinate  $\varphi$  abhängt ( $\varphi$  ist also zyklisch) ist der zugehörige Impuls

$$p_\varphi = \frac{\partial L}{\partial \dot{\varphi}} = mR^2 \dot{\varphi} , \quad (2.18)$$

eine Konstante der Bewegung. Damit ist also auch die Winkelgeschwindigkeit  $\dot{\varphi}$  eine Konstante der Bewegung und durch die Anfangsbedingungen  $\dot{\varphi} = \dot{\varphi}_0$  festgelegt. Bestimmen wir ausserdem den Winkel zur Startzeit  $t = 0$  auf  $\varphi_0$ , so ergibt sich sofort

$$\varphi(t) = \varphi_0 + t\dot{\varphi}_0 ,$$

die Winkelfunktion ist also eindeutig für alle Zeiten festgelegt.

Für die zweite generalisierte Koordinate  $z$  des Systems erhalten wir die Lagrange Bewegungsgleichung

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{z}} = m\ddot{z} = \frac{\partial L}{\partial z} = -kz ,$$

und daraus

$$\ddot{z} = -\omega^2 z \quad \text{mit} \quad \omega = \sqrt{\frac{k}{m}} . \quad (2.19)$$

Die allgemeine Lösung dieser Bewegungsgleichung lässt sich darstellen in der Form

$$z(t) = A \cos(\omega t) + B \sin(\omega t),$$

wobei die Konstanten  $A$  und  $B$  durch die Startbedingungen festzulegen sind. Wir haben hier die Lagrange Bewegungsgleichungen bestimmt und gelöst.

Als Alternative dazu sollen jetzt die entsprechenden Hamiltonschen Bewegungsgleichungen bestimmt werden. Der kanonische Impuls für die generalisierte Koordinate  $\varphi$  ist bereits in (2.18) definiert und die entsprechende Gleichung für die Koordinate  $z$  liefert

$$p_z = \frac{\partial L}{\partial \dot{z}} = m\dot{z}. \quad (2.20)$$

Damit können wir die Hamiltonfunktion gemäß der Definition (2.6) berechnen

$$\begin{aligned} H &= p_\varphi \dot{\varphi} + p_z \dot{z} - L \\ &= mR^2 \dot{\varphi}^2 + m\dot{z}^2 - \frac{m}{2} (R^2 \dot{\varphi}^2 + \dot{z}^2) + \frac{k}{2} (R^2 + z^2) \\ &= \frac{m}{2} (R^2 \dot{\varphi}^2 + \dot{z}^2) + \frac{k}{2} (R^2 + z^2) \\ &= T + V \end{aligned}$$

Wir haben also bis zu diesem Punkt nur bestätigt, dass die Hamiltonfunktion auch in diesem Fall die Gesamtenergie angibt. Dies ist aber keine Überraschung, da auch bei diesem Beispiel die Zwangsbedingungen zeitunabhängig sind und das Potenzial nicht von der Geschwindigkeit abhängt, so dass die Voraussetzungen für (2.16) gegeben sind.

Für die Anwendung der Hamiltonschen Bewegungsgleichung ist es nun erforderlich, die Hamiltonfunktion als Funktion der Koordinaten und Impulse zu bestimmen. Dies geschieht durch

$$H = \frac{p_\varphi^2}{2mR^2} + \frac{p_z^2}{2m} + \frac{k}{2} (R^2 + z^2). \quad (2.21)$$

Da  $\varphi$  eine zyklische Koordinate ist, ist, wie bereits diskutiert,  $p_\varphi$  eine Konstante der Bewegung und somit ist auch der erste Summand, genau so wie der dritte, lediglich eine Konstante. Wir betrachten die Hamiltonfunktion also lediglich als Funktion der Koordinate  $z$  und des zugehörigen Impulses  $p_z$ . Dies führt zu den Hamiltonschen Bewegungsgleichungen

$$\begin{aligned} \dot{p}_z &= -\frac{\partial H}{\partial z} = -kz \\ \dot{z} &= \frac{\partial H}{\partial p_z} = \frac{p_z}{m} \end{aligned}$$

Diese zwei Differenzialgleichungen erster Ordnung entsprechen natürlich der Lagrange Bewegungsgleichung (2.20), was man am einfachsten dadurch verifiziert, dass man die zweite dieser Gleichungen noch einmal nach der Zeit ableitet:

$$\ddot{z} = \frac{\dot{p}_z}{m} = -\frac{k}{m}z.$$

Dementsprechend ist natürlich auch die Lösung der Hamiltonschen Bewegungsgleichungen identisch mit der der Lagrange Gleichungen.

Als ein zweites Beispiel betrachten wir die Bewegung einer Masse  $m$  in einem konservativen zentralen Kraftfeld, das durch das Zentralpotenzial  $V(r)$  mit  $r$  dem Abstand vom Kraftzentrum und Koordinatenursprung. In diesem Fall bieten sich die Kugelkoordinaten  $(r, \vartheta, \varphi)$  als geeignete generalisierte Koordinaten an, und man erhält für die Lagrange-funktion

$$L = \frac{m}{2} (\dot{r}^2 + r^2 \dot{\vartheta}^2 + r^2 \sin^2 \vartheta \dot{\varphi}^2) - V(r). \quad (2.22)$$

Bei der Bewegung in einem Zentralfeld bleibt der Drehimpuls erhalten, was auch bedeutet, dass die Bewegung in einer Ebene erfolgt. Das Koordinatensystem kann also so festgelegt werden, dass  $\varphi = 0$ , dafür aber  $\vartheta$  Werte zwischen 0 und  $2\pi$  annehmen kann. Damit entfällt der dritte Term in (2.22) und der Winkel  $\vartheta$  wird zu einer zyklischen Koordinate. Daraus folgt wiederum, dass der zugehörige Impuls

$$p_{\vartheta} = \frac{\partial L}{\partial \dot{\vartheta}} = mr^2 \dot{\vartheta} = l, \quad (2.23)$$

eine Konstante der Bewegung ist. Diese Konstante entspricht dem Betrag des Drehimpulses des Teilchens  $l$ . Mit dem kanonischen Impuls

$$p_r = \frac{\partial L}{\partial \dot{r}} = m\dot{r} \quad (2.24)$$

ergibt sich dann die Hamiltonfunktion zu

$$H = T + V = \frac{p_r^2}{2m} + \frac{l^2}{2mr^2} + V(r) \quad (2.25)$$

und die kanonischen Bewegungsgleichungen schreiben sich

$$\begin{aligned} \dot{r} &= \frac{\partial H}{\partial p_r} = \frac{p_r}{m} \\ \dot{p}_r &= -\frac{\partial H}{\partial r} = -\frac{\partial}{\partial r} \underbrace{\left( V(r) + \frac{l^2}{2mr^2} \right)}_{:=V_{\text{effektiv}}(r)}. \end{aligned} \quad (2.26)$$

Diese Bewegungsgleichungen für  $r(t)$  haben also mathematisch die gleiche Gestalt, wie für die Bewegung der Masse  $m$  in einem eindimensionalen Problem mit einem effektiven Potenzial

$$V_{\text{effektiv}}(r) = V(r) + \frac{l^2}{2mr^2}. \quad (2.27)$$

Für ein attraktives Potenzial  $V(r)$  sind in Abb. 2.2 dieses Potenzial und ein zugehöriges effektives Potenzial skizziert. Für  $l \neq 0$  liefert der Zusatzterm einen repulsiven Beitrag, der für  $r \rightarrow 0$  divergiert. Dieser sogenannte **Zentrifugalterm** hat seinen Ursprung in der kinetischen Energie, wie man aus der obigen Herleitung leicht sieht. Er bringt zum Ausdruck, dass man bei festem Wert des Drehimpulses  $l$  hohe Winkelgeschwindigkeiten und damit einen großen Beitrag zur kinetischen Energie erbringen muss, wenn der Abstand des Teilchens  $r$  vom Zentrum klein wird.

Die Form des effektiven Potenzials und unsere Erfahrungen mit mechanischen Problemen in einer Dimension lassen uns erahnen, wie die Lösungen für  $r(t)$  aussehen werden. Bringt

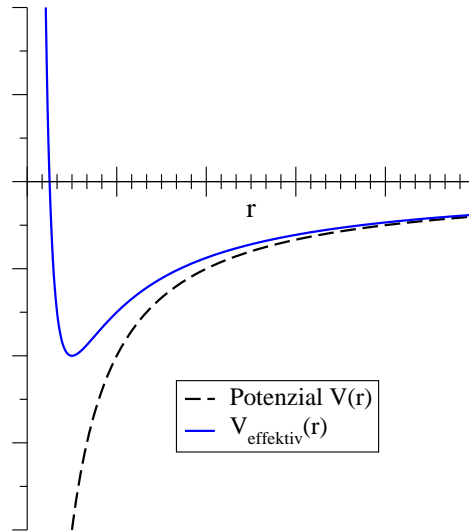


Abbildung 2.2: Potenzial plus Zentrifugalterm in einem Zentralfeldproblem (siehe (2.27)).

man das Teilchen auf eine Umlaufbahn, bei der der Abstand  $r$  gerade der Position des Minimums von  $V_{\text{effektiv}}$  entspricht und ist  $\dot{r} = 0$ , so erwarten wir eine stabile Bahn mit festem Wert für  $r$ . Wird zu einer Zeit  $t_0$  das System gestartet mit einem Wert von  $r$ , der von diesem Minimalwert abweicht, so erwarten wir eine Schwingung der Funktion  $r(t)$  um dieses Minimum. Die Frequenz dieser Schwingung könnte man durch eine Taylor-Entwicklung des effektiven Potenzials um das Minimum leicht abschätzen.

Diese Schwingung für  $r(t)$  beschreibt natürlich nur den Zeitverlauf der radialen Koordinate. Damit sind aber über (2.23) auch die aktuellen Winkelgeschwindigkeiten  $\dot{\vartheta}(t)$  bestimmt ( $l$  bleibt ja konstant), was uns auch erlaubt, den aktuellen Positionswinkel  $\vartheta(t)$  zu berechnen.

Ein weiteres Beispiel von großer Bedeutung ist die Behandlung eines Teilchens mit der Masse  $m$  und der Ladung  $e$  in einem elektromagnetischen Feld. Im Abschnitt 1.3 haben wir gezeigt, dass die elektromagnetischen Kräfte auf eine Ladung durch ein geschwindigkeitsabhängiges Potenzial beschrieben wird, (siehe (1.73)), so dass sich eine Lagrangefunktion ergibt

$$L = \frac{m}{2} \dot{\vec{r}}^2 - e\Phi(\vec{r}) + e\vec{A}(\vec{r}) \cdot \dot{\vec{r}}. \quad (2.28)$$

Dabei entsprechen z.B. die kartesischen Koordinaten des Teilchens den generalisierten Koordinaten ( $r_i = q_i$ ,  $i = x, y, z$ ) und entsprechendes gilt für die Geschwindigkeiten. Die Felder  $\Phi$  und  $\vec{A}$  sind die üblichen Potentiale zur Bestimmung der elektromagnetischen Felder. Mit dieser Lagrangefunktion ergeben sich die kanonischen Impulse

$$p_i = \frac{\partial L}{\partial \dot{r}_i} = m\dot{r}_i + eA_i, \quad (2.29)$$

wobei,  $A_i$  die entsprechende kartesische Komponente des Vektorfeldes  $\vec{A}$  bezeichnet. Damit ergibt sich die Hamiltonfunktion

$$H = \vec{p} \cdot \dot{\vec{r}} - L$$

$$\begin{aligned}
&= m\dot{r}^2 + e\vec{A} \cdot \dot{\vec{r}} - \left( \frac{m}{2}\dot{r}^2 - e\Phi + e\vec{A} \cdot \dot{\vec{r}} \right) \\
&= \frac{m}{2}\dot{r}^2 + e\Phi.
\end{aligned} \tag{2.30}$$

Dies ist gerade die gesamte Energie als Summe der kinetischen Energie plus der potenziellen Energie der Ladung im “elektrostatischen” Potenzial. Dieses Ergebnis war nicht zu erwarten, da wir es ja hier mit einem geschwindigkeitsabhängigen Potenzial zu tun haben, so dass die Voraussetzungen für (2.16) nicht erfüllt sind.

Sind die elektromagnetischen Potenziale konstant, also unabhängig von der Zeit, so hängt auch  $H$  nicht explizit von der Zeit ab und ist somit eine Konstante der Bewegung (siehe(2.15)).

Zur Bestimmung der Bewegungsgleichung müssen wir den Ausdruck (2.30) umschreiben in eine Funktion der kanonischen Impulse. Mit (2.29) ergibt dies

$$H = \frac{(\vec{p} - e\vec{A})^2}{2m} + e\Phi. \tag{2.31}$$

Man kann dieses Ergebnis auch in dem folgenden Rezept für die **minimale Substitution** zusammenfassen: Ausgehend von der Hamiltonfunktion des freien Teilchens

$$H = \frac{\vec{p}^2}{2m}$$

erhält man die Hamiltonfunktion des geladenen Teilchens im elektromagnetischen Feld durch folgende Ersetzungen:

$$\begin{aligned}
H &\Rightarrow H - e\Phi \\
\vec{p} &\Rightarrow \vec{p} - e\vec{A}.
\end{aligned} \tag{2.32}$$

In einem weiteren Schritt wollen wir uns nun die Hamiltonfunktion zur Beschreibung der relativistischen Kinematik ansehen. In einem ersten Teil wollen wir uns davon überzeugen, dass die Kinematik eines freien Teilchens, auf das also keine Kräfte wirken, durch die Lagrangefunktion

$$L = -mc^2 \sqrt{1 - \frac{\dot{r}^2}{c^2}}, \tag{2.33}$$

beschrieben wird. Dabei steht  $c$  für die Lichtgeschwindigkeit und  $\dot{\vec{r}}$  ist die Geschwindigkeit des Teilchens im aktuellen Koordinatensystem. Diese Lagrangefunktion sieht auf den ersten Blick etwas merkwürdig aus, insbesondere da ja nach unseren bisherigen Regeln die Lagrangefunktion für ein freies Teilchen durch die kinetische Energie gegeben sein soll, der Ausdruck (2.33) aber negative Werte liefert. Mit den folgenden Argumenten wollen wir uns aber davon überzeugen, dass (2.33) eine richtige Wahl für die Lagrangefunktion darstellt.

- Als erstes betrachten wir den nichtrelativistischen Grenzfall dieser Lagrangefunktion, also den Fall für den der Quotient  $\dot{r}^2/c^2$  klein ist im Vergleich zu 1. Dann können

wir die Taylorentwicklung der Wurzel in (2.33) betrachten

$$\begin{aligned} L &= -mc^2 \left[ 1 - \frac{1}{2} \frac{\dot{r}^2}{c^2} - \frac{1}{8} \left( \frac{\dot{r}^2}{c^2} \right)^2 + \dots \right] \\ &= -mc^2 + \frac{1}{2} m \dot{r}^2 + \dots \end{aligned} \quad (2.34)$$

In diesem Grenzfall liefert  $L$  also gerade den nichtrelativistischen Ausdruck für die kinetische Energie minus einer Konstanten ( $mc^2$ ), die ja für die Bewegungsgleichung oder bei der Formulierung des Hamiltonschen Prinzips keine Rolle spielt.

- Nach dem Hamiltonschen Prinzip legt das System ja den Weg zurück, für den die Wirkung

$$S = \int_{t_1}^{t_2} L dt, \quad (2.35)$$

ein Extremum ist. Dieses Variationsverfahren sollte natürlich in jedem Koordinatensystem das gleiche Ergebnis liefern. Der Wert des Integrals, beziehungsweise der Integrand sollte also invariant unter einer Lorentztransformation sein. Nun hängt aber der jeweilige Wert der Lagrangefunktion  $L$  in (2.33) von der Geschwindigkeit des Teilchens ab und damit natürlich auch vom Koordinatensystem. Andererseits wissen wir aber natürlich, dass sich auch die Zeit  $t$  und damit auch die Differentialform  $dt$  bei einer Lorentztransformation ändert. Die Eigenzeit, also die Zeit, die im Koordinatensystem gemessen wird, das sich mit dem Teilchen mitbewegt ist eine Lorentzskalar also invariant unter Lorentztransformation. Diese Eigenzeit  $\tau$  ergibt sich aus der Zeit, die in einem Koordinatensystem gemessen wird, das sich mit der Geschwindigkeit  $|\dot{\vec{r}}|$  relativ zum Eigensystem bewegt durch

$$\tau = t \sqrt{1 - \frac{\dot{r}^2}{c^2}},$$

was ja für die zugehörigen Differentialformen bedeutet, dass

$$d\tau = \sqrt{1 - \frac{\dot{r}^2}{c^2}} dt. \quad (2.36)$$

Damit können wir den Integranden im Wirkungsintegral (2.35) umschreiben

$$L dt = -mc^2 \sqrt{1 - \frac{\dot{r}^2}{c^2}} dt = -mc^2 d\tau.$$

Dieser Ausdruck ist also das Produkt zweier Lorentzskalare: der Differentialform der Eigenzeit  $d\tau$  und der invarianten Zahl  $-mc^2$ , die wir ja als Lagrangefunktion in dem System interpretieren können, in dem sich das Teilchen nicht bewegt und in dem dann auch die Eigenzeit gemessen wird.



- Schliesslich liefert, und das ist ja eigentlich das entscheidende Merkmal, die Lagrangefunktion (2.33) aber auch die richtige Bewegungsgleichung. So gilt ja

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{r}_i} = \frac{d}{dt} \frac{m \dot{r}_i}{\sqrt{1 - \frac{\dot{r}^2}{c^2}}} = \frac{d}{dt} p_i = F_i = \frac{\partial L}{\partial r_i}. \quad (2.37)$$

Die Zeitableitung der Komponenten des relativistischen Impulses

$$p_i = \frac{m \dot{r}_i}{\sqrt{1 - \frac{\dot{r}^2}{c^2}}}, \quad (2.38)$$

entspricht der entsprechenden Komponente der Kraft  $F_i$ , was ja gerade die Aussage der Newtonschen Bewegungsgleichung ist.

Mit den Komponenten des kanonischen Impulses  $p_i$  in (2.38) können wir nun die Hamiltonfunktion bestimmen:

$$\begin{aligned} H &= \sum_i p_i \dot{r}_i - L \\ &= \frac{m \dot{r}^2}{\sqrt{1 - \frac{\dot{r}^2}{c^2}}} + mc^2 \sqrt{1 - \frac{\dot{r}^2}{c^2}} \\ &= \frac{mc^2}{\sqrt{1 - \frac{\dot{r}^2}{c^2}}} = E. \end{aligned} \quad (2.39)$$

Nach (2.16) ist dies also die Energie des freien Teilchens. Wir können uns leicht davon überzeugen, dass im nichtrelativistischen Grenzfall dieser Ausdruck neben der Ruheenergie des Massenpunktes

$$E = mc^2$$

als nächsten Term die nichtrelativistische kinetische Energie enthält.

Für die weiteren Betrachtungen wollen wir die Hamiltonfunktion beziehungsweise den Ausdruck für die Energie als Funktion des kanonischen Impulses darstellen. Dazu quadrieren wir die letzte Zeile von (2.39)

$$\begin{aligned} E^2 &= \frac{m^2 c^4}{1 - \frac{\dot{r}^2}{c^2}} \\ &= \frac{m^2 c^4 - m^2 c^2 \dot{r}^2}{1 - \frac{\dot{r}^2}{c^2}} + \frac{m^2 c^2 \dot{r}^2}{1 - \frac{\dot{r}^2}{c^2}} \\ &= m^2 c^4 + c^2 \vec{p}^2. \end{aligned} \quad (2.40)$$

Dies bedeutet also, dass die Hamiltonfunktion für ein relativistisches freies Teilchen lautet

$$E = H(p) = \sqrt{m^2 c^4 + c^2 \vec{p}^2}. \quad (2.41)$$

Wir können diese Energie - Impulsbeziehung aber auch umschreiben auf die Form

$$m^2 c^2 = \frac{E^2}{c^2} - \vec{p}^2 = p_0^2 - \sum_i p_i^2 = p^\mu p_\mu, \quad (2.42)$$

mit dem kontravarianten Vektor

$$p^\mu = \begin{pmatrix} E/c \\ \vec{p} \end{pmatrix}. \quad (2.43)$$

Der dreidimensionale Impulsvektor  $\vec{p}$  wird also ergänzt durch die zeitartige (nullte) Komponente  $p^0 = E/c$ , so dass das Betragsquadrat dieses Vierektors den Lorentzskalar  $m^2 c^2$  liefert (siehe (2.42)).

Mit diesen Bezeichnungen bekommt auch die minimale Substitution (2.32) eine neue Bedeutung. Wir wissen aus der Elektrodynamik, dass die Potenzialfelder einen Lorentzvektor bilden

$$A^\mu = \begin{pmatrix} \Phi/c \\ \vec{A} \end{pmatrix}.$$

Damit schreibt sich nun die Regel der minimalen Substitution

$$p^\mu \Rightarrow p^\mu - eA^\mu \quad (2.44)$$

Die Wirkung der elektromagnetischen Felder wird dadurch berücksichtigt, dass der kontravariante Impulsvektor des freien Teilchen ergänzt wird durch die Addition des kontravarianten Vektors der Potenzialfelder multipliziert mit der negativen Ladung des Teilchens ( $-e$ ), was natürlich wieder zu einem kontravarianten Vektor führt. Durch diese einfache Addition der Vektoren wird also auf einfachste (minimalste) Art gewährleistet, dass die Berücksichtigung der elektromagnetischen Felder die Kovarianz der Bewegungsgleichungen unter Lorentztransformationen nicht zerstört.

Diese minimale Substitution führt zu einer relativistischen Hamiltonfunktion eines geladenen Teilchens im elektromagnetischen Feld der Form

$$H = \sqrt{m^2 c^4 + c^2 (\vec{p} - e\vec{A})^2} + e\Phi. \quad (2.45)$$

## 2.3 Kanonische Transformationen

Ein wichtiger Unterschied zwischen dem Hamilton Formalismus und dem Lagrange Formalismus besteht darin, dass wir verschiedene Arten von Koordinaten benutzen um eine Position des Systems im Phasenraum zu bezeichnen. Im Lagrange Formalismus kennzeichnet man einen Punkt im Phasenraum, also z.B. den Startpunkt eines Systems, durch die Angabe von  $n$  generalisierten Koordinaten  $q_i(t)$  und den zugehörigen Geschwindigkeiten  $\dot{q}_i(t)$ . Wenn wir die Beschreibung vereinfachen wollen, also z.B. das System der gekoppelten Bewegungsgleichungen vereinfachen wollen, so kann man versuchen eine geeignetere Basis zu finden. Das bedeutet man versucht andere generalisierte Koordinaten  $Q_i$  zu finden, in denen sich die Bewegungsgleichungen vereinfachen. Eine solche Transformation der Koordinaten

$$q_i \implies Q_i = Q_i(q_1, \dots, q_n), \quad (2.46)$$

legt auch die Transformation der zugehörigen Geschwindigkeiten  $\dot{q}_i \implies \dot{Q}_i$  fest. Die Freiheit eine geeignete Basis des  $2n$  dimensionalen Phasenraumes zu finden ist also auf eine Transformation des  $n$  dimensionalen Raumes der Koordinaten beschränkt. Man bezeichnet Transformationen vom Typ (2.46) deshalb als **Punkttransformationen**.

Im Hamilton Formalismus wird der Phasenraum durch  $n$  generalisierte Koordinaten und  $n$  generalisierte Impulse definiert. Es besteht dabei keine so eindeutige Verknüpfungen zwischen den  $q_i$  und  $p_i$ , wie das bei den Koordinaten und Geschwindigkeiten des Lagrange Formalismus der Fall. Es erhebt sich deshalb die Frage: Gibt es im Hamilton Formalismus allgemeinere Transformationen als die Punkttransformationen des Lagrange Formalismus? Gibt es Transformationen auf "neue" Koordinaten  $Q_i$  und Impulse  $P_i$ , die über die Punkttransformationen hinausgehen also vom Typ

$$\begin{aligned} q_i & \implies Q_i = Q_i(q_1 \dots q_n, p_1 \dots p_n) \\ p_i & \implies P_i = P_i(q_1 \dots q_n, p_1 \dots p_n) \end{aligned}, \quad (2.47)$$

sind, aber dennoch zu Hamiltonschen Bewegungsgleichungen führen. Wir nennen eine solche Transformation eine **Kanonische Transformation** wenn die Hamiltonschen Bewegungsgleichungen sowohl für die ursprüngliche Hamiltonfunktion  $H$  mit ihren Satz von Koordinaten und Impulsen

$$H(p_i, q_i, t) \quad \text{mit} \quad \begin{aligned} \dot{q}_i &= \frac{\partial H}{\partial p_i} \\ \dot{p}_i &= -\frac{\partial H}{\partial q_i} \end{aligned}, \quad (2.48)$$

als auch für die "neuen" Koordinaten, Impulse und daraus gebildete Hamiltonfunktion  $\mathcal{H}$

$$\mathcal{H}(P_i, Q_i, t) \quad \text{mit} \quad \begin{aligned} \dot{Q}_i &= \frac{\partial \mathcal{H}}{\partial P_i} \\ \dot{P}_i &= -\frac{\partial \mathcal{H}}{\partial Q_i} \end{aligned}, \quad (2.49)$$

gelten.

Der Hauptsatz über die Kanonischen Transformationen besagt nun, dass eine Kanonische Transformation genau dann existiert, wenn es für diese Transformation eine **Erzeugende Funktion**  $F_1(Q_i, q_i, t)$  gibt in der Art, dass die Transformation (2.47) festgelegt ist durch die Beziehungen

$$P_i = -\frac{\partial F_1}{\partial Q_i}$$

$$\begin{aligned} p_i &= \frac{\partial F_1}{\partial q_i} \\ \mathcal{H} &= H + \frac{\partial F_1}{\partial t}. \end{aligned} \quad (2.50)$$

Bevor wir an den Beweis dieser Behauptung herangehen, sollen diese Begriffe an einem einfachen Beispiel verdeutlicht werden. Dazu betrachten wir die Hamiltonfunktion des harmonischen Oszillators in einer Dimension

$$H(p, q) = \frac{p^2}{2m} + \frac{k}{2}q^2, \quad (2.51)$$

mit den Bewegungsgleichungen

$$\dot{p} = -\frac{\partial H}{\partial q} = -kq \quad \text{und} \quad \dot{q} = \frac{\partial H}{\partial p} = \frac{p}{m}. \quad (2.52)$$

Die Transformation soll erzeugt werden durch die Funktion

$$F_1(Q, q) = qQ.$$

Mit (2.50) erhalten wir also die Transformationsgleichungen

$$P = -\frac{\partial F_1}{\partial Q} = -q \quad \text{sowie} \quad p = \frac{\partial F_1}{\partial q} = Q \quad \text{und} \quad \mathcal{H} = H. \quad (2.53)$$

Die Koordinaten und Impulse haben bei dieser Transformation also gerade ihre Bedeutung vertauscht. Mit (2.53) ergibt sich die Hamiltonfunktion in den "neuen" Variablen

$$\mathcal{H} = \frac{Q^2}{2m} + \frac{k}{2}P^2,$$

mit den Bewegungsgleichungen

$$\dot{P} = -\frac{\partial \mathcal{H}}{\partial Q} = -\frac{Q}{m} \quad \text{und} \quad \dot{Q} = \frac{\partial \mathcal{H}}{\partial P} = kP.$$

Übersetzt man diese Gleichungen mit den Transformationen (2.53)

$$\dot{P} = -\dot{q} = -\frac{Q}{m} = -\frac{p}{m} \quad \text{und} \quad \dot{Q} = \dot{p} = kP = -kq,$$

so erhält man zu (2.52) äquivalente Gleichungen. In beiden Koordinatensystemen wird also die gleiche Bewegung beschrieben, wie wir es ja auch von einer kanonischen Transformation erwarten.

Zum Beweis des Hauptsatzes über die Kanonischen Transformationen erinnern wir uns zunächst einmal daran, dass die Hamiltonschen Bewegungsgleichungen aus dem modifizierten Hamiltonschen Prinzip (2.10) folgen. Die Forderung nach einer Kanonischen Transformation ist also gleichbedeutend damit, dass aus

$$\delta \int_{t_1}^{t_2} dt \left[ \sum_{i=1}^n p_i \dot{q}_i - H(p_i, q_i, t) \right] = 0$$

folgt, dass

$$\delta \int_{t_1}^{t_2} dt \left[ \sum_{i=1}^n P_i \dot{Q}_i - \mathcal{H}(P_i, Q_i, t) \right] = 0$$

folgt und umgekehrt, aus der zweiten Gleichung die erste. Die beiden Integrale sollten sich allenfalls um eine Konstante unterscheiden, die unabhängig vom Weg ist. Dies erreicht man genau dann, wenn sich die beiden Integranden nur um die totale Ableitung einer Funktion  $F_1(Q_i, q_i, t)$  nach der Zeit unterscheiden, d.h.

$$\sum_{i=1}^n p_i \dot{q}_i - H(p_i, q_i, t) = \sum_{i=1}^n P_i \dot{Q}_i - \mathcal{H}(P_i, Q_i, t) + \frac{dF_1(Q_i, q_i, t)}{dt}. \quad (2.54)$$

Der letzte Term integriert über die Zeit liefert ja

$$\int_{t_1}^{t_2} \frac{dF_1(Q_i, q_i, t)}{dt} dt = F_1(Q_i(t_2), q_i(t_2), t_2) - F_1(Q_i(t_1), q_i(t_1), t_1)$$

also eine Zahl die unabhängig ist vom Weg, da ja die Endpunkte  $Q_i(t_{1/2})$  bzw.  $q_i(t_{1/2})$  bei der Variation des Weges festgehalten werden sollen. Damit entspricht die Forderung nach einer kanonischen Transformation, der Gültigkeit von (2.54). Wir werden uns nun davon überzeugen, dass aus der Gültigkeit von (2.54) die Transformationsregeln (2.50) folgen. Dazu schreiben wir

$$\frac{dF_1(Q_i, q_i, t)}{dt} = \sum_{i=1}^n \frac{\partial F_1}{\partial q_i} \dot{q}_i + \frac{\partial F_1}{\partial Q_i} \dot{Q}_i + \frac{\partial F_1}{\partial t}.$$

Setzt man dies in (2.54), ergibt sich nach kleinen Umformungen

$$\sum_{i=1}^n \dot{q}_i \left( p_i - \frac{\partial F_1}{\partial q_i} \right) - H(p_i, q_i, t) = \sum_{i=1}^n \dot{Q}_i \left( P_i + \frac{\partial F_1}{\partial Q_i} \right) - \mathcal{H}(P_i, Q_i, t) + \frac{\partial F_1}{\partial t}. \quad (2.55)$$

Da die alten Koordinaten und die neuen hier als voneinander unabhängig angesehen werden, wird diese Gleichung nur dann identisch erfüllt, wenn

$$\left( p_i - \frac{\partial F_1}{\partial q_i} \right) = 0 \quad \text{und} \quad \left( P_i + \frac{\partial F_1}{\partial Q_i} \right) = 0. \quad (2.56)$$

Es bleibt dann nur noch von (2.55)

$$-H(p_i, q_i, t) = -\mathcal{H}(P_i, Q_i, t) + \frac{\partial F_1}{\partial t}. \quad (2.57)$$

Die Gleichungen (2.56) und (2.57) entsprechen aber genau den Transformationsgleichungen (2.50). Wir haben also aus der Forderung nach einer Kanonischen Transformation (bzw. aus Gl.(2.54)) die Existenz der erzeugenden Funktion  $F_1$  mit den Eigenschaften (2.50) hergeleitet. Man kann den Weg aber auch umkehren: Aus den Transformationsgleichungen (2.56) und (2.57) folgt (2.55) und damit (2.54) also schliesslich die Behauptung, dass die Transformation kanonisch ist.

Wir haben hier eine erzeugende Funktion  $F_1$  als Funktion von alten und neuen Koordinaten betrachtet. Man mag aber eventuell vorziehen, dass man an Stelle von  $Q_i$  und  $q_i$  als

unabhängige Koordinaten lieber die alten Koordinaten und neuen Impulse als unabhängig betrachten möchte. Dazu generieren wir eine erzeugende Funktion der Form

$$F_2(q_i, P_i, t) = \sum_{i=1}^n P_i Q_i + F_1(q_i, Q_i, t). \quad (2.58)$$

Also eine Umformung ähnlich wie bei der Legendre Transformation von  $L$  nach  $H$ . Wir verifizieren, dass  $F_2$  in der Tat von den unabhängigen Variablen  $q_i$  und  $P_i$  abhängt und betrachten dazu die Differenzialform

$$dF_2 = \sum_{i=1}^n \left[ P_i dQ_i + Q_i dP_i + \underbrace{\frac{\partial F_1}{\partial q_i}}_{=p_i} dq_i + \underbrace{\frac{\partial F_1}{\partial Q_i}}_{=-P_i} dQ_i \right] + \frac{\partial F_1}{\partial t} dt, \quad (2.59)$$

wobei die Beziehungen von (2.50) eingesetzt wurden. Da sich der erste und der vierte Term in der Summe kompensieren, treten nur Terme auf, die proportional zu  $dP_i$ ,  $dq_i$  und  $dt$  sind. Dies bedeutet aber, dass  $F_2$  eine Funktion der unabhängigen Variablen  $P_i$ ,  $q_i$  und  $t$  ist mit der Differenzialform

$$dF_2 = \sum_{i=1}^n \left[ \frac{\partial F_2}{\partial P_i} dP_i + \frac{\partial F_2}{\partial q_i} dq_i \right] + \frac{\partial F_2}{\partial t} dt, \quad (2.60)$$

ist. Der Vergleich von (2.59) und (2.60) zeigt, dass

$$Q_i = \frac{\partial F_2}{\partial P_i} \quad \text{und} \quad p_i = \frac{\partial F_2}{\partial q_i} \quad \text{sowie} \quad \frac{\partial F_2}{\partial t} = \frac{\partial F_1}{\partial t} = \mathcal{H} - H. \quad (2.61)$$

Alternativ kann man auch erzeugende Funktionen  $F_3(Q_i, p_i, t)$  und  $F_4(P_i, p_i, t)$  betrachten, die (hier ohne Beweis) Koordinatentransformationen der Form

$$P_i = -\frac{\partial F_3}{\partial Q_i} \quad \text{und} \quad q_i = -\frac{\partial F_3}{\partial p_i}, \quad (2.62)$$

$$Q_i = \frac{\partial F_4}{\partial P_i} \quad \text{und} \quad q_i = -\frac{\partial F_4}{\partial p_i}, \quad (2.63)$$

erzeugen. Beachte dabei insbesondere das Auftreten der Vorzeichen bei einzelnen dieser Gleichungen.

Wir wollen uns nun davon überzeugen, dass insbesondere die Punkttransformationen (2.46), die ja auch im Rahmen des Lagrange Formalismus möglich sind, einen Spezialfall einer Kanonischen Transformation darstellen. Wir betrachten dazu eine Punkttransformation in der Form, dass die

$$Q_i = f_i(q_j, t), \quad (2.64)$$

also als beliebige Funktionen  $f_i$  der  $q_j$  gegeben sind. Nehmen wir nun eine erzeugende Funktion vom Typ  $F_2(q_i, P_i, t)$  in der Form

$$F_2(q_j, P_j, t) = \sum_i f_i(q_j, t) P_i, \quad (2.65)$$

so ergeben die zugehörigen Transformationsgleichungen (2.61)

$$\begin{aligned} Q_i &= \frac{\partial F_2}{\partial P_i} = f_i(q_j, t) \\ p_i &= \frac{\partial F_2}{\partial q_i} = \sum_j P_j \frac{\partial f_j}{\partial q_i}. \end{aligned} \quad (2.66)$$

Die erste Gleichung liefert gerade die geforderte Punkttransformation. Damit die Transformation kanonisch ist, wird gleichzeitig durch die zweite Gleichung die Transformation der Impulskoordinaten festgelegt.

Als Spezialfall einer Punkttransformation wollen wir uns nun den Fall ansehen, dass die neuen Koordinaten  $Q_i$  durch eine orthogonale Transformation aus den  $q_j$  bestimmt werden. In diesem Fall ist also

$$Q_i = f_i(q_j) = \sum_j a_{ij} q_j \quad \text{in Matrixform:} \quad \vec{Q} = A \vec{q}.$$

Dabei gilt für eine orthogonale Transformation, dass für die Matrix  $A$ , die durch die Matrixelemente  $a_{ij}$  definiert ist, gilt,  $A^{-1} = A^t$ , die transponierte Matrix ist auch die inverse zur  $A$ . Die Funktion  $F_2$  hat dann die Gestalt

$$F_2 = \sum_{i,j} a_{ij} q_j P_i$$

Damit ergibt sich für die Transformation der Impulse

$$p_i = \sum_j a_{ji} P_j = \sum_j a_{ij}^t P_j.$$

In der Matrixschreibweise bedeutet dies:

$$\vec{p} = A^t \vec{P} = A^{-1} \vec{P} \iff \vec{P} = A \vec{p},$$

die Impulse unterliegen der gleichen Transformation wie die Koordinaten.

Wir kommen noch einmal auf das Beispiel des Harmonischen Oszillators zurück mit der Hamiltonfunktion in (2.51) und den Bewegungsgleichungen in (2.52). Aus diesen Bewegungsgleichung ergibt sich

$$\ddot{q} = -\omega^2 q \quad \text{mit} \quad \omega^2 = \frac{k}{m}$$

und daraus die allgemeine Lösung

$$q(t) = A \sin(\omega t + \varphi), \quad (2.67)$$

mit Konstanten  $A$  und  $\varphi$  für die Amplitude und Phase, die aus den Anfangsbedingungen zu bestimmen sind. An diesem Beispiel wollen wir demonstrieren, dass man im Prinzip durch die Konstruktion einer geeigneten Kanonischen Transformation, die Bewegungsgleichungen vereinfachen kann. Dies ist vielleicht für das hier gewählte Beispiel nicht wirklich erforderlich, aber das Beispiel soll ja nur dazu dienen, das Prinzip zu demonstrieren.

Wir wählen als erzeugende Funktion, eine Funktion vom Typ  $F_1$  in der Form

$$F_1(Q, q) = \frac{m\omega}{2} q^2 \cot(Q). \quad (2.68)$$

Mit (2.50) definiert diese Funktion die kanonische Transformation

$$\begin{aligned} p &= \frac{\partial F_1}{\partial q} = m\omega q \cot(Q) \\ P &= \frac{\partial F_1}{\partial Q} = \frac{m\omega q^2}{2} \frac{1}{\sin^2 Q}. \end{aligned} \quad (2.69)$$

Aus der zweiten dieser Gleichungen ergibt sich

$$q^2 = \frac{2P \sin^2 Q}{m\omega}. \quad (2.70)$$

Zusammen mit dem Quadrat der ersten Gleichung erhält man also

$$p^2 = m^2 \omega^2 q^2 \cot^2 Q = 2m\omega P \cos^2 Q. \quad (2.71)$$

Mit (2.70) und (2.71) haben wir 2 Gleichungen durch die die alten Impulse und Koordinaten durch die neuen  $P$  und  $Q$  dargestellt sind. Wir können deshalb die Hamiltonfunktion umschreiben

$$\mathcal{H} = H = \frac{p^2}{2m} + \frac{k}{2} q^2 = \omega P \cos^2 Q + k \frac{P \sin^2 Q}{m\omega} = \omega P. \quad (2.72)$$

Bei der letzten Umformung wurde benutzt, dass  $k/m = \omega^2$ . Damit haben wir eine sehr einfache Darstellung der Hamiltonfunktion  $\mathcal{H} = \omega P$ , die auch zu einfachen Bewegungsgleichung führt. Wegen

$$\dot{P} = -\frac{\partial \mathcal{H}}{\partial Q} = 0$$

ist der Impuls  $P$  eine Konstante der Bewegung, nämlich gerade

$$P = \frac{\mathcal{H}}{\omega} = \frac{E}{\omega}$$

also die Energie dividiert durch die Winkelgeschwindigkeit  $\omega$ . Auch die zweite Bewegungsgleichung ist sehr leicht integriert

$$\dot{Q} = \frac{\partial \mathcal{H}}{\partial P} = \omega \quad \Rightarrow \quad Q(t) = \omega t + \varphi$$

mit einer Konstanten  $\varphi$ , die noch zu bestimmen ist. Eingesetzt in die Gleichung (2.70) ergibt sich daraus für die ursprüngliche Koordinate

$$q(t) = \sqrt{\frac{2P}{m\omega}} \sin(\omega t + \varphi) = \sqrt{\frac{2E}{m\omega^2}} \sin(\omega t + \varphi),$$

also genau das Ergebnis (2.67).

Die Lösung der Differenzialgleichungen aus den Bewegungsgleichungen war in den neuen Variablen  $P$  und  $Q$  ein Kinderspiel. Das Problem ist aber: Wie findet man eine Kanonische



Transformation, die die Bewegungsgleichungen so vereinfacht? Wie kommt man auf eine erzeugende Funktion wie (2.68) in diesem Beispiel?

Im Prinzip bietet der **Hamilton Jacobi** Formalismus einen Weg, geeignete erzeugende Funktionen zu bestimmen. Im Fall, dass die Hamiltonfunktion nicht explizit von der Zeit abhängt, wird man versuchen eine erzeugende Funktion zu finden, bei der die Hamiltonfunktion, die ja in diesem Fall eine Konstante der Bewegung ist, genau einem generalisierten Impuls, sagen wir  $P_1$  entspricht, also

$$\mathcal{H} = P_1 \quad (2.73)$$

Damit sind alle Impulse  $P_i$  Konstanten und wie man aus den Bewegungsgleichungen sofort sieht ebenso die Koordinaten  $Q_2$  bis  $Q_n$ . Lediglich für  $Q_1$  erhalten wir

$$Q_1(t) = t + \varphi.$$

Die Lösung der Bewegungsgleichungen ist also trivial, wie bestimmen wir aber die erzeugende Funktion  $F_1$ ? Wir wollen dieses Verfahren am Beispiel des eindimensionalen Harmonischen Oszillators skizzieren. Gesucht ist also eine Funktion  $F_1(Q, q)$  mit

$$p = \frac{\partial F_1}{\partial q} \quad \text{und} \quad P = -\frac{\partial F_1}{\partial Q}. \quad (2.74)$$

Ausserdem soll aber auch gelten

$$\mathcal{H} = P = H = \frac{1}{2m}p^2 + \frac{kq^2}{2}.$$

Setzt man die Beziehungen (2.74) in diese Gleichung ein, so ergibt sich

$$-\frac{\partial F_1}{\partial Q} = \frac{1}{2m} \left( \frac{\partial F_1}{\partial q} \right)^2 + \frac{kq^2}{2}. \quad (2.75)$$

Dies ist eine partielle Differenzialgleichung für die erzeugende Funktion  $F_1$ . Die Lösung dieser Differenzialgleichung wird dadurch erschwert, dass sie nichtlinear ist (die Ableitung der Funktion  $F_1$  nach  $q$  tritt quadratisch aus). So führt dieser Hamilton Jacobi Formalismus auch nur in sehr wenigen Beispielen zu Vereinfachungen der Rechnungen.

## 2.4 Poisson Klammern

In diesem Abschnitt soll ein mathematisches Werkzeug eingeführt werden, die sogenannten Poisson Klammern. Diese Poisson Klammern sind eigentlich nur eine abkürzende Beschreibung, es wird sich aber zeigen, dass man mit dieser Abkürzung einige komplexe Zusammenhänge sehr elegant formulieren kann. Darüber hinaus besitzen die Poisson Klammern ein Analogon in der Quantenmechanik, die Kommutator Klammern. Über diese Brücke Poisson Klammern und Kommutator Klammern können viele Verbindungen zwischen der Quantenmechanik und der Klassischen Mechanik verifiziert werden.

Zunächst wollen wir den Begriff der **Dynamischen Variablen** einführen. Die Entwicklung eines mechanischen Systems mit  $n$  Freiheitsgraden wird ja durch seine Trajektorie im Phasenraum beschrieben. Sein Zustand ist also durch die Angabe von  $n$  generalisierten Koordinaten  $q_i$  und entsprechenden kanonischen Impulsen  $p_i$  eindeutig definiert. Eine dynamische Variable ist einfach eine beliebige Funktion

$$f(q_1 \dots q_n, p_1 \dots p_n, t) \quad \text{oder} \quad g(q_1 \dots q_n, p_1 \dots p_n, t),$$

dieser Koordinaten und Impulse. Beispiele für solche dynamischen Variablen sind also etwa die Hamiltonfunktion, die kinetische Energie, der gesamte Drehimpuls des Systems oder aber auch eine Koordinate  $f = q_i$  oder ein Impuls  $f = p_j$ .

Für zwei dynamische Variable  $f$  und  $g$  wird die **Poisson Klammer** durch zwei geschweifte Klammern angezeigt und ist definiert durch

$$\{f, g\} := \sum_{i=1}^n \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial g}{\partial q_i} \frac{\partial f}{\partial p_i}. \quad (2.76)$$

Das Ergebnis ist im Allgemeinen wiederum eine Funktion, die von den Koordinaten und Impulsen des Systems abhängt, also ebenfalls eine dynamische Variable. Die Poisson Klammer ist also wie die Summation oder die Multiplikation eine mathematische Verknüpfung, die jeweils zwei Elementen einer Menge (beziehungsweise einer Gruppe oder eines Körpers) ein Element dieser Menge zuordnet.

Dabei ergeben sich unter anderem die folgenden Eigenschaften, die man leicht beweisen kann ( $f, g, h$  sind dynamische Variable,  $\lambda$  eine Konstante):

$$\{g, f\} = -\{f, g\} \quad \text{und} \quad \{f, f\} = 0, \quad (2.77)$$

$$\{g + f, h\} = \{g, h\} + \{f, h\}, \quad (2.78)$$

$$\{\lambda f, g\} = \lambda \{f, g\}, \quad (2.79)$$

$$\{gf, h\} = f\{g, h\} + \{f, h\}g. \quad (2.80)$$

Aus (2.77) sieht man, dass das Ergebnis der Poisson Klammer von der Reihenfolge der Argumente  $f$  und  $g$  abhängt. Die mathematische Verknüpfung  $\{g, f\}$  ist also nicht kommutativ. Ja im Gegenteil wegen  $\{g, f\} = -\{f, g\}$  nennt man diese Verknüpfung antikommutativ. Andererseits erfüllen die Poisson Klammern das Distributivgesetz bezüglich der Addition von dynamischen Variablen (2.78) und der Multiplikation mit einer Konstanten  $\lambda$  (2.79).

Die Poisson Klammern erfüllen kein Assoziativgesetz, es gilt also im allgemeinen

$$\{f, \{g, h\}\} \neq \{\{f, g\}, h\}.$$

Stattdessen gilt die sogenannte **Jacobi Identität**

$$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0. \quad (2.81)$$

Der Beweis dieser Identität ist etwas mühsam aber durch direktes Ausrechnen leicht zu erzielen.

Die Poisson Klammern sind sehr hilfreich, um Konstanten der Bewegung zu identifizieren. Eine dynamische Variable  $f$ , die nicht explizit von der Zeit abhängt, ist genau dann eine Konstante der Bewegung, wenn die Poisson Klammer von  $f$  mit der Hamiltonfunktion  $H$  des Systems verschwindet. Zum Beweis berechnen wir einfach

$$\begin{aligned} \frac{df}{dt} &= \frac{\partial f}{\partial t} + \sum_{i=1}^n \left[ \frac{\partial f}{\partial q_i} \dot{q}_i + \frac{\partial f}{\partial p_i} \dot{p}_i \right] \\ &= \frac{\partial f}{\partial t} + \sum_{i=1}^n \left[ \frac{\partial f}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial H}{\partial q_i} \right] \\ &= \frac{\partial f}{\partial t} + \{f, H\}, \end{aligned} \quad (2.82)$$

womit die Behauptung bewiesen ist. Bei dem Übergang zur zweiten Zeile wurden die Hamiltonschen Bewegungsgleichungen benutzt, in der dritten Zeile die Definition der Poisson Klammer.

Als einfaches Anwendungsbeispiel zeigen wir noch einmal, dass der Wert der Hamiltonfunktion konstant ist, wenn  $H$  nicht explizit von der Zeit abhängt (siehe (2.15)). Die Anwendung von (2.82) liefert

$$\frac{dH}{dt} = \underbrace{\frac{\partial H}{\partial t}}_{=0 \text{ nach Vor.}} + \{H, H\} \stackrel{\text{siehe (2.77)}}{=} 0.$$

Mit dem Ergebnis (2.82) kann man auch sehr leicht die Hamiltonschen Bewegungsgleichungen verifizieren und darstellen. Mit  $f = p_k$  gilt ja

$$\dot{p}_k = \{p_k, H\} = \sum_{i=1}^n \underbrace{\frac{\partial p_k}{\partial q_i}}_{=0} \frac{\partial H}{\partial p_i} - \underbrace{\frac{\partial p_k}{\partial p_i}}_{=\delta_{ki}} \frac{\partial H}{\partial q_i} = -\frac{\partial H}{\partial q_k}.$$

Entsprechend für  $f = q_k$  ergibt sich

$$\dot{q}_k = \{q_k, H\} = \sum_{i=1}^n \underbrace{\frac{\partial q_k}{\partial q_i}}_{=\delta_{ki}} \frac{\partial H}{\partial p_i} - \underbrace{\frac{\partial q_k}{\partial p_i}}_{=0} \frac{\partial H}{\partial q_i} = \frac{\partial H}{\partial p_k}.$$

Bis zu diesem Punkt haben wir stets einen Satz von generalisierten Koordinaten und Impulsen genommen, um die Poisson Klammern zu berechnen. Die Frage erhebt sich natürlich, ob das Ergebnis von der Wahl dieser Basis des Phasenraumes abhängt. Man kann jedoch zeigen, dass die Berechnung der Poisson Klammer invariant unter einer kanonischen Transformation der Koordinaten und Impulse ist. Sind also  $f(p, q)$  und  $g(p, q)$

wie zuvor dynamische Variable definiert mit den Koordinaten  $p$  und  $q$ , ist eine Kanonische Transformation der Form  $(p, q) \rightarrow (P, Q)$  gegeben und bezeichnen  $\tilde{f}(P, Q)$  und  $\tilde{g}(P, Q)$  die entsprechenden Variablen, dargestellt in den neuen Koordinaten, so dass

$$\tilde{f}(P, Q) = f(p, q) \quad \text{und} \quad \tilde{g}(P, Q) = g(p, q)$$

dann ist das auch die Poisson Klammer invariant unter der Koordinatentransformation, also

$$\{f, g\}_{p,q} = \{\tilde{f}, \tilde{g}\}_{P,Q}, \quad (2.83)$$

wobei die unteren Indices ( $p, q$  beziehungsweise  $P, Q$ ) angeben in welcher Basis die Poissonklammer berechnet wird.

Der Beweis von (2.83) ist etwas aufwändig und wird deshalb hier nicht ausgeführt.<sup>1</sup> Das Ergebnis ist aber sofort plausibel, wenn man sich vor Augen hält, dass ja die zeitliche Entwicklung einer dynamischen Variable unabhängig von der Basis sein muss, in der man diese Observable darstellt. Nun gilt nach (2.82)

$$\frac{df}{dt} = \{f, H\}_{p,q} = \frac{d\tilde{f}}{dt} = \{\tilde{f}, \tilde{H}\}_{P,Q}.$$

Da dies für eine beliebige Funktion  $f$  und auch für jede Hamiltonfunktion gilt, sollte (2.83) auch allgemein gelten.

Eine andere Anwendung sind die sogenannten **Fundamentalklammern**, die wie man durch Ausrechnen leicht verifizieren kann, gegeben sind durch

$$\begin{aligned} \{p_i, p_j\} &= \{q_i, q_j\} = 0, \\ \{q_i, p_j\} &= \delta_{ij}. \end{aligned} \quad (2.84)$$

Diese Ergebnisse für die Fundamentalklammern sind so charakteristisch, dass sogar gilt: Eine Transformation  $q_i, p_j \rightarrow Q_i, P_j$  ist genau dann kanonisch, wenn auch für die neuen Koordinaten gilt:

$$\{P_i, P_j\} = \{Q_i, Q_j\} = 0 \quad \text{und} \quad \{Q_i, P_j\} = \delta_{ij}.$$

Zum Beweis berechnen wir

$$\begin{aligned} \dot{Q}_j &= \{Q_j, H\} = \sum_l \left[ \frac{\partial Q_j}{\partial q_l} \frac{\partial H}{\partial p_l} - \frac{\partial Q_j}{\partial p_l} \frac{\partial H}{\partial q_l} \right] \\ \dot{P}_j &= \{P_j, H\} = \sum_l \left[ \frac{\partial P_j}{\partial q_l} \frac{\partial H}{\partial p_l} - \frac{\partial P_j}{\partial p_l} \frac{\partial H}{\partial q_l} \right]. \end{aligned} \quad (2.85)$$

Stellen wir uns nun die Hamiltonfunktion dargestellt in den neuen Variablen vor, so ergibt sich

$$\begin{aligned} \frac{\partial H}{\partial p_l} &= \sum_k \left[ \frac{\partial \tilde{H}}{\partial Q_k} \frac{\partial Q_k}{\partial p_l} + \frac{\partial \tilde{H}}{\partial P_k} \frac{\partial P_k}{\partial p_l} \right] \\ \frac{\partial H}{\partial q_l} &= \sum_k \left[ \frac{\partial \tilde{H}}{\partial Q_k} \frac{\partial Q_k}{\partial q_l} + \frac{\partial \tilde{H}}{\partial P_k} \frac{\partial P_k}{\partial q_l} \right] \end{aligned} \quad (2.86)$$

<sup>1</sup>der Beweis findet sich z.B. in W.Nolting: Theoretische Physik 2, Abschnitt 2.4.2

Setzt man diese Gleichungen in die erste Zeile von (2.85) so ergibt sich nach einer einfachen Umgruppierung

$$\begin{aligned}\dot{Q}_j &= \sum_k \left[ \frac{\partial \widetilde{H}}{\partial Q_k} \{Q_j, Q_k\} + \frac{\partial \widetilde{H}}{\partial P_k} \{Q_j, P_k\} \right] \\ &= \frac{\partial \widetilde{H}}{\partial P_j}\end{aligned}\quad (2.87)$$

Die letzte Gleichung muss erfüllt sein, damit die Transformation kanonisch ist. Diese Gleichung ist aber genau dann erfüllt, wenn  $\{Q_j, Q_k\} = 0$  und  $\{Q_j, P_k\} = \delta_{jk}$ . Entsprechend können wir (2.86) in die zweite Zeile von (2.85) einsetzen und erhalten wieder nach Umformungen

$$\begin{aligned}\dot{P}_j &= \sum_k \frac{\partial \widetilde{H}}{\partial Q_k} \{P_j, Q_k\} + \frac{\partial \widetilde{H}}{\partial P_k} \{P_j, P_k\} \\ &= -\frac{\partial \widetilde{H}}{\partial Q_j}.\end{aligned}\quad (2.88)$$

Diese letzte Gleichung (die notwendig erfüllt sein muss, damit die Transformation kanonisch ist) ist genau dann erfüllt, wenn  $\{P_j, P_k\} = 0$  und  $\{P_j, Q_k\} = -\delta_{jk}$ , also genau die Fundamentalklammern gelten.

Als eine weitere Anwendung der Poisson Klammern wollen wir nun zeigen, dass die Poisson Klammer von zwei Konstanten der Bewegung selbst eine Konstante der Bewegung ist, d.h.:

$$\frac{df}{dt} = 0 \quad \text{und} \quad \frac{dg}{dt} = 0 \quad \implies \quad \frac{d\{f, g\}}{dt} = 0. \quad (2.89)$$

Zum Beweis schreiben wir die Zeitableitung der Poisson Klammer gemäß (2.82)

$$\begin{aligned}\frac{d\{f, g\}}{dt} &= \frac{\partial \{f, g\}}{\partial t} + \{\{f, g\}, H\} \\ &= \left\{ \frac{\partial f}{\partial t}, g \right\} + \left\{ f, \frac{\partial g}{\partial t} \right\} - \{H, \{f, g\}\} + \{H, \{f, g\}\} + \{f, \{g, H\}\} + \{g, \{H, f\}\}.\end{aligned}$$

Bei dem Übergang zur zweiten Zeile haben wir (2.77) auf die Poisson Klammer mit der Hamiltonfunktion in der ersten Zeile angewandt. Ausserdem wurde ein Null in Form einer Jacobi Identität (2.81) hinzugefügt. Da sich der dritte und vierte Term in dieser zweiten Zeile kompensieren, erhalten wir nach zweimaligem Anwenden der Rechenregel (2.77) auf den letzten Term in der Zeile

$$\begin{aligned}\frac{d\{f, g\}}{dt} &= \left\{ \frac{\partial f}{\partial t}, g \right\} + \{\{H, f\}, g\} + \left\{ f, \frac{\partial g}{\partial t} \right\} + \{f, \{g, H\}\} \\ &= \left\{ \frac{df}{dt}, g \right\} + \left\{ f, \frac{dg}{dt} \right\} \\ &= 0.\end{aligned}$$

Bei dem Übergang zur zweiten Zeile wurde wieder (2.82) angewandt. Da aber die Poisson Klammer von  $0 = df/dt = dg/dt$  (siehe Voraussetzung (2.89)) mit einer dynamischen

Variable verschwindet folgt der Übergang zur letzten Zeile, womit der Beweis erbracht ist.

Eine interessante Anwendung dieses Ergebnisses ergibt sich für den Drehimpuls eines Teilchens. Wir betrachten dazu die Bewegung eines Teilchens in einem konservativen Kraftfeld ohne Zwangsbedingungen. Die kartesischen Koordinaten des Teilchens  $(x, y, z)$  und die zugehörigen Impulskomponenten  $(p_x, p_y, p_z)$  sind also eine geeignete Basis des Phasenraumes. In diesen generalisierten Koordinaten und Impulsen, ergeben sich für die Darstellung der kartesischen Komponenten des Drehimpulses

$$\begin{aligned} l_x &= yp_z - zp_y \\ l_y &= zp_x - xp_z \\ l_z &= xp_y - yp_x. \end{aligned} \tag{2.90}$$

Berechnet man z.B. die Poisson Klammer für die Komponenten  $l_x$  und  $l_y$ , so ergibt sich mit Anwendung des Distributivgesetzes (2.78)

$$\{l_x, l_y\} = \{yp_z, zp_x\} - \{yp_z, xp_z\} - \{zp_y, zp_x\} + \{zp_y, xp_z\}.$$

Zur weiteren Behandlung betrachten wir den ersten Term auf der rechten Seite und wenden die Regel (2.80) wiederholt an. Daraus ergibt sich

$$\begin{aligned} \{yp_z, zp_x\} &= \underbrace{\{y, z\}}_{=0} p_z p_x + \underbrace{\{y, p_x\}}_{=0} p_z z + \underbrace{\{p_z, z\}}_{=-1} y p_x + \underbrace{\{p_z, p_x\}}_{=0} y z \\ &= -yp_x. \end{aligned}$$

Die Ergebnisse für die Poisson Klammern in der ersten Zeile ergeben sich durch die Fundamental Klammern (2.84). Entsprechende Behandlung der anderen Terme in der obigen Gleichung liefert schliesslich

$$\{l_x, l_y\} = -yp_x + xp_y = l_z. \tag{2.91}$$

Ganz analog kann man auch zeigen, dass die Beziehung auch gilt, wenn man in dieser Gleichung die Komponenten  $l_x, l_y, l_z$  zyklisch vertauscht, also:

$$\{l_x, l_y\} = l_z \quad \text{und} \quad \{l_x, l_y\} = l_z. \tag{2.92}$$

Die Komponenten des Drehimpulses sind also keine unabhängigen dynamischen Variablen. Sind nämlich zwei Komponenten, also z.B.  $l_x$  und  $l_y$  Konstanten der Bewegung, so gilt mit (2.89) und (2.91), dass auch die dritte Komponente  $l_z$  konstant ist.

## 2.5 Noether Theorem: Symmetrien und Konstanten

Den Zusammenhang zwischen Symmetrien und Konstanten der Bewegung haben wir bereits im Abschnitt 1.7 angesprochen. Hier soll gezeigt werden, dass jeder Symmetrie des Systems eine Erhaltungsgröße zugeordnet werden kann. Zu jeder kanonischen Transformation, bei deren Anwendung sich die Hamilton Funktion nicht ändert, können wir eine dynamische Variable definieren, die bei der Entwicklung des Systems erhalten bleibt, also eine Konstante der Bewegung ist. Dieser Zusammenhang wurde als Theorem von Emmy Noether<sup>2</sup> formuliert. Dieses Noether Theorem wird häufig im Rahmen des Lagrange Formalismus formuliert (siehe z.B. Kapitel 15 in T. Fließbach: Mechanik). Hier werden wir es im Rahmen des Hamilton Formalismus behandeln. Es spielt eine besonders große Rolle in der Feldtheorie.

Zur Formulierung des Noether Theorems definieren wir eine **infinitesimale kanonische Transformation**

$$\begin{array}{l} p_i \\ q_i \end{array} \quad \Longrightarrow \quad \begin{array}{l} \alpha_i = p_i + \delta p_i \\ \beta_i = q_i + \delta q_i \end{array}, \quad (2.93)$$

als eine kanonische Transformation, bei der die neuen Koordinaten und Impulse sich nur um infinitesimal kleine Korrekturen,  $\delta q_i$  und  $\delta p_i$  aus den ursprünglichen Koordinaten und Impulsen ergeben. Damit die Transformation kanonisch ist, muss es eine erzeugende Funktion geben, mit der diese Transformation beschrieben wird. Wir betrachten dazu eine erzeugende Funktionen vom Typ  $F_2$  (siehe (2.61)) und machen den Ansatz

$$F_2(q_i, \alpha_i) = \sum_{i=1}^n \alpha_i q_i + \varepsilon f(q_i, \alpha_i). \quad (2.94)$$

Dabei ist  $f(q_i, \alpha_i)$  eine beliebige Funktion der alten Koordinaten und neuen Impulsen und der Faktor  $\varepsilon$  soll infinitesimal klein sein, so dass die Beiträge dieses zweiten Terms  $\varepsilon f$  zur erzeugenden Funktion  $F$  infinitesimal sind. Mit dieser erzeugenden Funktion und den Regeln (2.61) ergibt sich eine kanonische Transformation

$$\begin{array}{l} p_i \\ \beta_i \end{array} = \begin{array}{l} \frac{\partial F_2}{\partial q_i} \\ \frac{\partial F_2}{\partial \alpha_i} \end{array} = \begin{array}{l} \alpha_i + \varepsilon \frac{\partial f}{\partial q_i} \\ q_i + \varepsilon \frac{\partial f}{\partial \alpha_i} \end{array}. \quad (2.95)$$

Der Vergleich von (2.95) mit (2.93) zeigt, dass wir den Absatz (2.94) geschickt gewählt haben, denn die Änderungen der Koordinaten und Impulse ergeben sich zu

$$\begin{array}{l} \delta p_i \\ \delta q_i \end{array} = \begin{array}{l} -\varepsilon \frac{\partial f(q_i, \alpha_i)}{\partial q_i} \\ \varepsilon \frac{\partial f(q_i, \alpha_i)}{\partial \alpha_i} \end{array}. \quad (2.96)$$

---

<sup>2</sup>Dieses Theorem ist nach der Mathematikerin und Physikerin Emmy Noether (1882-1935) benannt. Als Jüdin musste Emmy Noether 1933 ihren Lehrstuhl für Mathematik an der Universität Göttingen verlassen. Sie flüchtete in die USA und lehrte am *Institute for Advanced Study* in Princeton und am *Bryn Mawr College*

Die Änderungen  $\delta p_i$  und  $\delta q_i$  sind also proportional zu  $\varepsilon$  und damit infinitesimal klein. Für die weitere Behandlung betrachten wir das Argument  $\alpha_i$  in der Funktion  $f$  und entwickeln die Funktion  $f$  um den Punkt  $\alpha_j = p_j$  im Sinne eine Taylor Entwicklung

$$f(\alpha_i, q_i) = f(p_i, q_i) + \sum_j \frac{\partial f(p, q)}{\partial p_j} \delta p_j + \dots$$

Diese Taylor Entwicklung kann nach dem zweiten Glied abgebrochen werden, da die  $\delta p_j$  infinitesimal sind und Terme, die quadratisch in infinitesimalen Größen sind, vernachlässigt werden sollen. Damit ergibt sich also für

$$\delta p_i = -\varepsilon \frac{\partial f(q_i, \alpha_i)}{\partial q_i} = -\varepsilon \frac{\partial f(q_i, p_i)}{\partial q_i} - \varepsilon \frac{\partial}{\partial q_i} \sum_j \frac{\partial f(p, q)}{\partial p_j} \delta p_j + \dots$$

Auch hier werden wieder die Terme zweiter und höherer Ordnung in infinitesimalen Größen vernachlässigt und die Korrekturterme (2.96) ergeben sich zu

$$\begin{aligned} \delta p_i &= -\varepsilon \frac{\partial f(q_i, p_i)}{\partial q_i} = \varepsilon \{p_i, f\}, \\ \delta q_i &= \varepsilon \frac{\partial f(q_i, p_i)}{\partial p_i} = \varepsilon \{q_i, f\}. \end{aligned} \quad (2.97)$$

Die zweite Gleichung in jeder dieser Zeilen ergibt sich direkt aus der Definition der Poisson Klammern (2.76). Die infinitesimalen Abweichungen der neuen Koordinaten und Impulse von den alten sind also ausschliesslich durch die Funktion  $f$ , beziehungsweise durch die entsprechenden Poisson Klammern von  $f$  mit seinen Argumenten definiert. Deshalb heisst  $f$  auch die **Erzeugende Funktion der infinitesimalen Transformation**.

Wir betrachten nun die Änderung einer dynamischen Variable  $g$ , wenn die Argumente dieser dynamischen Variablen infinitesimal verändert werden

$$\delta g := g(\alpha_i, \beta_i) - g(p_i, q_i). \quad (2.98)$$

Ist  $f$  die erzeugende Funktion dieser infinitesimalen Transformation der Variablen  $(p, q \rightarrow \alpha, \beta)$ , so berechnet sich diese Änderung gemäß

$$\delta g = \varepsilon \{f, g\}. \quad (2.99)$$

Zum Beweis dieser Beziehung berechnen wir in einer Taylor Entwicklung

$$\begin{aligned} g(\alpha_i, \beta_i) &= g(p_i, q_i) + \sum_i \left[ \frac{\partial g}{\partial q_i} \delta q_i + \frac{\partial g}{\partial p_i} \delta p_i \right] + \dots \\ &= g(p_i, q_i) + \sum_i \left[ \frac{\partial g}{\partial q_i} \varepsilon \frac{\partial f}{\partial p_i} + \frac{\partial g}{\partial p_i} \left( -\varepsilon \frac{\partial f}{\partial q_i} \right) \right] \\ &= g(p_i, q_i) + \varepsilon \{f, g\}. \end{aligned} \quad (2.100)$$

Die durch Punkte angedeuteten Terme in der ersten Reihe werden beim Übergang zur zweiten Zeile vernachlässigt, da sie quadratisch in den infinitesimalen Änderungen sind. Bei dem Übergang zur zweiten Zeile wurden die Beziehungen (2.97) eingesetzt. Die dritte Zeile und damit der Beweis von(2.99) ergibt sich aus der Definition der Poisson Klammer.

Nach diesen Vorüberlegungen können wir das Noether Theorem formulieren



Sei  $f$  eine dynamische Variable, die nicht explizit von der Zeit abhängt. Die Hamiltonfunktion des Systems  $H$  ist genau dann invariant unter der infinitesimalen Transformation, die durch  $f$  erzeugt wird (d.h.  $\delta H$  definiert gemäß (2.99) ist identisch null), wenn  $f$  eine Konstante der Bewegung ist.

Nach den geleisteten Vorarbeiten ist der Beweis dieses Theorems sehr einfach: Nach (2.99) ist  $H$  genau dann invariant unter der infinitesimalen Transformation, wenn die Poisson Klammer  $\{f, H\} = 0$ . Dies ist aber nach (2.82) gleichbedeutend damit, dass  $f$  eine Konstante der Bewegung ist. Ausgedrückt in Symbolen kann der Beweis also auf eine Zeile reduziert werden

$$\delta H = 0 \quad \Leftrightarrow \quad \{f, H\} = 0 \quad \Leftrightarrow \quad \frac{df}{dt} = \frac{\partial f}{\partial t} = 0. \quad (2.101)$$

Im weiteren Verlauf dieses Abschnittes sollen verschiedene Beispiele für die Anwendung dieses Theorems besprochen werden. Als erstes Beispiel wollen wir den Fall behandeln, dass die Hamilton Funktion von einer bestimmten Koordinate  $q_i$  nicht abhängt. Damit ändert sich der Wert der Hamilton Funktion nicht, wenn diese Koordinate um einen Wert  $\varepsilon$  modifiziert wird. Die Hamilton Funktion ist also invariant unter der Transformation

$$\begin{aligned} p_j &\Rightarrow \alpha_j = p_j + 0 \\ q_j &\Rightarrow \beta_j = q_j + \varepsilon \delta_{i,j}. \end{aligned} \quad (2.102)$$

Wir können uns nun davon überzeugen, dass dieses Transformation durch die Funktion

$$f = p_i \quad (2.103)$$

erzeugt wird. Dazu verifizieren wir mit (2.97)

$$\begin{aligned} \delta p_j &= -\varepsilon \frac{\partial p_i}{\partial q_j} = 0, \\ \delta q_j &= \varepsilon \frac{\partial p_i}{\partial p_j} = \varepsilon \delta_{i,j}. \end{aligned} \quad (2.104)$$

Nach dem Noether Theorem ist also die Tatsache, dass  $H$  invariant unter einer Änderung der Koordinate  $q_i$  ist, gleichbedeutend damit, dass der zugehörige kanonische Impuls eine Konstante der Bewegung ist. Natürlich ist dieses Ergebnis längst bekannt etwa durch die Bewegungsgleichung

$$\dot{p}_i = -\frac{\partial H}{\partial q_i},$$

aber mit diesem trivialen Beispiel sollten die Begriffe, die bei der Formulierung des Noether Theorems auftreten, verdeutlicht werden.

In einem zweiten Beispiel betrachten wir ein System aus vielen Teilchen ohne Zwangsbedingungen so dass die kartesischen Koordinaten,  $x_i, y_i, z_i$ , der einzelnen Teilchen  $i$  als generalisierte Koordinaten und die zugehörigen Komponenten des Newtonschen Impulses,  $p_{xi}, p_{yi}, p_{zi}$ , als kanonische Impulse betrachtet werden können. Wir betrachten dann als ein Beispiel für die Erzeugende Funktion einer infinitesimalen kanonischen Transformation die Funktion

$$f_x = \sum_{i=1}^N p_{xi}, \quad (2.105)$$

das ist die  $x$ -Komponente des Gesamtimpulses des Systems. Nach (2.97) erzeugt diese Funktion  $f_x$  die folgende kanonische Transformation für die Impulskomponenten

$$\begin{aligned}\delta p_{xj} &= \varepsilon \left\{ p_{xj}, \sum_{i=1}^N p_{xi} \right\} \\ &= \varepsilon \sum_{i=1}^N \{ p_{xj}, p_{xi} \} \\ &= 0.\end{aligned}\tag{2.106}$$

Bei dem Übergang zur zweiten Zeile wurden die Rechenregeln für die Poisson Klammern benutzt und bei dem Übergang zur dritten Zeile die Eigenschaften der Fundamentalklammern (2.84).

Natürlich kann man ganz analog auch zeigen, dass die Funktion  $f_x$  aus (2.105) auch keine Transformation der anderen kartesischen Impulskomponenten  $p_{yj}$  und  $p_{zj}$  erzeugt. Andererseits ergibt sich für

$$\begin{aligned}\delta x_j &= \varepsilon \left\{ x_j, \sum_{i=1}^N p_{xi} \right\} \\ &= \varepsilon \sum_{i=1}^N \{ x_j, p_{xi} \} \\ &= \varepsilon \sum_{i=1}^N \delta_{ji} \\ &= \varepsilon,\end{aligned}\tag{2.107}$$

während die  $y_i$  und  $z_i$  durch  $f_x$  nicht modifiziert werden. Die Funktion  $f_x$  erzeugt also eine Transformation, bei der die  $x$  Koordinaten aller Teilchen um die gleiche Strecke  $\varepsilon$  verschoben werden, also eine Translation des Gesamtsystems in  $x$ -Richtung. Das Noether Theorem besagt also in dieser Anwendung: Ist die Hamiltonfunktion invariant unter einer solchen Translation in der  $x$ -Richtung, dann impliziert dies, dass die  $x$ -Komponente des Gesamtimpulses,  $f_x = \sum_i p_{xi}$ , eine Konstante der Bewegung ist.

Natürlich gilt dies analog auch für die  $y$  und  $z$  Komponente. Ist also eine Hamiltonfunktion invariant unter einer Translation in eine beliebige Raumrichtung, so ist der Gesamtimpuls des Systems eine Erhaltungsgröße.

Als drittes Beispiel betrachten wir die infinitesimale Transformationen, die durch eine kartesische Komponente des Gesamtdrehimpulses des Systems also z.B. durch die  $z$  Komponente

$$f = \sum_i (x_i p_{yi} - y_i p_{xi}),\tag{2.108}$$

erzeugt wird. Dazu berechnen wir die zugehörige Transformation der Koordinaten wie z.B.

$$\begin{aligned}\delta x_j &= \varepsilon \left\{ x_j, \sum_{i=1}^N (x_i p_{yi} - y_i p_{xi}) \right\} \\ &= \varepsilon \sum_{i=1}^N (\{ x_j, x_i p_{yi} \} - \{ x_j, y_i p_{xi} \})\end{aligned}$$

$$\begin{aligned}
&= \varepsilon \sum_{i=1}^N \left( \underbrace{\{x_j, x_i\}}_{=0} p_{yi} + \underbrace{\{x_j, p_{yi}\}}_{=0} x_i - \underbrace{\{x_j, p_{xi}\}}_{=\delta_{ij}} y_i - \underbrace{\{x_j, y_i\}}_{=0} p_{xi} \right) \\
&= -\varepsilon y_j \\
\delta y_j &= \varepsilon x_j \\
\delta z_j &= 0 \\
\delta p_{xj} &= -\varepsilon p_{yj} \\
\delta p_{yj} &= \varepsilon p_{xj} \\
\delta p_{zj} &= 0.
\end{aligned} \tag{2.109}$$

Bezeichnen wir also mit  $\vec{a}$  den Ortsvektor  $\vec{r}_i$  oder Impulsvektor  $\vec{p}_i$  eines Teilchens  $i$ , so werden also all diese Vektoren durch die infinitesimale Transformation übergeführt in

$$\begin{aligned}
a_x &\Rightarrow \widetilde{a}_x = a_x - \varepsilon a_y \\
a_y &\Rightarrow \widetilde{a}_y = a_y + \varepsilon a_x \\
a_z &\Rightarrow \widetilde{a}_z = a_z
\end{aligned} \tag{2.110}$$

beziehungsweise dargestellt in Matrixschreibweise

$$\begin{aligned}
\widetilde{\vec{a}} &= \begin{pmatrix} 1 & -\varepsilon & 0 \\ \varepsilon & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} \\
&= \lim_{\varepsilon \rightarrow 0} \begin{pmatrix} \cos \varepsilon & -\sin \varepsilon & 0 \\ \sin \varepsilon & \cos \varepsilon & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix}.
\end{aligned} \tag{2.111}$$

Die infinitesimale Transformation, die durch die  $z$ -Komponente des Gesamtdrehimpulses  $f = l_z$  in (2.108) erzeugt wird entspricht also einer Drehung aller Vektoren  $\vec{a}$  um die  $z$ -Achse mit dem infinitesimalen Winkel  $\varepsilon$  (beziehungsweise einer Drehung des Koordinatensystems um den Winkel  $-\varepsilon$ ). Ändert sich die Hamiltonfunktion bei einer solchen Transformation nicht, ist ihr Wert also invariant gegenüber einer Drehung um die  $z$ -Achse, so ist nach dem Noether Theorem, die  $z$ -Komponente des Gesamtdrehimpulses eine Erhaltungsgröße für die Entwicklung des Systems. Man sagt die Hamiltonfunktion ist **rotationsinvariant** bezüglich der  $z$ -Achse. Entsprechendes lässt sich natürlich auch für die  $x$ - und  $y$ -Achse formulieren.

Als letztes Beispiel wollen wir die Hamiltonfunktion als erzeugende Funktion einer infinitesimalen Transformation heranziehen. Hängt die Hamiltonfunktion nicht explizit von der Zeit ab, so ist sie ja eine Konstante der Bewegung (siehe z.B.(2.15)). Mit (2.97) ergeben sich die folgenden Ausdrücke für die Transformationen

$$\begin{aligned}
\delta p_i &= \varepsilon \{p_i, H\} = \varepsilon \dot{p}_i \\
\delta q_i &= \varepsilon \{q_i, H\} = \varepsilon \dot{q}_i
\end{aligned} \tag{2.112}$$

Dies bedeutet also für die kanonischen Impulse und Koordinaten

$$\begin{aligned}
p_i(t) &\Rightarrow p_i(t) + \varepsilon \dot{p}_i = p_i(t + \varepsilon) \\
q_i(t) &\Rightarrow q_i(t) + \varepsilon \dot{q}_i = q_i(t + \varepsilon).
\end{aligned} \tag{2.113}$$

Die Koordinaten und Impulse zur Zeit  $t$  werden also transformiert in die Koordinaten und Impulse zur Zeit  $t + \varepsilon$  also um eine infinitesimale Zeit verschoben. Die Hamiltonfunktion ist also die erzeugende Funktion für eine Verschiebung in der Zeit.

### 2.5.1 Infinitesimale Drehungen und Drehungen um einen vorgegebenen Winkel

Zum Abschluss dieses Abschnittes sollen Zusammenhänge zwischen einer infinitesimalen Drehung und einer Drehung um einen endlichen, von Null verschiedenen Winkel erläutert werden. Dabei wollen wir uns zunächst auf Drehungen von Vektoren um die  $z$ -Achse beschränken und schreiben die Matrix für die infinitesimale Drehung in (2.111) in der Form

$$(\mathbb{1} + \varepsilon A_z) \quad \text{mit} \quad A_z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (2.114)$$

und der Eins-Matrix  $\mathbb{1}$ . Wir werden uns nun davon überzeugen, dass die Rotation um einen Winkel  $\varphi$  in  $N$  Teilschritte von Drehungen jeweils um den Winkel  $\varphi/N$  unterteilt werden kann. Lässt man die Zahl der Teilschritte gegen Unendlich wachsen,  $N \rightarrow \infty$ , so sind die einzelnen Drehwinkel infinitesimal und wir können für die Matrix der Gesamtdrehung schreiben

$$R_z(\varphi) = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} = \lim_{N \rightarrow \infty} \left( \mathbb{1} + \frac{\varphi}{N} A_z \right)^N. \quad (2.115)$$

Zum Beweis dieser Beziehung berechnen wir zunächst einmal die Potenzen der Matrix  $A_z$  zu

$$\begin{aligned} A_z^2 &= \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ A_z^3 &= -A_z \\ A_z^4 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ A_z^5 &= A_z, \end{aligned} \quad (2.116)$$

womit sich natürlich auch die höheren Potenzen ergeben. Berechnet man nun mit der Binomialformel (die Klammern bezeichnen die Binomialkoeffizienten)

$$\begin{aligned} \left( \mathbb{1} + \frac{\varphi}{N} A_z \right)^N &= \sum_{i=0}^N \binom{N}{i} \left( \frac{\varphi}{N} A_z \right)^i \\ &= \mathbb{1} + \binom{N}{1} \frac{\varphi}{N} A_z + \binom{N}{2} \frac{\varphi^2}{N^2} A_z^2 + \dots \end{aligned} \quad (2.117)$$

so kann man sich die einzelnen Matrixelemente der resultierenden Matrix berechnen. Als Beispiel soll das Element in der ersten Zeile und Spalte herangezogen werden. Wegen der

Ergebnisse für die Potenzen  $A_z^i$  in (2.116) ergeben sich nur Beiträge zu diesem Element für die geraden Potenzen von  $A_z$  in der Form

$$\begin{aligned} 1 - \binom{N}{2} \frac{\varphi^2}{N^2} + \binom{N}{4} \frac{\varphi^4}{N^4} - \dots \\ = 1 - \frac{N(N-1)}{2!} \frac{\varphi^2}{N^2} + \frac{N(N-1)(N-2)(N-3)}{4!} \frac{\varphi^4}{N^4} - \dots \end{aligned} \quad (2.118)$$

Im Grenzfall  $N \rightarrow \infty$  ergibt sich daraus

$$1 - \frac{1}{2!} \varphi^2 + \frac{1}{4!} \varphi^4 - \dots = \cos \varphi$$

also genau das Ergebnis für das erste Matrixelement in (2.115). In analoger Weise kann man (2.115) auch für die anderen Matrixelemente verifizieren.

Eine andere Sichtweise ergibt sich aus der folgenden Überlegung. Bezeichnen wir mit  $\vec{a}(\varphi)$  den Vektor, der durch Drehung um den Winkel  $\varphi$  aus den Vektor  $\vec{a} = \vec{a}(0)$  entstanden ist, also mit der Bezeichnung aus (2.115)

$$\vec{a}(\varphi) = R_z(\varphi) \vec{a}(0).$$

Dann ergibt sich  $\vec{a}(\varphi + d\varphi)$ , der durch eine zusätzliche Drehung um den infinitesimalen Winkel  $d\varphi$  entsteht durch

$$\vec{a}(\varphi + d\varphi) = (\mathbb{1} + d\varphi A_z) \vec{a}(\varphi)$$

Dies kann man auch als Differentialgleichung schreiben

$$\frac{d\vec{a}}{d\varphi} = A_z \vec{a}$$

mit der Lösung

$$\vec{a}(\varphi) = e^{\varphi A_z} \vec{a}(0).$$

Natürlich müssen wir noch erklären, wie die Matrix  $A_z$  als Exponent der Zahl  $e$  zu verstehen ist. Die Definition ist so zu verstehen, dass wir die Exponentialfunktion in Reihe entwickeln. Mit den Ergebnissen aus (2.116) ergibt sich dann

$$e^{\varphi A_z} = \sum_{i=0}^{\infty} \frac{\varphi^i}{i!} A_z^i = R_z(\varphi).$$

Zum guten Schluss seien jetzt Drehungen um  $x$ ,  $y$  und  $z$ -Achse um infinitesimale Winkel  $d\Omega_x$ ,  $d\Omega_y$  und  $d\Omega_z$  betrachtet. Zunächst überzeugen wir uns davon, dass bei infinitesimalen Winkeln die Reihenfolge der Drehungen um die verschiedenen Achsen irrelevant ist. So gilt z.B.

$$\begin{aligned} R_z(d\Omega_z) R_x(d\Omega_x) &= (\mathbb{1} + d\Omega_z A_z) (\mathbb{1} + d\Omega_x A_x) \\ &= \mathbb{1} + d\Omega_z A_z + d\Omega_x A_x + \underbrace{d\Omega_z A_z d\Omega_x A_x}_{=0} \\ &= (\mathbb{1} + d\Omega_x A_x) (\mathbb{1} + d\Omega_z A_z) \\ &= R_x(d\Omega_x) R_z(d\Omega_z). \end{aligned} \quad (2.119)$$

Der Grund für diese Identität liegt darin, dass das Glied quadratisch in infinitesimalen Winkeln  $d\Omega_z d\Omega_x$  in der zweiten Zeile unterdrückt werden kann. Damit ergibt sich also eine allgemeine Rotation um die infinitesimalen Winkel  $d\Omega_x$ ,  $d\Omega_y$  und  $d\Omega_z$

$$\begin{aligned} R(d\vec{\Omega}) &= (\mathbb{1} + d\Omega_x A_x + d\Omega_y A_y + d\Omega_z A_z) \\ &= \begin{pmatrix} 1 & -d\Omega_z & d\Omega_y \\ d\Omega_z & 1 & -d\Omega_x \\ -d\Omega_y & d\Omega_x & 1 \end{pmatrix}. \end{aligned} \quad (2.120)$$

Damit ergibt sich für einen entsprechend gedrehten Vektor

$$\begin{aligned} \tilde{\vec{a}} &= R(d\vec{\Omega})\vec{a} \\ &= \vec{a} + \begin{pmatrix} 0 & -d\Omega_z & d\Omega_y \\ d\Omega_z & 0 & -d\Omega_x \\ -d\Omega_y & d\Omega_x & 0 \end{pmatrix} \vec{a} \\ &= \vec{a} + d\vec{\Omega} \times \vec{a}. \end{aligned} \quad (2.121)$$

Der Übergang zur letzten Zeile kann leicht dadurch verifiziert werden, dass man einerseits das Produkt der Matrix in der zweiten Zeile mit  $\vec{a}$  berechnet und andererseits das Vektorprodukt aus  $d\vec{\Omega}$ , mit den kartesischen Komponenten  $d\Omega_x$ ,  $d\Omega_y$  und  $d\Omega_z$  und dem Vektor  $\vec{a}$  bestimmt. Betrachten wir nun den Vektor  $\vec{a}$  einerseits in einem mitrotierenden (körperfesten) Koordinatensystem und andererseits in einem Koordinatensystem, in dem dieser Vektor  $\vec{a}$  nach einer Zeit  $dt$  in den Vektor  $\tilde{\vec{a}}$  gedreht ist, so ergibt sich für zeitliche Änderungen in diesem Laborsystem

$$\frac{d\tilde{\vec{a}}}{dt} = \left( \frac{d\vec{a}}{dt} \right)_{\text{Lab}} = \left( \frac{d\vec{a}}{dt} \right)_{\text{Kö}} + \frac{d\vec{\Omega}}{dt} \times \vec{a}, \quad (2.122)$$

die Beziehung zwischen Zeitableitungen von Vektorfunktionen im Laborsystem und in einem körperfesten Koordinatensystem, das mit der Winkelgeschwindigkeit

$$\vec{\omega} = \frac{d\vec{\Omega}}{dt}$$

mitrotiert (Vergleiche Vorlesung Physik I).

# Kapitel 3

## Nichtlineare Systeme, Chaos

### 3.1 Logistische Gleichung

In diesem Abschnitt wird eine einfache Iterationsvorschrift behandelt, die **Logistische Gleichung**, der Form

$$x_{n+1} = f_a(x_n) = a x_n (1 - x_n) \quad \text{für } 0 \leq x_n \leq 1 \quad \text{und} \quad a \geq 0. \quad (3.1)$$

Diese Gleichung definiert eine mathematische Funktion  $f_a(x)$ , die jedem Punkt  $x_n$  einen Bildpunkt  $x_{n+1}$  zuordnet. Startet man mit einem beliebigem Wert  $x_0$  aus dem Intervall  $[0, 1]$ , so definiert diese Gleichung eine Folge von Punkten  $x_n$  aus diesem Intervall. Diese Folge von Punkten hat zunächst nichts mit einem System aus der Mechanik zu tun. Wir werden aber im nächsten Abschnitt entsprechende Folgen von Vektoren  $\vec{x}_n$  betrachten, die jeweils einen Punkt im Phasenraum eines physikalischen Systems identifizieren. Sieht man einen solchen Phasenraumpunkt  $\vec{x}_n$  als Identifikation des Systems zu einem Zeitpunkt  $t_n = n\Delta t$  an, so beschreibt die Folge von Vektoren die  $\vec{x}_n$  den Weg der zeitlichen Entwicklung des Systems im Phasenraum durch die Angabe der Phasenraumvektoren für eine Folge von diskreten Zeitpunkten.

Die Betrachtung der logistischen Gleichung dient hier nur als einfaches Beispiel in einem eindimensionalen Raum, an dem wir bestimmte Begriffe einführen und verdeutlichen wollen. Andererseits kann man die logistische Gleichung aber auch als ein einfaches Modell für die zeitliche Entwicklung eines biologischen Systems ansehen. Die dimensionslose Zahl  $x$  steht dabei für die Bevölkerungsdichte einer betrachteten Art bezogen auf eine maximal mögliche Bevölkerungsdichte  $x = 1$ . Der Parameter  $a$  in der logistischen Gleichung (3.1) steht dabei für das Wachstum der Bevölkerung von einer Generation  $n$  zur nächsten Generation  $n + 1$ . Würde auf der rechten Seite der Gleichung einfach der lineare Term  $a x_n$  stehen, so würde die Bevölkerung für  $a > 1$  exponentiell wachsen, beziehungsweise bei  $a < 1$  entsprechend schrumpfen. Eine solche lineare Beziehung liefert also ein monotones Verhalten. Das System wird erst interessant durch den zusätzlichen Faktor  $(1 - x_n)$ , der die Gleichung in eine nichtlineare Beziehung umwandelt. Dieser Faktor  $(1 - x_n)$  kann bei der biologischen Interpretation der logistischen Gleichung als Kontrollfaktor interpretiert werden, der etwa die Begrenzung einer Bevölkerungsdichte durch begrenzte Nahrungsressourcen simuliert.

Betrachten wir als Beispiel die logistische Gleichung mit dem “Wachstumsparameter”

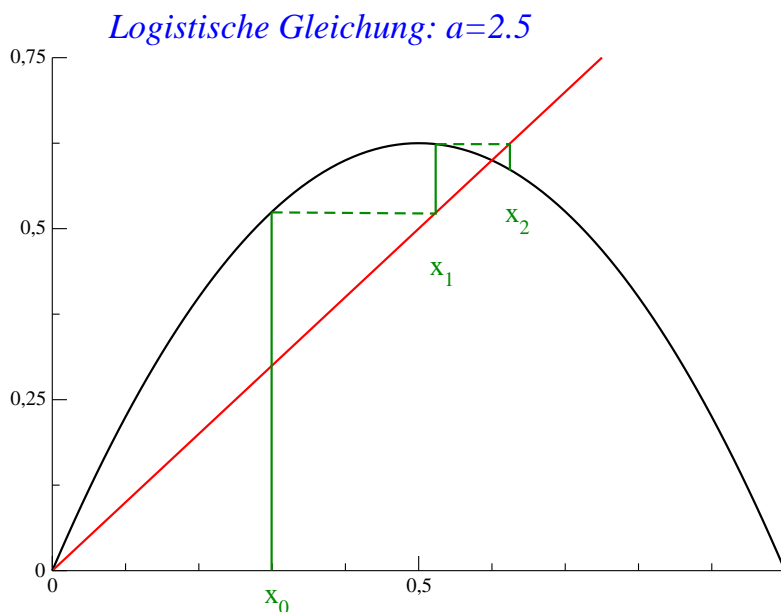


Abbildung 3.1: Grafische Konstruktion der ersten Glieder der Folge, die durch die logistische Gleichung mit  $a=2.5$  definiert ist, ausgehend von  $x_0=0.3$

$a = 2.5$  und einem Startwert  $x_0 = 0.3$ . Damit ergibt sich nach (3.1)

$$x_1 = a * x_0 * (1 - x_0) = 2.5 * 0.3 * (1 - 0.3) = 0.525,$$

und daraus durch nochmalige Anwendung

$$x_2 = 2.5 * 0.525 * (1 - 0.525) = 0.6234375$$

und so weiter. Diese Folge von Punkten  $x_i$  lässt sich auch sehr leicht grafisch erzeugen und darstellen wie in in Abb. 3.1 demonstriert wird. In dieser Abbildung sind die Funktionen

$$y = x \quad \text{und} \quad y = f_a(x) = a * x - a * x^2, \quad (3.2)$$

als rote Gerade beziehungsweise schwarze Kurve dargestellt. Wir beginnen mit dem Wert  $x_0$  und bestimmen  $x_1 = y = f_a(x_0)$  (siehe grüne Gerade, die zum Wert  $x_1 = f_a(x_0)$  auf der Abszisse (y-Achse) führt). Durch die gestrichelte Verbindungslinie mit der Geraden  $y = x$  erhalten wir den zugehörigen Punkt  $x_1$  auf der Ordinate (x-Achse). Die durchgezogene Linie führt uns dann zum Punkt  $y = f_a(x_1)$  und bestimmt damit den Wert  $x_2$ . Durch diese Konstruktion erhält man also die Folge  $x_i$ , die durch die logistische Gleichung definiert ist.

Verfolgt man eine solche Iteration der logistischen Gleichung für  $a=2.5$  entweder mit der grafischen Darstellung von Abb. 3.1 oder durch sie explizite Berechnung mit (3.1) so stellt man nach einigen Iterationsschritten fest, dass diese Iteration zum Wert  $x_n = x_{n+1} = 0.6$  führt und zwar unabhängig davon mit welchem Startwert man beginnt. Der Punkt  $x_n = 0.6$  ist (bei  $a=2.5$ ) ein **Fixpunkt**. Ist man bei dieser Iteration einmal bei einem solchen Fixpunkt gelandet, so reproduzieren weitere Iterationen immer wieder den gleiche Wert. Für solche Fixpunkte gilt

$$x_{n+1} = f_a(x_n) = x_n, \quad (3.3)$$



Für diese Fixpunkte ergibt sich also die Bestimmungsgleichung

$$\begin{aligned}x &= f_a(x) \\x &= ax - ax^2 \\ax^2 + (1 - a)x &= 0\end{aligned}\tag{3.4}$$

Diese Gleichung besitzt 2 Lösungen, die in der Grafik von Abb. 3.1 gerade durch die Schnittpunkte der beiden Funktionen aus (3.2) gegeben sind, nämlich

$$x = 0 \quad \text{und} \quad x = 1 - \frac{1}{a} = (0.6 \text{ für } a = 2.5).\tag{3.5}$$

Warum führen aber alle Iterationen mit  $a > 1$  auf den Fixpunkt  $x = 0.6$  und keine auf den Fixpunkt  $x = 0$ ? Man bezeichnet Fixpunkte einer Gleichung, also Werte  $x = x_n$ , die (3.3) erfüllen, die Iterationen “anziehen”, als **Attraktoren**. Für einen solchen Attraktor muss offensichtlich gelten, dass in der Umgebung dieses Attraktors zwei Bildpunkte der Abbildung  $f_a(x)$  näher zusammenliegen, als die Punkte für die die Funktion berechnet wird also:

$$|f(x + \varepsilon) - f(x)| < |(x + \varepsilon) - x| = |\varepsilon|.\tag{3.6}$$

Nur wenn diese Bedingung erfüllt ist, wird die Folge der Punkte, die durch Iteration der Abbildung  $f_a$  definiert ist zum Fixpunkt hin konvergieren. Das Kriterium (3.6) ist äquivalent zu

$$\frac{|f(x + \varepsilon) - f(x)|}{|\varepsilon|} \rightarrow \left| \frac{df}{dx} \right| < 1.\tag{3.7}$$

Für die logistische Gleichung bedeutet dies

$$\left| \frac{df}{dx} \right| = |a - 2ax| < 1.\tag{3.8}$$

Falls  $a > 1$  ist diese Gleichung offensichtlich nicht erfüllt für den Fixpunkt  $x = 0$ . Andererseits erfüllt der Fixpunkt  $x = 1 - 1/a$  aber das Kriterium für Werte von  $a$  im Bereich  $1 < a < 3$ .

Im Fall  $a < 1$  erfüllt der Fixpunkt  $x = 0$  das Kriterium (3.8) für einen Attraktor.

Was geschieht mit der Folge, die durch die logistische Gleichung definiert ist, wenn der Wert für  $a$  größer als 3 ist? Das Beispiel für  $a = 3.3$  und einem Startwert  $x_0 = 0.2$  ist in Abb. 3.2 dargestellt. Für diesen aber auch für alle anderen Startwerte entwickelt sich die Iterationen nach einigen Startgliedern zu einer Folge, die stets zwischen den Werten  $x = 0.513\dots$  und  $x = 0.794\dots$  hin und her pendelt. Es liegt also ein stabiler Kreislauf vor. Die beiden Fixpunkte dieses Kreislaufes werden berechnet durch

$$x_{n+2} = f_a(x_n) = f_a(f_a(x_n)) = x_n.\tag{3.9}$$

Für den Fall der logistischen Gleichung liefert diese Bedingung für die doppelt iterierte Funktion die Bestimmungsgleichung

$$x = a \left( a(x - x^2) - [a(x - x^2)]^2 \right),$$

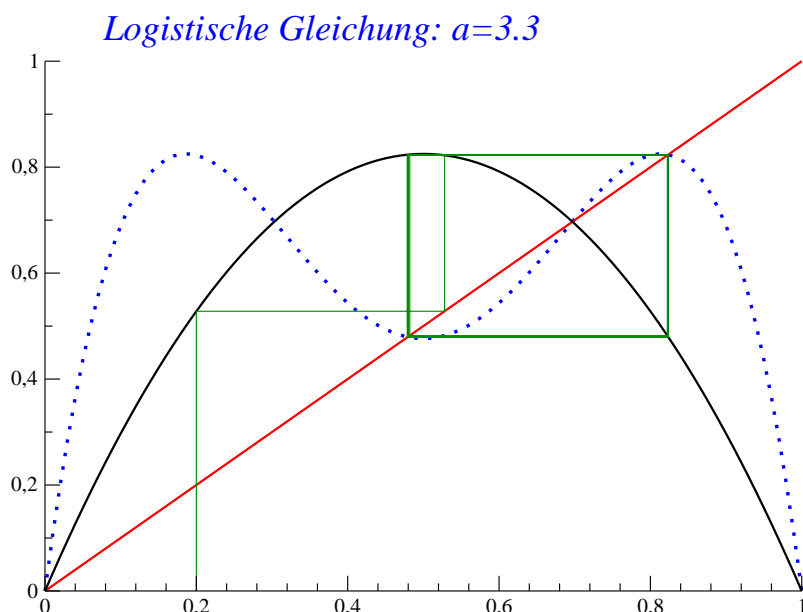


Abbildung 3.2: Grafische Konstruktion der Folge, die durch die logistische Gleichung mit  $a=3.3$  definiert ist, ausgehend von  $x_0=0.2$

beziehungsweise nach der entsprechenden Umformung

$$a^3x^4 - 2a^3x^3 + a^3x^2 + a^2x^2 - a^2x - x = 0. \quad (3.10)$$

Die Nullstellen dieser Gleichung (in Abb. 3.2 grafisch dargestellt durch die Schnittpunkte der Gerad  $y = x$  mit der Funktion  $y = f_a(f_a(x))$ , das ist die blaue gestrichelte Kurve) entsprechen den beiden Lösungen der einfachen Iteration in (3.5) und den Lösungen bei

$$x = \frac{a+1}{2a} \pm \sqrt{\frac{(a-3)(a+1)}{2a}}. \quad (3.11)$$

Diese beiden Punkte sind die Attraktoren im Bereich  $3 < a < 3.45$ . Im Fall der logistischen Gleichung entwickeln sich bei dem Wert des Parameters  $a = 3$  aus einem Attraktor 2 Attraktoren. Man spricht hier von einer **Bifurkation**.

Um das Verhalten der Attraktoren der logistischen Gleichung etwas systematischer darzustellen werden in der Abb. 3.3 als Funktion des Parameters  $a$  Glieder der Folge dargestellt, die sich aus der logistischen Gleichung nach etwa 100 Startschritten ergeben.

Die Ergebnisse im linken Teil der Abb. 3.3 für den Bereich  $2.5 < a < 4$  zeigen die Bifurkation bei  $a = 3$  und weitere Bifurkationen bei  $a = 3.45$ . Für höhere Werte von  $a$  gibt es Bereiche, in denen sich keine Attraktoren ausbilden. In diesem Bereich bildet die logistische Gleichung eine chaotische Folge aus.

Aber auch in diesem chaotischen Bereich bilden sich immer wieder Inseln der Stabilität aus, also Intervalle für den Parameter  $a$  bei denen sich stabile Kreisläufe ausbilden. Als Beispiel sind im rechten Teilbild der Abb. 3.3 die Attraktoren der logistischen Gleichung für ein Intervall in der Nähe des Wertes  $a = 3.74$  dargestellt. In diesem Bereich ergibt sich ein stabiler Kreislauf mit 5 Attraktoren, wie in 3.4 dargestellt. Auch hier findet man

### 3.1.

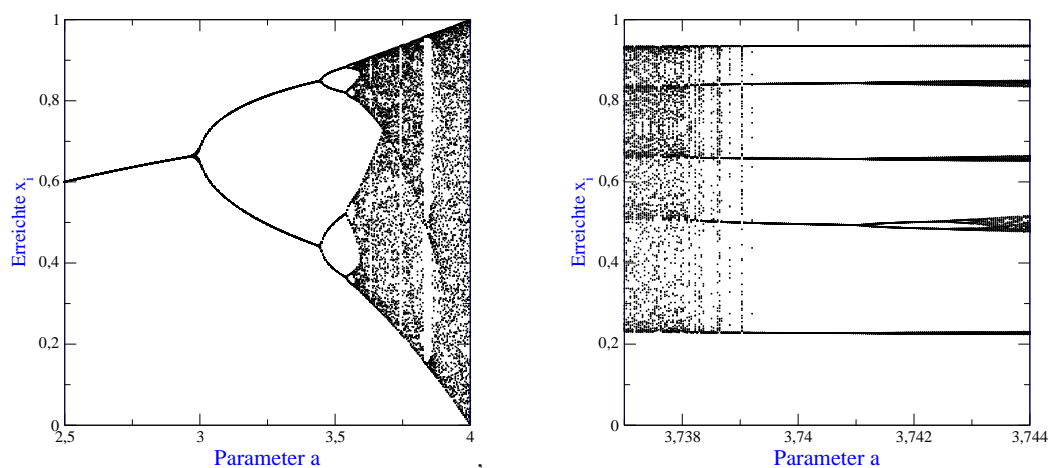


Abbildung 3.3: Darstellung der Attraktoren der logistischen Gleichung für verschiedene Werte des Parameters  $a$ . Glieder der Folge,  $x_n$ ,  $n = 100 \dots 200$  sind jeweils durch einen Punkt dargestellt.

jedoch, dass durch weiteres Anwachsen des Parameters  $a$  aus einzelnen Attraktoren durch Bifurkation weitere Paare von Attraktoren entstehen.

Zum Schluss dieses Abschnittes sollen einige Charakteristika der logistischen Gleichung zusammengefasst werden:

- Die logistische Gleichung (3.1) generiert eine Folge von Punkten  $x_n$ , die wegen der Nichtlinearität der Abbildungsfunktion  $f_a(x)$  zu Fixpunkten mit  $x = f_a(x)$  führt.
- Bei diesen Fixpunkten unterscheidet man zwischen Attraktoren, Fixpunkte auf die sich die Iteration zubewegt und nicht-stabile Fixpunkte. Ein Attraktor ist ein Fixpunkt für den auch das Kriterium (3.7) erfüllt ist.
- Durch die Veränderung des Kontrollparameters  $a$  der logistischen Gleichung wird das Verhalten der erzeugten Folge zum Teil drastisch verändert. Insbesondere:
  - ändert sich der Wert der Attraktoren,
  - bilden sich aus einem Attraktor mehrere (Bifurkationen),
  - verschwinden die Attraktoren, es bilden sich Bereiche aus, in denen die Folge chaotisch wird.
  - Auch in diesem chaotischen Bereich bilden sich wieder Inseln der Stabilität aus, in denen Attraktoren wieder Bifurkationen zeigen. Man spricht von einer Selbstähnlichkeit des Verhaltens auf verschiedene Skalen.

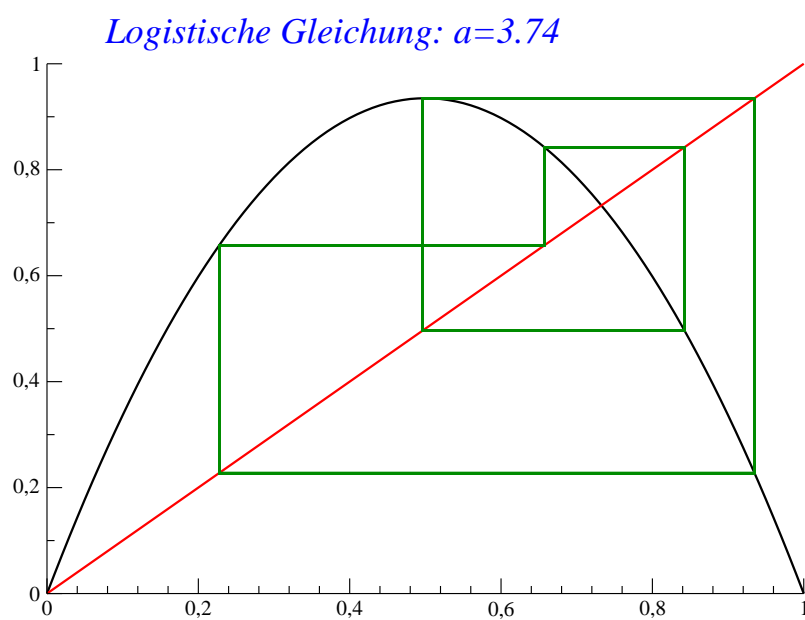


Abbildung 3.4: Grafische Konstruktion der Folge, die durch die logistische Gleichung mit  $a=3.74$  definiert ist.

Abbildung 3.5: *Bewegung einer Kugel in einer Doppelmulde*

## 3.2 Attraktoren in physikalischen Systemen

Wie kann man die Begriffe, die im vorhergehenden Abschnitt diskutiert wurden, auf physikalische Systeme übertragen. Was sind kritische Punkte, was sind Attraktoren bei der zeitlichen Entwicklung eines physikalischen Systems? Als ein einfaches Beispiel betrachten wir eine Kugel, die in einer Raumrichtung in einer Doppelmulde, wie in Abb. 3.5 skizziert, hin und her rollen kann.

Das Experiment zeigt, dass die Kugel einige Male hin und her rollt und schliesslich entweder in der linken oder der rechten Mulde zur Ruhe kommt. Diese Ruhelagen mit der Koordinate  $x = x_L$  oder  $x = x_R$  und jeweils einem Impuls  $p = 0$  sind offensichtlich zwei Attraktoren für die Trajektorien, die die Bewegungen dieses Systems im 2-dimensionalen  $(x, p)$  Phasenraum beschreiben. Je nachdem wie wir das System starten, werden wir über kurz oder lang zu einem dieser Attraktoren gelangen. Wenn die Energie der Kugel zum Start groß genug ist, dass die Kugel über das Maximum im Zentrum hinweg rollen kann, dann hängt der erreichte Endpunkt sehr genau von der Startbedingung ab. Ist die Energie kleiner als die potenzielle Energie im zentralen Maximum, dann wird die Kugel in der Mulde landen, in der sie gestartet ist. Wir haben also hier eine Bifurkation der Attraktoren als Funktion der Energie. Außerdem hat das System noch einen kritischen Punkt bei  $x = x_M$ , der Position des zentralen Maximums, und  $p = 0$ . Wird das System genau an diesem kritischen Punkt gestartet, so verbleibt das System dort. Dieser Zustand ist aber labil und ein kritischer Punkt, der kein Attraktor ist.

Die potenzielle Energie dieser Kugel in der Doppelmulde im Schwerfeld der Erde ist proportional zum Wert der Funktion  $f(x)$ , die diese Doppelmulde darstellt. Wenn wir die generalisierte Koordinate des Systems mit  $q = x$  bezeichnen, so wäre also

$$V(q) = aq^4 + bq^2, \quad (3.12)$$

mit geeigneten Parameter  $a$  (positiv) und  $b$  (negativ) eine Parametrisierung des Potentials der Doppelmulde. Die Hamiltonfunktion lautet also

$$H = \frac{p^2}{2m} + aq^4 + bq^2,$$

und die Bewegungsgleichungen des Systems sind gegeben durch

$$\dot{q} = \frac{\partial H}{\partial p} = \frac{p}{m}$$

$$\dot{p} = -\frac{\partial H}{\partial q} + R(p) = -4aq^3 - 2bq - \alpha p. \quad (3.13)$$

Bei der zweiten Gleichung wurde zu der Hamiltonschen Gleichung für die Bewegung in dem konservativen Kraftfeld, das durch  $V(q)$  beschrieben wird noch ein Term für die Reibungskraft  $R(p) = -\alpha p$  hinzugefügt, der die Verlangsamung der Kugel durch die Reibung beschreibt. Zur Beschreibung der Trajektorie im Phasenraum zerlegen wir die Zeit der Bewegung in diskrete Zeitintervalle  $\Delta t$  und berechnen die Position  $q_i$  und  $p_i$  zur Zeit  $t = i\Delta t$ .

$$q_i = q(t_i) = q(i\Delta t) \quad \text{und} \quad p_i = p(t_i) = p(i\Delta t). \quad (3.14)$$

Die einzelnen Punkte dieser Trajektorie erhalten wir dann für hinreichend kleine Werte des Zeitintervalls  $\Delta t$  durch

$$\begin{aligned} q_{i+1} &= q_i + \dot{q}(t_i)\Delta t = q_i + \Delta t \frac{p_i}{m} \\ p_{i+1} &= p_i + \dot{p}(t_i)\Delta t = p_i + \Delta t (-4aq_i^3 - 2bq_i - \alpha p_i). \end{aligned} \quad (3.15)$$

Dies entspricht dem einfachen Euler Verfahren zur Lösung der 2 gekoppelten Differenzialgleichungen für  $q(t)$  und  $p(t)$ . Nach dem gleichen Prinzip, anwendbar allerdings auch für größere Zeitschritte  $\Delta t$ , funktionieren auch die numerisch etwas stabileren Integrationsmethoden etwa nach dem Runge-Kutta Schema. Als einfachste der Runge-Kutta Methoden betrachten wir hier das Verfahren zweiter Ordnung. Dabei berechnet man die Zeitableitungen  $\dot{q}$  und  $\dot{p}$  nicht wie im Euler Verfahren am Startpunkt des Zeitintervalls  $[t_i, t_{i+1}]$ , also  $t_i$ , sondern in der Mitte des Intervalls, also zur Zeit

$$t_{i+} = t_i + \frac{1}{2}\Delta t.$$

Zur Berechnung der Zeitableitungen an dieser Zwischenzeit  $t_{i+}$  benötigt man dann aber auch eine Abschätzung für die Position  $q_{i+}$  und den Impuls  $p_{i+}$  zu dieser Zwischenzeit. Diese lassen sich als Mittelwert zwischen den Werten zur Zeit  $t_i$  und  $t_{i+1}$  bestimmt nach dem Eulerverfahren abschätzen und man erhält

$$\begin{aligned} q_{i+} &= q_i + \frac{1}{2}\dot{q}(q_i, p_i, t_i)\Delta t \\ p_{i+} &= p_i + \frac{1}{2}\dot{p}(q_i, p_i, t_i)\Delta t. \end{aligned}$$

Damit ergibt sich dann

$$\begin{aligned} q_{i+1} &= q_i + \dot{q}(q_{i+}, p_{i+}, t_{i+})\Delta t \\ p_{i+1} &= p_i + \dot{p}(q_{i+}, p_{i+}, t_{i+})\Delta t. \end{aligned}$$

Während die Fehlerabschätzung des Eulerverfahrens einen Fehler der Ordnung  $\Delta t^2$  ergibt ist bei diesem einfachsten Runge-Kutta Verfahren der Fehler proportional zu  $\Delta t^3$ . Etwas aufwendiger sind Runge-Kutta Methoden höherer Ordnung.

Eine Implementation einer Runge-Kutta Integration von gekoppelten Differenzialgleichungen findet man z.B. im Software Paket **MATLAB**, das für dieses Beispiel kurz vorgestellt werden soll.

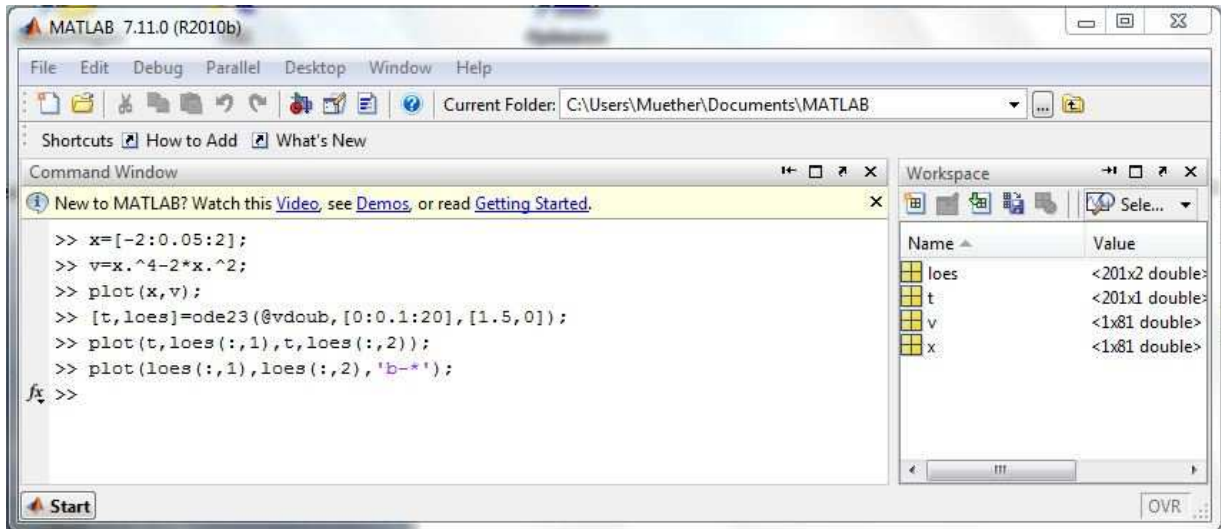


Abbildung 3.6: Beispiel einer MATLAB Sitzung, siehe Diskussion im Text

Ein Beispiel für eine kurze MATLAB Sitzung ist in Abb. 3.6 dargestellt. Nach Eröffnung der Sitzung werden die Anweisungen im Fenster “Command Window” eingegeben. Durch das Kommando

$$x = [-2 : 0.05 : 2];$$

wird ein Zeilenvektor mit Namen  $x$  erzeugt, dessen Komponenten die Werte von  $-2$  bis  $+2$  in Schritten von  $0.05$  enthält, und damit einen Vektor mit  $81$  Elementen darstellt. Auf das  $5$ . Element kann mit  $x(5)$  zugreifen und sollte dort den Wert  $-1.8$  vorfinden. Diesem Vektor soll nun ein Vektor  $v$  zugeordnet werden, der jeweils die Werte des Potentials  $v(x) = x^4 - 2 * x^2$  enthält. Dies erreicht man mit dem Statement

$$v = x.^4 - 2 * x.^2;$$

Zu beachten ist dabei die Operation  $.$ , die zu unterscheiden ist von der Operation  $^$ . Während mit  $x.^2$  oder auch  $x.*x$  die einzelnen Elemente von  $x$  quadriert werden und so wieder ein Vektor der gleichen Dimension erzeugt wird, würde  $x*x$  einer Matrixmultiplikation entsprechen, die aber nicht funktioniert, da die Zeilendimension und die Spaltendimension von  $x$  nicht übereinstimmen. Hier wäre  $x*transpose(x)$  denkbar, wobei die Funktion  $transpose$  die transponierte Matrix erzeugt, also hier einen Spaltenvektor. Damit wird  $x*transpose(x)$  also gerade das Skalarprodukt von  $x$  mit sich selbst generieren.

Das `plot` Kommando in der nächsten Zeile von Figur 3.6 sorgt dafür, dass die Punkte, deren horizontale Koordinate durch das jeweilige Element  $x(i)$  und vertikale Koordinate durch das entsprechende  $y(i)$  gegeben sind, dargestellt werden und durch eine Linie miteinander verbunden sind. `plot(x,v)` plottet also genau die Funktion  $v(x)$ , dargestellt durch die einzelnen Stützstellen.

Die Lösung des Systems von 2 gekoppelten Differenzialgleichungen vom Typ (3.13) erfolgt durch die nächste Zeile

$$[t, loes] = ode23(@vdoub, [0 : 0.1 : 20], [1.5, 0]); \quad (3.16)$$

```

1 function [ ydot ] = vdoub(t, y )
2 %Outputvector enthält die Zeitableitungen: ydot(1)=dq/dt; ydot(2)=dp/dt
3 %Input ist die Zeitvariable t und die Werte y(1)=q(t); y(2)=p(t)
4 %hier für die gedämpfte Schwingung (Dämpfungskonstante alpha) im
5 %Potenzial V(x) = ax^4 -bx^2 mit $m=1$
6 alpha=0.1;
7 a=1;
8 b=1;
9 ydot=[0;0];
10 ydot(1)=y(2);
11 ydot(2)=-4*a*y(1)^3+2*b*y(1)-alpha*y(2);
12 end
13
2 usages of "alpha" found
vdoub Ln 6 Col 11 OVR

```

Abbildung 3.7: Beispiel einer MATLAB Funktion, siehe Diskussion im Text

Die Lösung erfolgt durch den Aufruf der Prozedur `ode23`. MATLAB bietet neben dieser Prozedur ein ganzes Spektrum von Alternativen an. Wir wollen zunächst die Parameter dieser Prozedur betrachten. Das erste Argument bezieht sich auf den Namen der Funktion, in der die rechten Seiten der Differentialgleichungen spezifiziert sind. In unserem Beispiel trägt diese Funktion den Namen “`vdoub`” und wird deshalb mit dem Argument “`@vdoub`” adressiert.

Zur Definition dieser Funktion wird der Editor von Matlab aufgerufen und benutzt, um die Funktion z.B. durch die Darstellung in Abb. 3.7 zu definieren. Bevor wir jedoch auf die Definition dieser Funktion eingehen, wollen wir den Aufruf von `ode23` in (3.16) weiter beschreiben. Das zweite Argument in `ode23`, `[0 : 0.1 : 20]` definiert das Intervall von Stützstellen in der Zeit  $t$  für die die Lösung gefunden werden soll: In diesem Fall also 201 Stützstellen von 0 bis 20 im Abstand von jeweils 0.1. Das dritte Argument, der Vektor `[1.5, 0]` identifiziert die Randbedingung, also die Startwerte für die Lösung zur Zeit  $t = 0$ . In unserem Beispiel:

$$q(0) = 1.5 \quad \text{und} \quad p(0) = 0.$$

Das Ergebnis der Operation ist durch die linke Seite des Statements (3.16) identifiziert. Das erste Element von `[t, loes]`, also  $t$  bezeichnet den Vektor der Stützstellen in der Zeit  $t$ , das zweite Element ist eine Matrix mit 2 Zeilen und 201 Spalten. Die erste Zeile enthält den Lösungsvektor  $q(i)$ , während die zweite Zeile die zugehörigen  $p(i)$  beinhaltet. Die jeweilige Zeile wird dann als Vektor durch `loes(:, 1)`, bzw. `(loes(:, 2))` adressiert. Damit führt also

$$\text{plot}(t, \text{loes}(:, 1));$$

zu einer Darstellung von  $q(t)$ . Mit dem entsprechenden Kommando `plot(t, loes(:, 1), t, loes(:, 2))`; werden parallel  $q(t)$  und  $p(t)$  dargestellt (siehe oberer Teil von Abb.3.8). Die Darstellung der Trajektorie im Phasenraum im unteren Teil von Abb.3.8) erfolgt durch das Kommando `plot(loes(:, 1), loes(:, 2))`; Die zusätzliche Angabe `'b - *` dient zur Spezifizierung der Darstellung und sorgt hier dafür, dass die einzelnen Stützstellen durch das Symbol `*` herausgehoben sind.



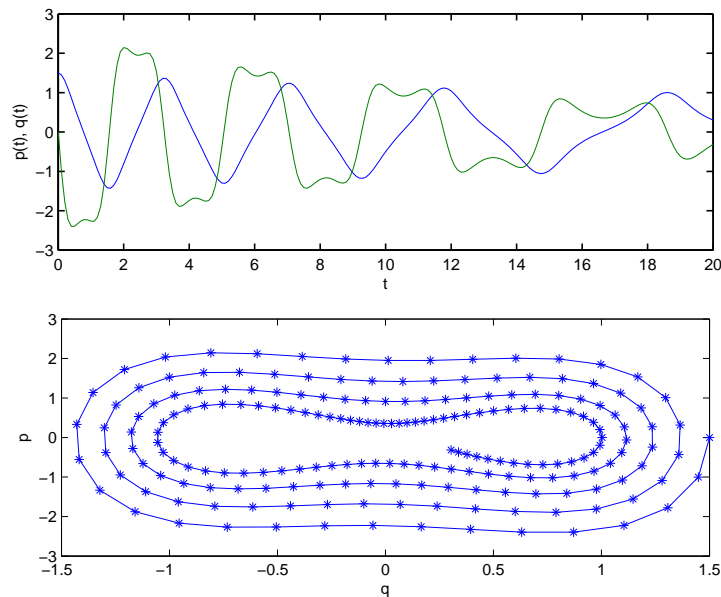


Abbildung 3.8: Lösung der Differenzialgleichung (3.13), siehe Diskussion im Text

Zur Abrundung seien noch ein paar Kommentare zur Definition der Funktion `vdoub` im Editor von MATLAB in der Abb. 3.7 gegeben. Im ersten Statement dem Funktionsstatement, sind neben dem Namen der Funktion die Inputparameter der Funktion,  $t$  und  $y$ , und der Output, die dazugehörigen Zeitableitungen,  $ydot$ , identifiziert. Dabei bezieht sich in unserem Beispiel  $t$  auf einen aktuellen Wert der Zeit, während  $y(1)$  den zugehörigen Ort  $q(t)$  und  $y(2)$  den entsprechenden Impuls  $p(t)$  bezeichnet. Die Zeilen beginnend mit dem Prozentzeichen sind reine Kommentare und haben keinen Einfluss auf den Ablauf der Rechnung. Die eigentliche Funktionszuweisung erfolgt in den Zeilen 10 und 11. Diese Funktion oder eine entsprechende muss im Editor eingegeben werden und als File `vdoub.m` im aktuellen Verzeichnis gespeichert sein.

Nach diesem kleinen Ausflug in die Anwendung eines Hilfsmittels zur numerischen Lösung von Bewegungsgleichungen kehren wir zur allgemeinen Diskussion zurück.

Für die weitere Diskussion greifen wir die Integration der Bewegungsgleichung nach (3.15) noch einmal auf und schreiben für die diskreten Punkte der Trajektorie im Phasenraum

$$\begin{aligned} x_{i+1}^1 &= q_{i+1} = q_i + \dot{q}_i \Delta t \\ x_{i+1}^2 &= p_{i+1} = p_i + \dot{p}_i \Delta t. \end{aligned} \quad (3.17)$$

Dabei haben wir die generalisierte Koordinate und die Impulse zur Zeit  $(i+1)\Delta t$  zu einem Vektor  $\vec{x}_{i+1}$  zusammengefasst. Dieser Vektor hat im obigen Beispiel die Dimension 2, im allgemeinen Fall, bei einem System mit  $s$  generalisierten Koordinaten die Dimension  $2s$ . Die Zeitableitungen  $\dot{q}_i$  und  $\dot{p}_i$  sind über die Bewegungsgleichungen wie z.B. (3.13) mit den Werten von  $q_i$  und  $p_i$  verknüpft. Die Gleichungen (3.17) nehmen also die Form an

$$x_{i+1}^\alpha = f^\alpha(x_i^1, \dots, x_i^{2s}) = x_i^\alpha + F^\alpha(x_i^1, \dots, x_i^{2s}) \Delta t. \quad (3.18)$$

Dabei sind die Funktionen  $F^\alpha$  über die Bewegungsgleichungen mit den Zeitableitungen

der entsprechenden Komponenten verknüpft

$$\dot{x}_i^\alpha = F^\alpha(x_i^1, \dots, x_i^{2s})\Delta t, \quad (3.19)$$

oder anders dargestellt für die  $2s$ -dimensionalen Vektoren

$$\dot{\vec{x}}_i = \vec{F}(\vec{x}_i).$$

Ein kritischer Punkt oder Fixpunkt für die Folge der  $\vec{x}_i$  ergibt sich genau für solche Werte von  $\vec{x}$  für die gilt

$$\dot{\vec{x}} = \vec{F} = \vec{0}, \quad (3.20)$$

also alle Funktionen  $F^\alpha(\vec{x}) = 0$  erfüllen.

Wann ist aber ein solcher Fixpunkt, sagen wir bei  $\vec{x} = \vec{x}_0$  ein Attraktor? Dazu sehen wir uns die Ergebnisse für  $F^\alpha$  für Werte von  $\vec{x}$  in der Umgebung des Fixpunktes  $\vec{x}_0$  an und entwickeln in einer Taylorreihe

$$F^\alpha(\vec{x}) = \underbrace{F^\alpha(\vec{x}_0)}_{=0, \text{wg. (3.20)}} + \sum_{\beta=1}^{2s} \frac{\partial F^\alpha}{\partial x^\beta} (x^\beta - x_0^\beta) + \dots \quad (3.21)$$

wobei wir in unmittelbarer Umgebung von  $\vec{x}_0$  die Beiträge höherer Ordnung in der Entwicklung vernachlässigen können. Die Ableitungen

$$A_{\alpha\beta} = \frac{\partial F^\alpha}{\partial x^\beta} \quad (3.22)$$

können zu einer Matrix der Dimension  $2s$  zusammengefasst werden. Die Eigenwerte dieser Matrix, heissen aus Gründen, die weiter unten deutlich werden, die **charakteristischen Exponenten** des Vektorfeldes  $\vec{F}$  am Punkt  $\vec{x}_0$ , also dem Punkt an dem diese Ableitungen bestimmt werden.

Wir wollen uns nun davon überzeugen, dass ein kritischer Punkt  $\vec{x}_0$  genau dann ein Attraktor ist, wenn alle charakteristischen Exponenten von  $\vec{F}$  am Punkt  $\vec{x}_0$  einen negativen Realteil haben. Anders ausgedrückt: Alle Eigenwerte der Matrix  $A$  müssen einen negativen Realteil besitzen.

Zum Beweis dieser Behauptung schreiben wir für die zeitliche Ableitung des Vektor  $\vec{x}$  in der Nähe des kritischen Punktes mit (3.21)

$$\dot{\vec{x}} = A(\vec{x} - \vec{x}_0). \quad (3.23)$$

Wäre diese Gleichung keine Vektorgleichung mit einer Matrix  $A$  sondern lediglich eine Gleichung in einer Dimension, so wäre sofort klar, dass

$$\vec{x}(t) - \vec{x}_0 = \exp((t - t_0)A)(\vec{x}(t_0) - \vec{x}_0), \quad (3.24)$$

die Lösung angibt, wie sich der Wert  $x(t)$  aus dem Anfangswert  $x(t_0)$  entwickelt. Diese Lösung gilt aber auch für Vektoren  $\vec{x}$  und entsprechende Matrizen  $A$ . In diesem Fall müssen wir aber darauf achten, dass die Exponentialfunktion mit einer Matrix als Argument durch ihre Reihe definiert ist und die Vektoren und Matrizen auf der rechten Seite

der Gleichung genau in der Reihenfolge geschrieben werden, wie dies in (3.24) geschehen ist.

Zur weiteren Beweisführung nehmen wir an, dass  $\vec{x}(t_0) - \vec{x}_0$  ein Eigenvektor der Matrix  $A$  sei mit dem komplexen Eigenwert  $\lambda = c + id$ . Die zeitliche Entwicklung dieses Eigenvektors ergibt sich dadurch, dass wir in (3.24) die Matrix  $A$  durch den Eigenwert  $\lambda$  ersetzen. Wir können also schreiben

$$\begin{aligned}\vec{x}(t) - \vec{x}_0 &= \exp((t - t_0)\lambda)(\vec{x}(t_0) - \vec{x}_0) \\ &= \exp((t - t_0)c) \exp(i(t - t_0)d)(\vec{x}(t_0) - \vec{x}_0).\end{aligned}\quad (3.25)$$

Die zweite Exponentialfunktion in der zweiten Zeile liefert ein oszillierendes Zeitverhalten und ist deshalb für die Entwicklung für große Zeiten  $t - t_0$  nicht so relevant. Ist aber  $c$ , der Realteil des Eigenwertes  $\lambda$  negativ, so wird für große Werte von  $t - t_0$  die Differenz  $\vec{x}(t) - \vec{x}_0$  gegen Null konvergieren. Die Trajektorie läuft auf den kritischen Punkt  $\vec{x}_0$  zu. Sind die Realteile aller Eigenwerte negativ, so gilt das für alle Eigenvektoren der Matrix  $A$  und damit für beliebige Vektoren, da sich diese ja als Linearkombination der Eigenvektoren darstellen lassen. Damit ist die Behauptung bewiesen.

Wir wollen diese Kriterien am Beispiel der Bewegungsgleichungen (3.13) mit  $a = 1/4$  und  $b = -1/2$ , also einem Beispiel der Doppelmulde demonstrieren. Die Bedingungen für die Fixpunkt (3.20) lauten in diesem Fall

$$\begin{aligned}\frac{p}{m} &= 0 \\ -q^3 + q - \alpha p &= 0\end{aligned}\quad (3.26)$$

was bedeutet, dass es 3 kritische Punkte gibt, jeweils mit  $p = 0$  und

$$q_1 = 0, \quad q_2 = 1, \quad \text{und} \quad q_3 = -1.$$

Die charakteristische Matrix hat die Gestalt

$$A = \begin{pmatrix} \frac{\partial F^1}{\partial q} & \frac{\partial F^1}{\partial p} \\ \frac{\partial F^2}{\partial q} & \frac{\partial F^2}{\partial p} \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{m} \\ -3q^2 + 1 & -\alpha \end{pmatrix}.$$

Man kann sich nun davon überzeugen, dass diese Matrix nur negative Eigenwerte hat für  $q_2 = 1$  und  $q_3 = -1$ , also den beiden Mulden, während bei  $q_1 = 0$  sich ein positiver Eigenwert für  $A$  ergibt, was natürlich besagt, dass der kritische Punkt bei  $q = 0$  kein Attraktor ist.

Für eindimensionale Probleme ist es sehr einfach die Trajektorie im 2-dimensionalen Phasenraum grafisch darzustellen. Dies wird aber sehr viel schwieriger, wenn man Phasenraumtrajektorien in höheren Dimensionen visualisieren will. Dabei hilft häufig die Technik des **Poincaré Schnittes**, die wir hier am Beispiel einer Bewegung in 2 Dimensionen demonstrieren wollen. Konkret betrachten wir die Bewegung in einer Potenzialfläche, die in den generalisierten Koordinaten  $q_1 = x$  und  $q_2 = y$  definiert sein soll in der Form

$$V(x, y) = \frac{1}{2}(x^2 + y^2) + x^2y - \frac{1}{3}y^3. \quad (3.27)$$

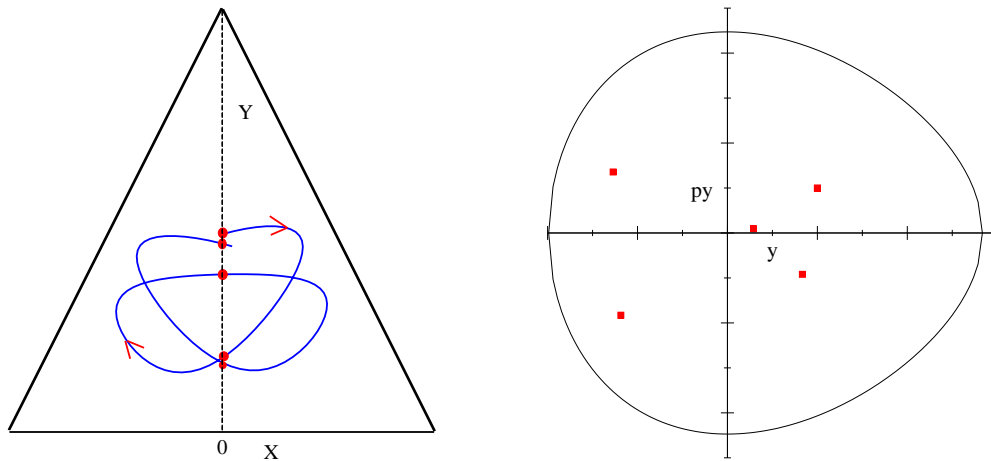


Abbildung 3.9: *Poincare Schnitt am Beispiel der Bewegung einer Masse in einem Henon Hiles Potenzial. Im linken Teilbild ist die Trajektorie in der  $x$ - $y$  Ebene dargestellt. Jedes Mal wenn die  $x = 0$  Achse geschnitten wird, werden im rechten Teilbild die Werte für  $y$  und  $p_y$  als Punkt dargestellt. Der Bereich des  $y$ - $p_y$  Phasenraumes, der bei der vorgegebenen Energie (hier  $E = 0.1$ ) erreicht werden kann ist im rechten Teilbild durch eine einhüllende Kurve dargestellt.*

Dies ist ein Potenzialmulde in der Umgebung des Koordinatenursprungs. Ein Teilchen mit der Masse  $m = 1$  wird in der Potenzialmulde verbleiben, wenn seine Energie kleiner als 0.16666 ist.

Ziel ist es, die Trajektorien dieser Bewegung der gebundenen Masse im Phasenraum darzustellen.

Da es sich bei dieser Bewegung um eine Bewegung in einem konservativen Kraftfeld handelt, reduziert sich bei vorgegebener Energie  $E$  die Zahl der Freiheitsgrade im Phasenraum von 4 auf 3. Sind 3 der Koordinaten,  $x$ ,  $y$ ,  $p_x$  und  $p_y$  bekannt, so ergibt sich der Wert für die 4. Koordinate durch die Energie. Die vorgegebene Energie  $E$  reduziert die möglichen Phasenraumpunkte auf eine 3-dimensionale Mannigfaltigkeit im 4-dimensionalen Phasenraum.

Die Technik des Poincare Schnittes besagt in dieser Situation, dass man die Trajektorie des Teilchens verfolgt, bis man an eine Stelle gelangt ist, bei der z.B. der Wert der  $x$  Koordinate identisch null ist. (Natürlich könnte man auch die  $y$ -Koordinate heranziehen oder einen anderen Wert auswählen, den diese Koordinate einnimmt.) Dies ist im linken Teilbild von Abb. 3.9 dargestellt.

Jedes Mal wenn diese Trajektorie die  $x = 0$  Achse schneidet (roter Punkt im linken Teilbild von Abb. 3.9), wird im rechten Teilbild durch einen Punkt der aktuelle Wert für  $y$  und  $p_y$  dargestellt. Eine solche Darstellung, wie sie sich im rechten Teil der Abb. 3.9 findet, nennt man Poincare Schnitt. Liest man aus einem Punkt des Poincare Schnittes in unserem Beispiel die Werte für  $y$  und  $p_y$  ab, so weiss man außerdem wegen der Konstruktionsvereinbarung, dass  $x = 0$  ist, und kann dann über die Ausnutzung der Energie alle 4 Koordinaten des aktuellen Punktes im Phasenraum konstruieren. Ein solcher Poincare Schnitt liefert also kein kontinuierliches Bild für die Trajektorie sondern für eine Folge von gewissen Zeitpunkten eine vollständige Information über diese Phasenraumtrajektorie.

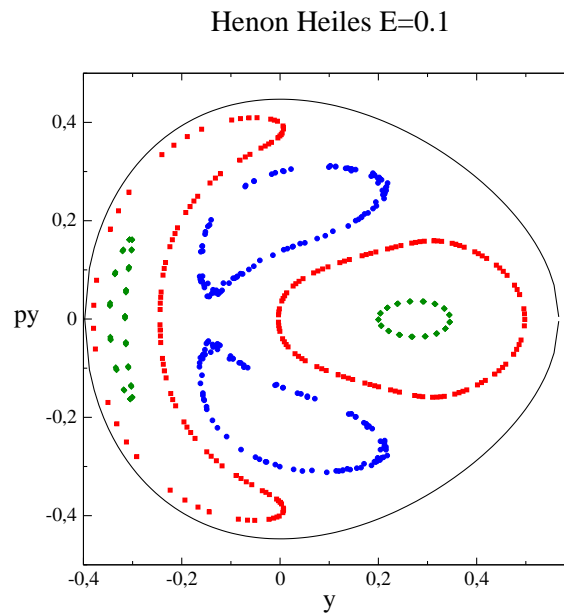


Abbildung 3.10: Poincare Schnitt für dasn Henon Hiles Potenzial bei  $E = 0.1$ . Die Trajektorien für verschiedene Anfangsbedingungen sind durch unterschiedliche Farben dargestellt. Weitere Information siehe Abb. 3.9

Wir haben hier nur ein Beispiel für die Anwendung des Poincare Schnittes erläutert. Natürlich kann man dieses Verfahren auch durch eine andere Auswahl der Konstruktionsbedingung und der dargestellten Koordinaten nutzen. Es kann auch für Probleme mit mehr als 2 Freiheitsgraden benutzt werden. In diesem Fall wird aber natürlich die Zahl der dargestellten Informationsfreiheitsgrade im Vergleich mit den vorhandenen Freiheitsgrade immer geringer.

In der Abb. 3.10 sind die Trajektorien für die Bewegung im Henon Hiles Potenzial bei einer Energie  $E = 0.1$  dargestellt. Die Trajektorien, die sich aus unterschiedlichen Anfangsbedingungen ergeben, sind durch Punkte in unterschiedlichen Farben dargestellt. Für jede Anfangsbedingung sind 200 Punkte visualisiert. Man sieht, dass die Trajektorien im Poincare Schnitt wohl definierte Figuren bilden. Bei einer vorgegebenen Startbedingung ist also offensichtlich nur ein wohl definierter Unterraum des Phasenraumes erreichbar, der geometrisch klar definiert ist.

Die Situation ändert sich, wenn man Bewegungen im Henon Hiles Potenzial bei höherer Energie betrachtet. Zwar gibt es bei der Darstellung in Abb. 3.11 auch Startbedingungen, bei denen die Trajektorie in einem klar umrissenen Bereich bleiben. Bei anderen Startbedingungen füllen die Punkte im Poincare Schnitt aber ganze Flächen aus, die sich für verschiedene Startbedingungen durchmischen. In diesem Fall wird der erreichbare Teil des Phasenraumes geometrisch sehr kompliziert. Entsprechend unsicher sind mögliche Vorhersagen, wie sich das System für lange Zeiten entwickeln wird. Das System ist in einem chaotischen Bereich.

Dieses Beispiel soll zeigen, wie man auch aus der reduzierten visuellen Information eines Poincare Schnittes, charakteristische Eigenschaften von Phasenraumtrajektorien und damit der dadurch beschriebenen Systeme ablesen kann.

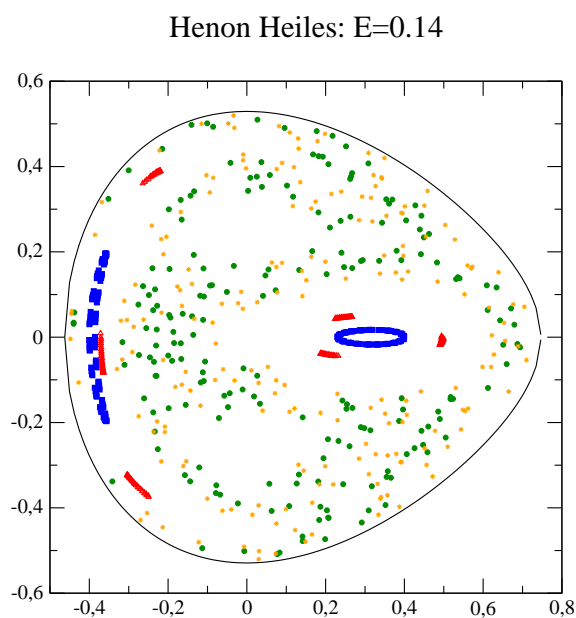


Abbildung 3.11: *Poincaré Schnitt für das Henon Heiles Potenzial bei  $E = 0.14$ . Die Trajektorien für verschiedene Anfangsbedingungen sind durch unterschiedliche Farben dargestellt. Weitere Information siehe Abb. 3.9*

### 3.3 Der Satz von Liouville

In diesem Abschnitt wird der sogenannte Volumensatz von Liouville diskutiert. Er gibt für Systeme, die mit Hamiltonschen Bewegungsgleichungen beschrieben werden (diese nennt man auch **Hamiltonsche Systeme**), einen gewissen Rahmen vor, wie weit sich Trajektorien, die zu einem Startzeitpunkt  $t_0$  einen gewissen Abstand haben, während der zeitlichen Entwicklung des Systems voneinander entfernen können.

Als eine Art Vorbemerkung soll aber zunächst festgehalten werden, dass für ein solches Hamiltonsches System zwei Trajektorien im Phasenraum sich nicht schneiden können. Für ein Hamiltonsches System ist ja durch die Angabe eines Punktes im Phasenraum, durch die Angabe der  $2s$  Koordinaten für den Startpunkt des Systems, die Entwicklung des Systems, die man ja mathematisch durch die Lösung der Bewegungsgleichungen bestimmen kann, eindeutig festgelegt. Durch die Angabe des Startpunktes liegt also die gesamte Trajektorie fest. Käme es nun zu einem Schnittpunkt zwischen zwei Trajektorien, so wäre ja dieser Schnittpunkt ein Punkt, der als Startpunkt der weiteren Entwicklung dieser Trajektorien angesehen werden kann. Die beiden Trajektorien würden also eine identische Entwicklung erfahren. Diese Trajektorien schneiden sich also nicht, sondern sind identisch.

Betrachten wir eine solche Trajektorie eines Hamiltonschen Systems im Phasenraum. Zur Zeit  $t_0$  definieren wir eine infinitesimale Umgebung um den Punkt den diese Trajektorie zur Zeit  $t_0$  einnimmt. Verfolgt man die Trajektorien, die den Startpunkt in der erwähnten Umgebung des Startpunktes der Referenztrajektorie besitzen, so ändert sich das Volumen, das diese Trajektorien im Phasenraum einnehmen, während der zeitlichen Entwicklung nicht. Dies ist die Aussage des Satzes von Liouville, den wir im folgenden beweisen werden.

Da sich, wie bereits erwähnt, Phasenraumtrajektorien nicht schneiden, bedeutet dies, dass keine Trajektorien aus diesem Volumen herauslaufen und keine in das Volumen eindringen werden, sie müssten sich dann ja mit Trajektorien auf dem Rand des Volumens schneiden. Die Dichte der Punkte im Phasenraum in der Umgebung eines Punktes, der sich den Bewegungsgleichungen entsprechend bewegt, ändert sich nicht. Wie wir noch diskutieren werden, kann sich dieses Volumen während der Entwicklung des Systems ändern, so dass sich der Abstand zwischen 2 Trajektorien bezüglich einer Koordinate des Phasenraumes stark ändert. Dies geht aber nur unter Beibehaltung des Volumens.

Zum Beweis des Liouville Theorems betrachten wir die diskutierte Phasenraumdicke und die totale Ableitung dieser Phasenraumdicke  $\rho$  nach der Zeit  $t$

$$\frac{d\rho}{dt} = \frac{\partial\rho}{\partial t} + \sum_{k=1}^s \frac{\partial\rho}{\partial q_k} \dot{q}_k + \frac{\partial\rho}{\partial p_k} \dot{p}_k. \quad (3.28)$$

Dabei beschreibt die partielle Ableitung von  $\rho$  nach der Zeit, die Änderung der Dichte, die sich ergibt, wenn man das Volumen, für das die Dichte berechnet wird, nicht mit der Entwicklung der Koordinaten  $q_k$  und  $p_k$  mitbewegt. Die weiteren Terme auf der rechten Seite der Gleichung (3.28) beschreiben die Beiträge zur Dichteänderung, die sich dadurch ergeben, dass sich das Volumen mit der Referenztrajektorie mitbewegen soll.

Wir bezeichnen die Koordinaten und Impulse der Referenztrajektorie mit  $q_{k0}(t)$ , bzw.  $p_{k0}(t)$ . In der Abb. 3.12 ist ein infinitesimales Volumen um diesen Referenzpunkt dargestellt. Die partielle Ableitung ergibt sich aus der Summe der Ströme durch die 4 Flächen, die in der Abbildung mit Ziffern I bis IV bezeichnet sind.

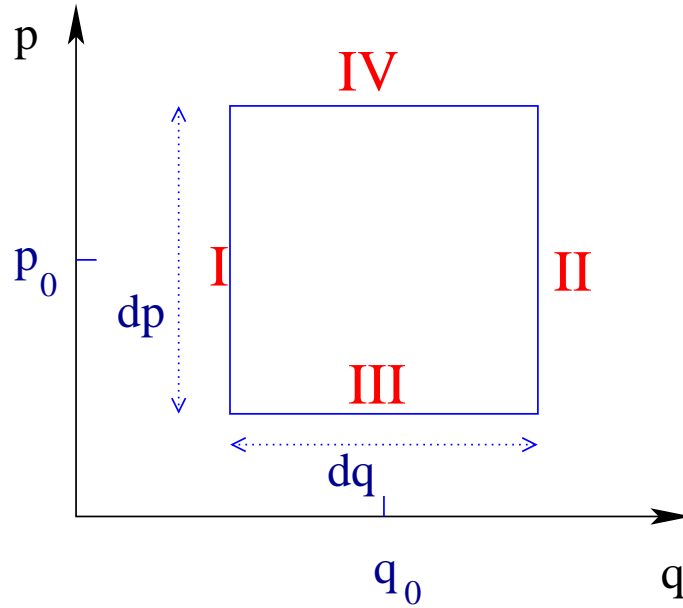


Abbildung 3.12: Volumen in einer 2-dimensionalen Projektion des Phasenraumes

Der Strom durch die Fläche I berechnet sich zu

$$\rho(I)\dot{q}_k(I) dp_k dF_k. \quad (3.29)$$

Dies ist das Produkt aus der Stromdichte an der Position der Fläche I mal der Größe der Fläche I. Die Stromdichte ist das Produkt aus der Dichte der Phasenraumpunkte  $\rho(I)$  multipliziert mit der Geschwindigkeit dieser Phasenraumpunkte senkrecht zur Oberfläche,  $\dot{q}_k$  an dieser Stelle. Ein positiver Wert für diese Stromdichte bedeutet, dass sich die Zahl der Phasenraumpunkte im Volumen erhöht. Die Fläche I ergibt sich als Produkt aus der Seitenlänge  $dp_k$  multipliziert mit der  $2(s-1)$  dimensionalen Fläche  $dF_k$ , die in der Abb. 3.12 nicht dargestellt werden kann.

Analog ergibt sich für den Strom durch die Fläche II

$$\rho(II)\dot{q}_k(II) dp_k dF_k = \left( \rho(I)\dot{q}_k(I) + \frac{\partial \rho \dot{q}_k}{\partial q_k} dq_k \right) dp_k dF_k. \quad (3.30)$$

Auf der rechten Seite haben wir die Stromdichte an der Stelle II entwickelt um den Punkt der Fläche I und diese Entwicklung nach dem zweiten Glied abgebrochen, da ja das Volumen und damit  $dq_k$  infinitesimal sein soll. Ein positiver Wert für diesen Strom durch die Fläche II entspricht einer Abnahme der Zahl der Phasenraumpunkte im Volumen  $dV$ .

Entsprechend können wir den Strom durch die Flächen III und IV berechnen zu

$$\begin{aligned} III : & \quad \rho(III)\dot{p}_k(III) dq_k dF_k \\ IV : & \quad \left( \rho(III)\dot{p}_k(III) + \frac{\partial \rho \dot{p}_k}{\partial p_k} dp_k \right) dq_k dF_k. \end{aligned} \quad (3.31)$$

Die gesamte Änderung der Phasenraumpunkte im Volumen  $dV$  pro Zeiteinheit, die be-



dingt ist durch einen Fluss der hier betrachteten Flächen, ergibt sich also zu

$$I - II + III - IV \rightarrow - \left( \frac{\partial \rho \dot{q}_k}{\partial q_k} + \frac{\partial \rho \dot{p}_k}{\partial p_k} \right) \underbrace{dq_k dp_k dF_k}_{=dV}. \quad (3.32)$$

Die partielle Ableitung der Dichte  $\rho$  nach der Zeit ergibt sich dadurch, dass man diesen Ausdruck durch das Volumen  $dV$  dividiert. Außerdem darf man natürlich nicht nur den Fluss durch die 4 Flächen betrachten, die in Abb. 3.12 explizit dargestellt und bisher behandelt sind, sondern muss entsprechende Beiträge der anderen  $s - 1$  Freiheitsgrade berücksichtigen. Dies führt uns auf

$$\begin{aligned} \frac{\partial \rho}{\partial t} &= - \sum_{k=1}^s \left( \frac{\partial \rho \dot{q}_k}{\partial q_k} + \frac{\partial \rho \dot{p}_k}{\partial p_k} \right) \\ &= - \sum_{k=1}^s \left( \frac{\partial \rho}{\partial q_k} \dot{q}_k + \rho \frac{\partial \dot{q}_k}{\partial q_k} + \frac{\partial \rho}{\partial p_k} \dot{p}_k + \rho \frac{\partial \dot{p}_k}{\partial p_k} \right). \end{aligned} \quad (3.33)$$

Für ein Hamiltonsches System gelten die Hamiltonschen Bewegungsgleichungen, so dass

$$\frac{\partial \dot{q}_k}{\partial q_k} = \frac{\partial^2 H}{\partial q_k \partial p_k} = \frac{\partial^2 H}{\partial p_k \partial q_k} = - \frac{\partial \dot{p}_k}{\partial p_k}.$$

Damit kompensieren sich die entsprechenden Terme in der zweiten Zeile von (3.33) und der Ausdruck reduziert sich auf

$$\frac{\partial \rho}{\partial t} = - \sum_{k=1}^s \left( \frac{\partial \rho}{\partial q_k} \dot{q}_k + \frac{\partial \rho}{\partial p_k} \dot{p}_k \right). \quad (3.34)$$

Setzt man dieses Ergebnis in (3.28) ein, so erhält man

$$\frac{d\rho}{dt} = 0, \quad (3.35)$$

was ja zum Beweis des Volumensatzes von Liouville zu zeigen war.

Zur weiteren Erläuterung der Bedeutung dieses Satzes wollen wir das Beispiel eines Pendels im Schwerfeld der Erde betrachten. Die generalisierte Koordinate dieses Systems ist z.B. der Winkel  $\varphi$ , den die Pendelstange mit der Richtung der Erdanziehungskraft bildet. Somit ist die potenzielle Energie gegeben durch

$$V = -mgl \cos \varphi = -mgl + \frac{mgl}{2} \varphi^2 - \frac{mgl}{24} \varphi^4 + \dots, \quad (3.36)$$

wobei  $m$  für die Masse des Pendels,  $l$  die Länge des Pendels und  $g$  für die Erdbeschleunigung stehen. Auf der rechten Seite dieser Gleichung haben wir die Taylorentwicklung dieses Potentials  $V$  für das Mathematische Pendel angedeutet. Bei kleinen Auslenkungen  $\varphi$  reduziert sich dieses Potential auf das Potential des harmonischen Oszillators mit einem quadratischem Term im Potential.

Im Grenzfall des harmonischen Oszillators sind die Winkelgeschwindigkeiten

$$\omega = \sqrt{\frac{g}{l}}$$

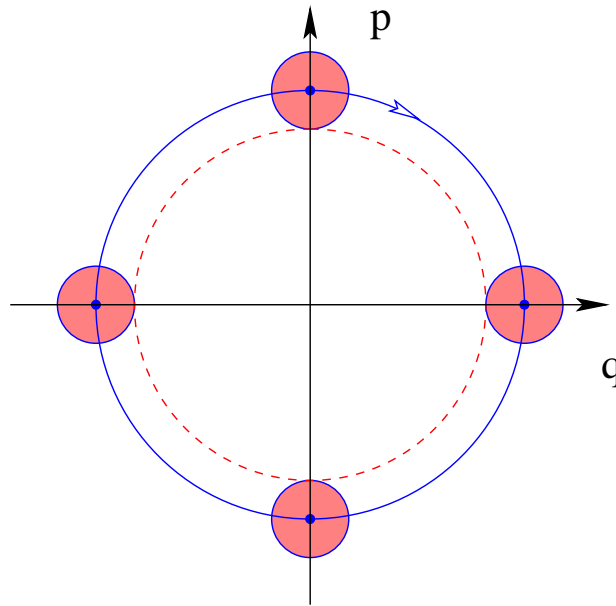


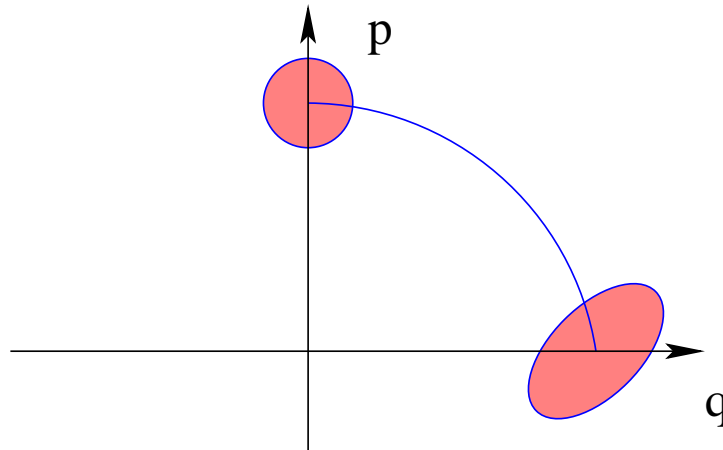
Abbildung 3.13: Phasenraumtrajektorien des harmonischen Oszillators

für die Pendelbewegungen unabhängig von der Amplitude der Schwingung und wir erhalten für eine Trajektorie, die zur Startzeit  $t = 0$  eine Auslenkung  $\varphi = 0$  und eine maximale Geschwindigkeit besitzt die Lösung

$$\varphi(t) = \varphi_0 \sin(\omega t). \quad (3.37)$$

Die entsprechende Phasenraumtrajektorie ist (bei geeigneter Skalierung der Orts- und Impulskoordinaten) ein Kreis, wie in Abb. 3.13 dargestellt ist. Wir betrachten nun Trajektorien, die in einer kreisförmigen Umgebung des Startpunktes beginnen. Zum Beispiel die Trajektorie, die ebenfalls durch  $q(t = 0) = 0$  gekennzeichnet ist aber eine etwas kleinere Geschwindigkeit besitzt. Da beim harmonischen Oszillator, die Winkelgeschwindigkeiten unabhängig von der Amplitude der Bewegung sind, wird diese Trajektorie nach der Zeit  $t = \pi/(2\omega)$  den Umkehrpunkt maximaler Auslenkung erreichen, wie die Referenztrajektorie. Entsprechendes gilt für die anderen Trajektorien, so dass sich das Volumen der Trajektorien (der kleine Kreis um den Referenzpunkt) während der Entwicklung überhaupt nicht ändert.

Betrachtet wir nun auch größere Auslenkungen des mathematischen Pendels aus der Ruhelage, also Winkel  $\varphi$  für die die harmonische Näherung des Potentials in (3.36) nicht mehr angemessen ist. In diesem Fall werden aus den linearen Bewegungsgleichungen des harmonischen Oszillators, nichtlineare Bewegungsgleichungen. Die Lösungen sind nicht mehr die einfachen trigonometrischen Funktionen mit konstanter Winkelgeschwindigkeit unabhängig von der Amplitude der Schwingung. Ohne auf die Details der mathematischen Lösungen dieser Bewegungsgleichungen eingehen zu wollen, kann man sich klar machen, dass durch die nichtlinearen Terme in den Bewegungsgleichungen solche Trajektorien, die mit einem größeren Impuls gestartet werden, länger brauchen bis sie zum Umkehrpunkt der Bewegung gelangen, als solche mit einem kleineren Impuls. Ist der Startimpuls so groß das das Pendel gerade den labilen Punkt bei  $\varphi = \pi$  erreicht, so wird es sogar unendlich lange dauern bis dieser Umkehrpunkt erreicht wird.

Abbildung 3.14: *Phasenraumtrajektorien des nicht-harmonischen Oszillators*

In diesem Fall wird das Volumen der Starttrajektorien offensichtlich verzerrt, etwa in die Form, wie sie in der Abb. 3.14 dargestellt ist. Nach dem Satz von Liouville bleibt aber für hinreichend kleine Volumina das Volumen der Phasenraumtrajektorien erhalten.

Stellen wir uns weiter vor, dass der Startimpuls des Referenzpunktes für das Mathematische Pendel so groß ist, dass dieses bis zum Überschlagpunkt kommt. Solche Trajektorien in der Umgebung, die einen noch größeren Startimpuls mitbekommen haben werden also über den Überschlagpunkt hinauschiessen. Die Winkelgeschwindigkeit beziehungsweise der zugehörige Impuls bleibt positiv und der Winkel macht einen Sprung von  $\varphi = \pi$  nach  $\varphi = -\pi$ . In diesem Fall kommt es also zu einer Aufteilung des Volumens in zwei Teilvolumina, so wie es in Abb. 3.15 dargestellt ist. Nach dem Liouville Theorem bleibt aber auch in diesem Fall das gesamte Volumen erhalten. Diese Aufteilung in Teilvolumina kann natürlich bei komplexeren Systemen, als das eindimensionale Pendel zu topologisch sehr viel komplizierteren Aufspaltungen der Volumina im Phasenraum führen. Als Beispiel seien die Ergebnisse für die Trajektorien im Henon Heils Potenzial angeführt. Dort kam es ja bei höheren Energien zur Ausbildung von chaotischen Prozessen, also komplizierten Mischungen im Phasenraum.

Fügt man den Hamiltonschen Bewegungsgleichungen Terme zu, mit denen die Reibungsverluste des Systems beschrieben werden, so gilt der Satz von Liouville für diese nicht-Hamiltonschen Systeme nicht. In diesem Fall kann man eine Reduktion des Volumens im Phasenraum mit der zeitlichen Entwicklung erwarten, da ja Reibung häufig zu einem stationären Attraktor des Systems führt.

Dabei unterscheidet man zwischen stabilen Systemen, das sind solche Systeme, bei denen die Volumina der betrachteten Phasenraumtrajektorien als Funktion der Zeit kleiner werden, und somit die verschiedenen anfangs benachbarten Trajektorien den gemeinsamen Attraktor zur gleichen Zeit erreichen.

Häufig erreichen aber Trajektorien, die in einer Nachbarschaft starten den Attraktor erst nach sehr unterschiedlichen Laufbahnen. Man spricht in diesem Fall von einer Bahnstabilität des Systems. Zu einer gegebenen Zwischenzeit können die Trajektorien weiter auseinanderliegen als zur Startzeit.

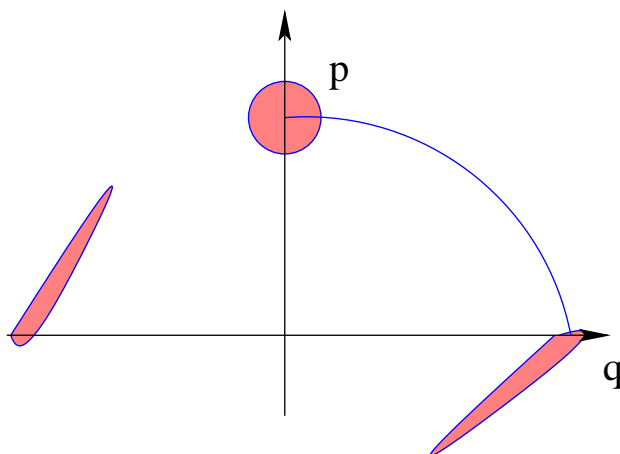


Abbildung 3.15: *Phasenraumtrajektorien des mathematischen Pendels mit Überschlag*

Gilt aber für 2 benachbarte Trajektorien, dass sie für  $t \rightarrow \infty$  den gleichen Endpunkt einnehmen, so spricht man von asymptotischer Stabilität. zu einer

### 3.4 Überlegungen zur Feldtheorie

In diesem Abschnitt zum Abschluss der Analytischen Mechanik wollen wir ganz kurz den Übergang von der klassischen Punktmechanik, also der Beschreibung der Bewegung von Punktteilchen zur Klassischen Feldtheorie andeuten. Dieser Übergang lässt sich in sehr intuitiver Weise am Beispiel der Linearen Kette darstellen.

Dazu wollen wir den Lagrange Formalismus auf eine eindimensionale Kette von Massenpunkten der Masse  $m$  anwenden, die im Fall des ruhenden Systems jeweils den Abstand  $a$  zum nächsten Nachbarn haben. Als generalisierte Koordinaten,  $q_i$ , benutzen wir die jeweilige Auslenkung des Teilchens  $i$  in dieser Kette aus seiner Ruhelage. Bei Annahme einer linearen Rückstellkraft haben wir damit die entsprechende potentielle Energie eines Harmonischen Oszillators mit einer Konstanten  $k$  für die Rückstellkraft in der Form

$$V = \sum_i \frac{1}{2} k (q_{i+1} - q_i)^2,$$

und der kinetischen Energie

$$T = \sum_i \frac{1}{2} m \dot{q}_i^2.$$

Damit ergibt sich also für die Lagrangefunktion

$$\begin{aligned} L &= \sum_i \frac{1}{2} \left[ m \dot{q}_i^2 - k (q_{i+1} - q_i)^2 \right] \\ &= \sum_i a \frac{1}{2} \left[ \frac{m}{a} \dot{q}_i^2 - k a \frac{(q_{i+1} - q_i)^2}{a^2} \right] \\ &= \sum_i a \mathcal{L}_i. \end{aligned} \quad (3.38)$$

Betrachten wir jetzt den Grenzfall, dass der Abstand zwischen den Massenpunkten infinitesimal wird

$$a \rightarrow dx,$$

so wird der Quotient

$$\frac{m}{a} \rightarrow \mu,$$

zur Massendichte  $\mu$  und

$$\frac{(q_{i+1} - q_i)}{a} \rightarrow \frac{dq}{dx},$$

zur Ableitung der lokalen Auslenkung  $q(x)$  nach der Position in der Kette  $x$ . Bezeichnet man jetzt noch das Produkt

$$k a \rightarrow \chi,$$

so ergibt sich aus der letzten Zeile von (3.38) für die Lagrangefunktion

$$L = \int dx \mathcal{L},$$

mit der Lagrangedichte  $\mathcal{L}$

$$\mathcal{L} = \frac{1}{2} \left[ \mu \left( \frac{dq}{dt} \right)^2 - \chi \left( \frac{dq}{dx} \right)^2 \right]. \quad (3.39)$$

Diese Lagrangedichte hängt also von den Ableitungen der Auslenkungsfunktion  $q(x, t)$  nach Ort und Zeit ab. Das Ziel der Klassischen Mechanik ist es nun, diese lokale Auslenkung der Kette,  $q(x, t)$  zu bestimmen. Dazu betrachten wir das Hamiltonsche Prinzip (1.99), das sich jetzt mit der Lagrangedichte darstellt in der Form

$$\delta \int \mathcal{L} dt dx = 0 \quad (3.40)$$

Auch hier haben wir es mit einem Euler - Lagrange Variationsprinzip zu tun, bei dem allerdings die Raum- und Zeitkoordinate gleich behandelt werden, was aus der Sicht der Relativitätstheorie natürlich eine sehr wichtige Eigenschaft ist. Dieses Variationsprinzip führt auf entsprechende Euler - Lagrange Gleichungen in der Form

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \frac{dq}{dt}} + \frac{d}{dx} \frac{\partial \mathcal{L}}{\partial \frac{dq}{dx}} - \frac{\partial \mathcal{L}}{\partial q} = 0. \quad (3.41)$$

Die Lagrangedichte (3.39) hängt von der Funktion  $q(x, t)$  explizit nicht ab, sondern nur von den Ableitungen dieser gesuchten Funktion nach Raum und Zeit. Damit entfällt der letzte Term auf der linken Seite von (3.41) und wir erhalten die Euler - Lagrange Gleichung

$$\begin{aligned} 0 &= \frac{d}{dt} \mu \frac{dq}{dt} - \frac{d}{dx} \chi \frac{dq}{dx} \\ &= \mu \frac{d^2 q}{dt^2} - \chi \frac{d^2 q}{dx^2}. \end{aligned}$$

Dies ist die Gleichung für die Ausbreitung einer Wellenstruktur in der Raumrichtung  $x$  mit einer Ausbreitungsgeschwindigkeit von

$$v = \sqrt{\frac{\chi}{\mu}}.$$

Dieses Beispiel für den Fall mit einer Raumrichtung lässt sich natürlich schnell verallgemeinern für 3 Raumrichtungen. Gesucht wird in diesem Fall also ein skalares Feld

$$\phi(\vec{r}, t),$$

also eine Funktion  $\phi$ , die jedem Raum - Zeitpunkt eine Zahl, also einen Skalar zuordnet. Deshalb der Name Skalarfeld. Die Dynamik dieses Skalarfeldes wird durch eine Lagrangedichte  $\mathcal{L}$  charakterisiert, die von dem Skalarfeld und den Ableitungen dieses Skalarfeldes nach den kartesischen Koordinaten des Ortsvektors, die wir im folgenden mit  $x_k$ ,  $k = 1, 2, 3$  bezeichnen wollen und der Ableitung von  $\phi$  nach der Zeit  $t$

$$\mathcal{L} \left( \phi, \frac{d\phi}{dx_1}, \frac{d\phi}{dx_2}, \frac{d\phi}{dx_3}, \frac{d\phi}{dt} \right).$$

Das Variationsprinzip

$$\delta \int \mathcal{L} dt d^3 r = 0$$

führt dann auf die Euler - Lagrange Gleichung

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \frac{d\phi}{dt}} + \sum_{k=1}^3 \frac{d}{dx_k} \frac{\partial \mathcal{L}}{\partial \frac{d\phi}{dx_k}} - \frac{\partial \mathcal{L}}{\partial \phi} = 0. \quad (3.42)$$

# Kapitel 4

## Einführung in die Quantenmechanik

### 4.1 Grenzen der klassischen Physik: “Die Physiker spinnen”

Bei vielen Mitmenschen, die in anderen Bereichen durchaus gebildet sind, findet man dass in der Physik ein sehr skurilles Weltbild entwickelt worden ist, das aber eigentlich mit dem wirklichen Leben nichts zu tun hat. Als Beispiel für ein solch unsinniges Weltbild wird dann etwa das Zwillingsparadoxon oder andere Phänomene der speziellen Relativitätstheorie herangezogen.

Da ja Licht kein Medium benötigt um sich auszubreiten, ist die Lichtgeschwindigkeit  $c$  in allen Bezugssystemen gleich. Wenn ich also einem Lichtsignal hinterher laufe, bewegt es sich, egal wie schnell ich mich bewege immer noch mit Lichtgeschwindigkeit von mir weg. Man kann Licht also, egal wie sich die Technik noch entwickeln sollte, niemals einholen oder gar überholen. Das klingt doch absurd, oder?

Aber in der Speziellen Relativitätstheorie gibt es noch merkwürdigere Vorhersagen. Die Konstanz der Lichtgeschwindigkeit führt notwendig dazu, dass es keine absolute Zeit gibt. Bei einem Übergang von einem Bezugssystem  $A$  zu einem anderen Bezugssystem  $B$ , das sich relativ zu  $A$  mit einer Geschwindigkeit  $v$  bewegt, müssen nicht nur die Koordinaten ineinander umgerechnet werden. Auch für die in dem jeweiligen Bezugssystem gemessene Zeit  $t_A$  bzw.  $t_B$  muss mittels der Lorentztransformation eine Umrechnung vorgenommen werden, wonach

$$t_B = t_A \sqrt{1 - \frac{v^2}{c^2}}.$$

Wenn ich mich also nur genügend bewege geht meine Uhr langsamer. Ein Zwilling, der oder die seinen Zwillingspartner verlässt, sich mit großer Geschwindigkeit  $v$  von ihm weg bewegt um dann nach einer gewissen Zeit wieder umzukehren, wird seinen Zwillingspartner wesentlich älter vorfinden, als er selbst.

Auch dies erscheint absurd, ist aber sowohl bei Elementarteilchen, die sich hinreichend schnell bewegen können, als auch bei großen Objekten (Menschen) immer wieder überzeugend nachgewiesen worden.

Warum erscheinen uns diese Phänomene absurd? Die Antwort ist natürlich darin be-

gründet, dass wir in unserem Alltag eigentlich nur Abläufe beobachten, bei denen sich Objekte relativ zueinander bewegen mit Geschwindigkeiten, die um viele Größenordnungen kleiner sind als die Lichtgeschwindigkeit. So kann man also etwa bei der relativistischen Beziehung zwischen der Energie und dem Impuls eines freien Teilchens mit der Ruhemasse  $m$

$$E = \sqrt{m^2 c^4 + c^2 (mv)^2},$$

den Faktor  $mc^2$  aus der Wurzel herausziehen und die verbleibende Wurzel nach Potenzen von  $v^2$  über  $c^2$  entwickeln

$$\begin{aligned} E &= mc^2 \sqrt{1 + \frac{v^2}{c^2}} \\ &= mc^2 \left[ 1 + \frac{1}{2} \frac{v^2}{c^2} - \frac{1}{8} \left( \frac{v^2}{c^2} \right)^2 + \dots \right] \\ &= mc^2 + \frac{1}{2} mv^2 - \dots \end{aligned}$$

Wenn der Quotient  $v^2/c^2$  sehr klein gegenüber 1 ist, was ja in unserer Alltagswelt immer gegeben ist, so reicht es vollkommen aus, die Terme quadratischer und höherer Ordnung in dieser Entwicklung zu vernachlässigen, sodass sich also neben der Ruheenergie der Teilchen  $E = mc^2$  gerade der nichtrelativistische Ausdruck für die kinetische Energie des Teilchens ergibt.

Zur Beschreibung unserer Alltagswelt reicht also die Klassische (nichtrelativistische) Mechanik. Diese Klassische Mechanik lässt sich als nichtrelativistischer Grenzfall ( $v \ll c$ ) aus der allgemeineren relativistischen Mechanik als Spezialfall kleiner Geschwindigkeiten herleiten. Es geht aber nicht umgekehrt, den allgemeinen relativistischen Ausdruck aus der Klassischen Mechanik herzuleiten.

Andererseits erscheinen uns die oben diskutierten Phänomene nur deshalb absurd, da unser Erfahrungen auf dem Leben in der Klassischen Welt basieren.

Das gleich gilt aber natürlich auch für den Übergang von der Klassischen Physik zur Physik der Quantenmechanik. Auch hier erfahren wir in unserem täglichen Leben nur die Klassische Welt. Die umfassendere Beschreibung liegt aber in der Quantenmechanik. Auch hier können wir nicht erwarten, dass alle Phänomene der Quantenmechanik auch in unsere Alltagswelt beobachtbar sind und uns deshalb verständlich erscheinen.

Genauso können wir auch in diesem Fall die Quantenmechanik nicht aus der Klassischen Mechanik herleiten, sondern müssen umgekehrt sicherstellen, dass der Klassische Grenzfall der Quantenmechanik, also die Welt bei der Wirkungen (z.B. das Produkt aus der Ortsverschiebung eines Teilchens und sein Impuls) groß sind gemessen am Planckschen Wirkungsquantum  $h = 2\pi\hbar$ , sich als Grenzfall der Quantenmechanik ergibt.

In der Abbildung 4.1 sind diese Zusammenhänge skizziert. So ist die Klassische Physik, der Eckpunkt, der mit der Ziffer 1 bezeichnet ist, einerseits der Grenzfall der Speziellen Relativitätstheorie für Geschwindigkeiten, für die  $c$  groß, also  $1/c$  klein ist, andererseits aber auch der Grenzfall der Quantenmechanik für Wirkungen, die im Verhältnis zum Planckschen Wirkungsquantum  $h$  groß sind.

Die relativistische Quantenmechanik oder relativistische Quantenfeldtheorie (Eckpunkt 4)



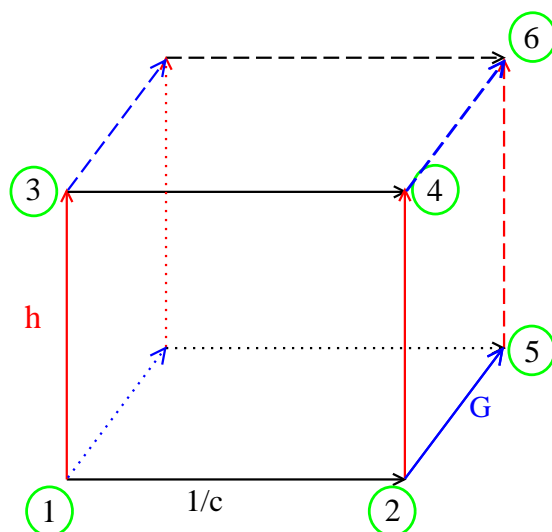


Abbildung 4.1: Gültigkeitsbereich unterschiedlicher Theorien in der Physik. In dieser Figur bezeichnet der Eckpunkt 1 den Bereich der Klassischen Physik, 2 den Bereich der Speziellen Relativitätstheorie, 3 den der nichtrelativistischen Quantenmechanik, 4 den der relativistischen Quantenmechanik und Quantenfeldtheorie, 5 den der Allgemeinen Relativitätstheorie und 6 den allgemeinsten Bereich der noch zu entwickelnden Quantisierung dieser Allgemeinen Relativitätstheorie

umfasst sowohl die Eigenschaften der Speziellen Relativitätstheorie als auch die der Quantenmechanik.

All die bisher genannten Theorien bezeichnen Systeme, bei denen die Gravitation als Newtonsche Näherung der Allgemeinen Relativitätstheorie (Eckpunkt 5) behandelt werden kann. Die umfassende Theorie, also die Quantisierung der Allgemeinen Relativitätstheorie ist bisher noch nicht formuliert worden. Dazu muss insbesondere eine Theorie für die Quantisierung der Geometrie der Raumzeit entwickelt werden. Eine solche Theorie ist insbesondere erforderlich, um die Kosmologie des sogenannten Urknalls zu beschreiben.

## 4.2 Experimentelle Hinweise auf die Quantenmechanik

In unserer bisherigen Betrachtung der Physik haben wir immer eine klare Abgrenzung zwischen Wellenphänomenen und Teilchenphänomenen ziehen können. Beispielsweise waren wir in der Lage alle Eigenschaften von Licht im Rahmen einer Wellentheorie zu deuten. Auf der anderen Seite konnten wir in der Mechanik und Thermodynamik die Eigenschaften von Atome, Moleküle und anderen Körpern stets als Teilchenphänomene erklären.

Gegen Ende des 19. Jahrhunderts und auch zum Anfang des 20. Jahrhunderts wurde in der Physik eine Reihe wichtiger Experimente durchgeführt, die nicht im Einklang mit der damals existierenden Theorie zu bringen waren. Zusammen mit der Entwicklung einiger neuen theoretischen Konzepte stellen diese Experimente die Geburtsstunde der Quantenphysik dar.

Zielsetzung diesen Abschnitts ist sowohl die Vorstellung dieser historisch wichtigen Entwicklungen als auch die Betrachtung einiger späterer Experimente, die die Grenzen der klassischen Physik aufzeigen. Insbesondere werden wir Experimente diskutieren, die

- die Teilcheneigenschaften des Lichts
- den Wellencharakter von Teilchen
- die Quantisierung physikalischer Größen

darstellen.

### 4.2.1 Der Fotoeffekt

Der Fotoeffekt wurde von Heinrich Hertz im Jahr 1887 im Rahmen seiner Untersuchungen zur Ausbreitung elektromagnetischer Strahlung entdeckt. Wenn man metallische Platten mit Licht bestrahlt, können Elektronen<sup>1</sup> emittiert werden. Die Emission findet aber nur statt, wenn die Frequenz  $\nu$  der elektromagnetischen Strahlung eine kritische Frequenz  $\nu_0$  überschreitet. Diese kritische Frequenz hängt vom Material, also dem Metall der Platte ab: z.B. für Lithium beträgt sie  $\nu_0 = 5.7 \times 10^{14}$  Hz entsprechend eine Wellenlänge von  $\lambda = 526$  nm (also grünes Licht). Der emittierte Strom der Ladungsträger ist direkt proportional zur Intensität des eingestrahnten Lichts.

Mit dem in der Abbildung 4.2 skizzierten Experimentaufbau kann man auch die kinetische Energie bestimmen, mit der die Elektronen aus der Metalloberfläche emittiert werden. Dazu legt man ein Spannung zwischen der Metalloberfläche und dem Auffangschirm für die emittiertem Elektronen an. Dieser Auffangschirm ist als gestrichelte Linie dargestellt und mit dem negativen Pol der Spannungsquelle verbunden, sodass nur solche Elektronen

---

<sup>1</sup>Der Beweis für die Emission von Elektronen kam natürlich erst später mit der Entdeckung des Elektrons durch J.J. Thomson im Jahr 1897.

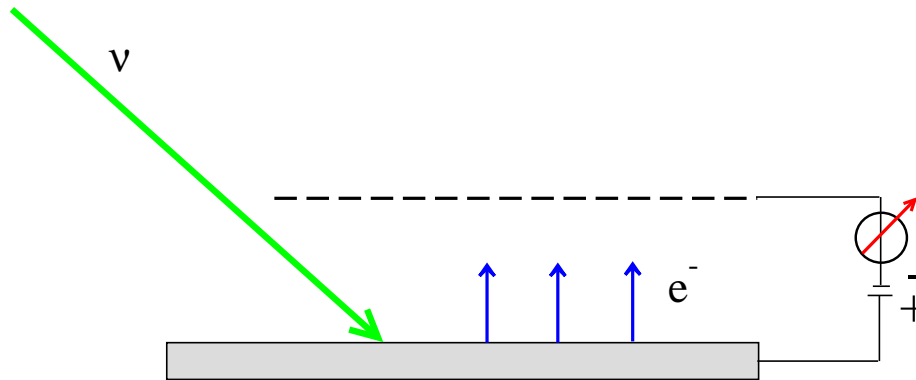


Abbildung 4.2: Schematischer Aufbau des Experiments zur Messung des Fotoeffekts

zu diesem Schirm gelangen, deren kinetische Energie bei der Emission groß genug ist um das “Stopping Potenzial”  $eU$  zu überwinden. Dabei zeigte sich, dass die Energie der emittierten Fotoelektronen nicht von der Intensität des Lichts sondern ebenfalls von der Frequenz des eingestrahlt Lichts abhängt.

Zur Erklärung dieses Fotoeffekts postulierte Einstein im Jahre 1905, dass Licht aus einer Ansammlung von sogenannten Lichtquanten besteht. Aus unserer heutigen Sicht sind diese Quanten mit den Photonen des Lichts zu identifizieren. Jedes Photon besitzt die Energie  $h\nu$ , wobei die Konstante  $h = 2\pi\hbar = 6.63 \times 10^{-34}$  Js die Planck’sche Wirkungskonstante bezeichnet, die Planck zur Erklärung der Schwarzkörperstrahlung zuvor eingeführt hatte (siehe unten). Die Energie der elektromagnetischen Welle wird durch die Absorption von Photonen auf die Materie übertragen. Dabei wird die Energie eines Elektrons im Metall durch die Absorption eines Photons um  $\Delta E = h\nu$  erhöht. Um die Bindungsenergie des Elektrons im Metall zu überwinden, muss die sogenannte Austrittsarbeit  $W_A$  aufgebracht werden und nur der Rest der Energie wird in Form kinetischer Energie dem Elektron zur Verfügung gestellt. Somit erhalten wir für die kinetische Energie:

$$\frac{1}{2}mv^2 = eU = h\nu - W_A. \quad (4.1)$$

Das Stopping Potenzial  $U_{SP}$  ist dann die Spannung, bei der gerade keine Elektronen mehr auf den Schirm gelangen und deshalb der Strom im angeschlossenen Kreis verschwindet. Aus der Frequenzabhängigkeit dieses Stopping Potenzials

$$U_{sp} = \frac{h}{e}\nu - \frac{W_A}{e}. \quad (4.2)$$

ergibt sich die Möglichkeit zur direkten Bestimmung der Planck’schen Wirkungskonstante und die Austrittsarbeit ist durch die kritische Frequenz des Fotoeffekts bestimmt

$$h\nu_0 = W_A.$$

Der Fotoeffekt ist ein Beispiel dafür, dass Licht, also elektromagnetische Wellen, Teilcheneigenschaften in der Art besitzen, dass die Energie in Form von festen Energiepaketen (Quanten) übertragen wird.

### 4.2.2 Schwarzkörperstrahlung

Unter Schwarzkörperstrahlung versteht man die Abstrahlung von elektromagnetischen Wellen eines Körpers, der im Kontakt mit einem Wärmebad steht, aber keine material-spezifische Abstrahlung erzeugt. Als ideale Realisierung einer solchen Schwarzkörperstrahlung können wir uns die elektromagnetischen Wellen in einem Hohlraum vorstellen. Durch den thermischen Kontakt mit der Wand des Hohlraums werden laufend Photonen, also die Energiequanten des elektromagnetischen Feldes, erzeugt und auch wieder absorbiert. Die Zahl der Photonen ist also nicht konstant.

Die Eigenschaften dieser Photonen, der Energiequanten des elektromagnetischen Feldes, werden zusammengefaßt durch (c bezeichnet die Lichtgeschwindigkeit):

$$\begin{aligned}\omega &= c|\vec{k}| = ck \\ k &= \frac{2\pi}{\lambda} \\ \vec{p} &= \hbar\vec{k} \quad \text{mit} \quad \hbar = \frac{h}{2\pi} \\ \varepsilon_p &= \hbar\omega = h\nu, \end{aligned} \tag{4.3}$$

wobei  $\hbar$  bzw.  $h$  wie üblich für das Plancksche Wirkungsquantum steht und  $\vec{k}$  als Wellenzahlvektor bezeichnet wird. Mit diesen Annahmen für die Energie und Impuls der Photonen, kann man nun die Frage stellen, wie viele Energiepakete  $\varepsilon = \hbar\omega = h\nu$  sich in dem Volumen des Hohlraums befinden. Die Wahrscheinlichkeit  $P_n$   $n$  dieser Energiepakete vorzufinden ist entsprechend der Boltzmann Verteilung proportional zu

$$P_n \sim e^{-n\varepsilon\beta} \quad \text{mit} \quad \beta = \frac{1}{k_B T} \tag{4.4}$$

wobei  $k_B$  die Boltzmann Konstante und  $T$  die Temperatur des Hohlraums bezeichnet. Die richtige Normierung

$$\sum_n P_n = 1$$

erhält man dann durch

$$P_n = \frac{e^{-n\varepsilon\beta}}{\sum_l e^{-l\varepsilon\beta}}.$$

Der statistische Mittelwert für die Anzahl der Photonen ergibt sich damit zu

$$\langle n \rangle = \sum_n n P_n \tag{4.5}$$

$$\begin{aligned} &= \sum_n n \frac{e^{-n\varepsilon\beta}}{\sum_l e^{-l\varepsilon\beta}} \\ &= -\frac{1}{\beta} \frac{d}{d\varepsilon} \log \left( \sum_n e^{-n\varepsilon\beta} \right). \end{aligned} \tag{4.6}$$

Die zweite Gleichung kann man einfach dadurch verifizieren, dass man die Differentiation des Logarithmus ausführt.

Nun entspricht aber das Argument der Logarithmusfunktion einer geometrischen Reihe

$$\sum_n e^{-n\varepsilon\beta} = \sum_n \left( e^{-\varepsilon\beta} \right)^n = \frac{1}{1 - e^{-\varepsilon\beta}}.$$

Setzt man diese Darstellung in Gl.(4.5) ein, so ergibt sich

$$\begin{aligned}
 \langle n \rangle &= -\frac{1}{\beta} \frac{d}{d\varepsilon} \log \left( \frac{1}{1 - e^{-\varepsilon\beta}} \right) \\
 &= \frac{1}{\beta} \frac{d}{d\varepsilon} \log (1 - e^{-\varepsilon\beta}) \\
 &= \frac{e^{-\varepsilon\beta}}{1 - e^{-\varepsilon\beta}} \\
 &= \frac{1}{e^{\varepsilon\beta} - 1}.
 \end{aligned} \tag{4.7}$$

Da nach den Gesetzen der statistischen Physik die Zahl der Zustände proportional dem Phasenraumvolumen dieser Zustände ist, gilt also für die Zahl der Photonen mit Impuls  $\vec{p}$  und Koordinate  $\vec{q}$  in dem Hohlraum

$$dN_{\vec{p}\vec{q}} = 2 \frac{d^3p d^3q}{h^3} \langle n \rangle,$$

wobei der Faktor 2 der Zahl der unabhängigen transversalen Polarisierungen für elektromagnetische Strahlungen entspricht und der Nenner  $h^3$  dafür sorgt, dass  $dN$  dimensionslos ist<sup>2</sup>.

Die Zahl der Photonen im Volumen  $V$  mit dem Betrag des Impulses  $p$  ergibt sich dann aus  $dN_{\vec{p}\vec{q}}$  durch Integration von  $d^3q$  und über die Kugelschale der Impulse vom Betrag  $p$  zu

$$dN_p = 2 \frac{V 4\pi p^2 dp}{h^3} \frac{1}{e^{\varepsilon\beta} - 1}. \tag{4.8}$$

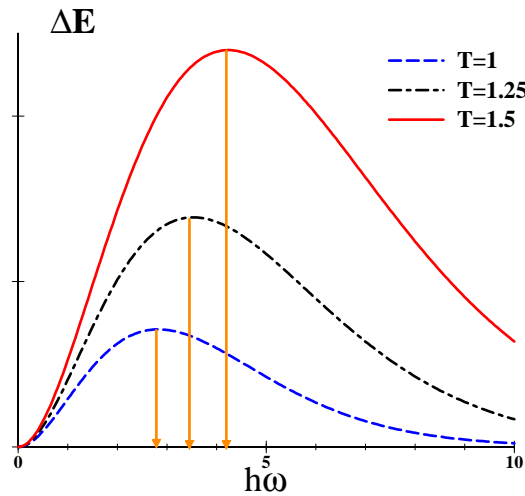
Die Energiedichte erhält man jetzt dadurch, dass man  $dN_p$  mit der Energie der Photonen  $\varepsilon = h\nu$  multipliziert und unter Ausnutzung der Energie Impulsbeziehung  $cp = h\nu$  aus (4.3) zu

$$\begin{aligned}
 dE &= \frac{1}{V} h\nu dN_p \\
 &= \frac{8\pi h \nu^3}{c^3} \frac{1}{e^{h\nu\beta} - 1} d\nu \\
 &= \frac{\hbar\omega^3}{c^3\pi^2} \frac{1}{e^{\hbar\omega\beta} - 1} d\omega.
 \end{aligned} \tag{4.9}$$

Dies entspricht der **Planck'schen Strahlungsformel**. Beispiele für dieses Energiespektrum der Strahlung eines Schwarzen Körpers bei verschiedenen Temperaturen  $T = 1/k_B\beta$  sind in Abb. 4.3 dargestellt.

Auf der einen Seite beobachtet man einen Abfall der Energiedichte bei Photonen kleiner Energie. Dieser Abfall hat seine Ursprung im Faktor  $\omega^3$  im Zähler des Ausdrucks in (4.9) und basiert natürlich zum Teil darauf, daß diese Photonen eine geringere Energie besitzen (eben  $\sim \omega$ ) und so natürlich jedes Photon einen geringeren Beitrag zu Energiedichte

<sup>2</sup>Dass das Volumen einer Elementarzelle gerade durch  $h^3$  gegeben ist mit  $h$  dem Planckschen Wirkungsquantum ist natürlich eigentlich erst ein Ergebnis der Quantenmechanik

Abbildung 4.3: Spektrum des Schwarzen Strahlers für verschiedene  $T$ .

leistet. Hinzu kommt aber insbesondere, dass die Zahl der Photonenzustände mit vorgegebener Energie bzw. Impuls mit dem Faktor  $p^2$  in (4.8) abfällt, was einen Faktor  $\sim \omega^2$  im Zähler des Ausdrucks in (4.9) beiträgt. Dieser Abfall bei kleinen Energien beruht also auf der klassischen Eigenschaft des Phasenraumes: In ein vorgegebenes Volumen *passen* weniger elektromagnetische Wellen mit kleiner Wellenzahl  $k$ , d.h. großer Wellenlänge  $\lambda$  (siehe (4.3)).

Zur genaueren Untersuchung des Verhaltens bei kleinen Energien, entwickeln wir die Exponentialfunktion im Nenner des Ausdrucks von (4.9) für kleine Werte von  $\hbar\omega\beta$

$$e^{\hbar\omega\beta} \approx 1 + \hbar\omega\beta$$

was, eingesetzt in die Plancksche Formel, die Strahlungsformel von Rayleigh und Jeans liefert

$$dE/d\omega \sim \omega^2$$

Für große  $\omega$  divergiert die Energiedichte nach dem Rayleigh Jeansschen Gesetz; die integrierte Energie des Schwarzen Strahlers wäre unendlich groß. Dieses Ergebnis bezeichnet man deshalb auch als *Ultraviolett katastrophe* der klassischen Rayleigh Jeans Formel.

Während also der Abfall der Planckschen Strahlungsformel bei kleinen Energien im Rahmen der klassischen Theorie zu verstehen ist, kann der Abfall zu großen  $\omega$  nur mit der Quantenmechanik verstanden werden. Wegen der Beziehung zwischen Frequenz und Energiequantum in (4.3) können elektromagnetische Wellen mit großem  $\omega$  nur mit großen Energiequanten erzeugt werden. Die Erzeugung solch großer Energiequanten ist aber bei einem Kontakt mit einem Wärmebad der Temperatur  $T$  unterdrückt.

Für große Werte  $\hbar\omega\beta$  geht die Plancksche Formel in die 1893 von Wien empirisch gefundene Strahlungsformel über. Das Verdienst von Planck besteht also vor allen Dingen darin, diese beiden empirischen Formeln miteinander verbunden zu haben.

Bisher haben wir die Strahlung des Schwarzen Körpers vor allen Dingen als Strahlung in einem Hohlraum interpretiert, bei dem die Temperatur der Wand für die elektromagnetischen Wellen verantwortlich ist. In guter Näherung kann aber auch z.B. die Sonne als

Quelle für die Strahlung eines *Schwarzen Körpers* ansehen: Die Sonne übernimmt also die Rolle der heißen Wand und wir, die Beobachter befinden uns im Hohlraum. So kann man mit Hilfe des Modells vom Schwarzen Strahler durch Messung des Spektrums eines Objekts Rückschlüsse auf dessen Temperatur ziehen. Diese Methode findet z.B. in der Astronomie Anwendung.

Beispielsweise liegt das Maximum bei Zimmertemperatur  $T \approx 300$  K im infraroten Bereich, wogegen bei Temperaturen von 5800 K, wie sie an der Sonnenoberfläche herrschen, hauptsächlich sichtbares Licht ( $\lambda \approx 500$  nm) emittiert wird.

Als weiteres Beispiel für die Strahlung eines Schwarzen Körpers sei die **Kosmische Hintergrundstrahlung** diskutiert. Im Jahre 1965 bemerkten Arno Penzias und Robert Wilson bei Messungen mit den ersten Kommunikationssatelliten ein Rauschen. Da es vollkommen unabhängig von der Beobachtungsrichtung war, wurde es von ihnen zunächst als Störgeräusch interpretiert. Als dieses Rauschen aber auch durch intensive Bemühungen um eine Verbesserung der Empfangsantennen und der Verstärker nicht unterdrückt werden konnte, suchte man nach anderen Erklärungen. Dabei wurde klar, daß Penzias und Wilson mit diesem Rauschen das Echo des Kosmologischen **Big Bang** registriert hatten.

Die Frequenz der Radiostrahlung entspricht der Strahlung eines Schwarzen Körpers bei der Temperatur  $T = 2,7$  K. Zur Erklärung dieser Hintergrundstrahlung muß man berücksichtigen, daß kurz nach dem Urknall die Materie wegen der hohen Temperaturen vollständig ionisiert war. Die freien Ladungen, Elektronen und Protonen, wechselwirken so stark mit den elektromagnetischen Wellen, dass die Materie und die Strahlung gleiche Temperatur besaßen. Erst bei einer Abkühlung des Weltalls auf  $T \approx 3000$  K, was der Bindungsenergie der Elektronen in einem Atom entspricht, konnten sich neutrale Atome bilden. Seit diesem Zeitpunkt, also etwa 380000 Jahre nach dem Urknall, sind die Wechselwirkungen zwischen Photonen und Materie erheblich schwächer. Die Folge ist eine Entkoppelung von Photonen und Materie,  $T_{\text{Mat}} \neq T_{\text{Phot}}$ . Der Weltraum sollte also leuchten wie ein Körper mit  $T = 3000$  K. Da die 3000 K-Strahlung vor sehr langer Zeit entstanden ist, sehen wir sie heute vom Rand des Universums isotrop einfallen. Dieser bewegt sich aber nahezu mit Lichtgeschwindigkeit von uns fort, was zur Folge hat, dass die bei uns ankommende Strahlung extrem rotverschoben ist:  $T = 2,7$  K.

Eine alternative, im Rahmen der Relativitätstheorie jedoch äquivalente Erklärung für diese Verschiebung ist, dass durch die Expansion des Weltalls eine Umskalierung in Richtung größerer Längen erfolgt ist, sodaß einer Welle mit einer bestimmten ursprünglichen Wellenlänge—entsprechend  $T = 3000$  K—also heute eine mit erheblich größerer Wellenlänge entspricht:  $T = 2,7$  K. Wegen der großen Bedeutung für die Kosmologie wurden Penzias und Wilson für ihre Entdeckung 1980 mit dem Nobelpreis ausgezeichnet.

### 4.2.3 Compton Effekt

Ein weiteres Experiment, das den Teilchencharakter des Lichts zeigt, ist die Comptonstreuung. Dazu betrachtet man die Streuung von Röntgenstrahlung oder kurzwelligem Licht der Wellenlänge  $\lambda_i$  an einem System, das freie Elektronen enthält, so findet man, dass sich die Wellenlänge des gestreuten Lichts  $\lambda_f$  als Funktion des Streuwinkels  $\theta$  mit dem Gesetz

$$\lambda_f - \lambda_i = \Lambda_c [1 - \cos \theta] , \quad (4.10)$$

beschreiben lässt, wobei  $\Lambda_c$  die sogenannte Comptonwellenlänge bezeichnet, die nach dem amerikanischen Physiker A.H. Compton benannt wurde, der sich um 1920 intensiv mit diesem Phänomen befasst hat.

Zur Beschreibung der Comptonstreuung betrachten wir Energie- und Impulserhaltung für die Streuung von Photonen an freien Elektronen, wobei wir annehmen, dass die Elektronen vor dem Stoß ruhen. Die Impulserhaltung

$$\vec{p}_i = \vec{p}_f + \vec{p}_e, \quad (4.11)$$

mit  $\vec{p}_i$  ( $\vec{p}_f$ ) dem Impuls des Photons vor (nach der Streuung) und  $\vec{p}_e$  dem Impuls des Elektrons nach dem Stoß. Quadriert man diese Gleichung der Impulserhaltung ergibt sich

$$\begin{aligned} \vec{p}_e^2 &= (\vec{p}_i - \vec{p}_f)^2 \\ &= \vec{p}_i^2 + \vec{p}_f^2 - 2\vec{p}_i \cdot \vec{p}_f \\ &= p_i^2 + p_f^2 - 2p_i p_f \cos \theta \\ &= \frac{1}{c^2} [E_i^2 + E_f^2 - 2E_i E_f \cos \theta], \end{aligned} \quad (4.12)$$

wobei z.B.  $E_i = cp_i$  die Energie des einlaufenden Photons bezeichnet. Aus der Energieerhaltung

$$E_i + m_e c^2 = E_f + c\sqrt{m_e^2 c^2 + p_e^2},$$

ergibt sich durch Quadrierung

$$p_e^2 = \frac{1}{c^2} [E_i^2 + E_f^2 + 2(E_i - E_f)m_e c^2 - 2E_i E_f] \quad (4.13)$$

Aus (4.12) und (4.13) zusammen ergibt sich dann

$$(E_i - E_f) = \frac{E_i E_f}{m_e c^2} (1 - \cos \theta),$$

bzw.

$$\frac{1}{E_f} - \frac{1}{E_i} = \frac{1}{m_e c^2} (1 - \cos \theta).$$

oder auch mit  $E_f = h\nu_f$

$$\frac{ch}{h\nu_f} - \frac{ch}{h\nu_i} = \frac{hc}{m_e c^2} (1 - \cos \theta).$$

Dies ist aber genau die Beziehung der Comptonstreuung 4.10 mit der Comptonwellenlänge für die Elektronen

$$\Lambda_c = \frac{h}{m_e c}. \quad (4.14)$$

All diese Experimente zeigen also die Teilcheneigenschaften der elektromagnetischen Wellen in Form von Photonen.



### 4.2.4 Doppelspaltexperiment mit Elektronen

Nach der Entdeckung des Welle - Teilchen Dualismus für die elektromagnetischen Wellen formuliert Louis de Broglie 1924 in seiner Doktorarbeit, dass auch Klassische Teilchen einen solchen Welle - Teilchen Dualismus aufweisen sollten. Danach sollte also z.B. ein Elektron, das mit einer Beschleunigungsspannung  $U$  auf eine kinetische Energie  $E = eU$  gebracht wird und damit einen Impuls

$$|\vec{p}| = \sqrt{2m_e E}$$

besitzt durch eine Welle der Wellenlänge (vergleiche Gl.(4.3))

$$\lambda = \frac{h}{|\vec{p}|} = \frac{h}{\sqrt{2m_e E}}, \quad (4.15)$$

beschrieben wird. Man bezeichnet diese Wellenlänge auch als **de Broglie Wellenlänge**.

Der Nachweis dieser Welleneigenschaft erfolgt natürlich am besten durch ein Interferenzexperiment. Welleneigenschaft von Licht entspricht, also z.B. durch die Interferenz einer solchen Materiewelle beim Durchgang durch einen Doppelspalt. Das Problem eines solchen Experimentes liegt allerdings darin, dass die de Broglie Wellenlänge eines Elektrons, das nur auf eine Energie von 50 keV beschleunigt wurde, lediglich  $5 \cdot 10^{-12}$  m beträgt, also 0.05 Angström und damit viel kleiner als die Ausmaße eines Atoms ist. Deshalb hielt man dieses Experiment auch lange Zeit für undurchführbar. Der historische erste Nachweis für die Materiewellen gelang Davisson und Germer 1927 durch die Beugung von Elektronen an einem Festkörper Einkristall. Nachdem diese Arbeit, die von de Broglie postulierte Materiewelle bestätigte, erhielt dieser 1929 den Nobelpreis.

Das Doppelspaltexperiment wurde erstmals durch den Tübinger Physiker Claus Jönsson 1959 in seiner Doktorarbeit in der Arbeitsgruppe von G. Möllenstedt realisiert. Wegen seiner überragenden Bedeutung als ein Schlüsselexperiment der Quantenmechanik wurde dieses Experiment in einer Umfrage der englischen physikalischen Gesellschaft zum schönsten Experiment aller Zeiten gewählt. Inzwischen gibt es entsprechende Experimente auch für "größere" Teilchen, wie z.B. Atome, Moleküle und Fullerene.

Wir wollen einige wesentlichen Ergebnisse eines solchen Experimentes diskutieren, ohne näher auf die Schwierigkeiten der experimentellen Umsetzung einzugehen. Schematisch ist ein solcher Doppelspaltversuch in der Abb. 4.4 dargestellt. Die Elektronen aus der Elektronenquelle werden durch die Beschleunigungsspannung  $U$  auf die Energi  $eU$  beschleunigt und sollten durch die Wellenlänge nach (4.15) charakterisiert sein.

Durch die Beugung dieser Materiewelle am Doppelspalt erhält man unter geeigneten Bedingungen auf dem Detektorschirm in der Tat ein Beugungsmuster von 2 interferierenden Teilstrahlen, was den Wellencharakter dieses Elektronenstrahls belegt.

Nun kann man einwenden, dass dieses Beugungsmuster kein Beleg für den Wellencharakter des Elektronenstrahls ist, sondern vielmehr auf der Wechselwirkung der Elektronen, die durch den oberen Spalt fliegen mit denen, die den unteren Spalt passieren. Deshalb fährt man die Intensität der Elektronenquelle soweit zurück, dass zunächst mal nur 1 Elektron durch den Apparat läuft. Dieses eine Elektron wird natürlich nicht das vollständige Interferenzmuster erzeugen, sondern an genau einer Stelle im Detektor ein Signal erzeugen. Dies belegt natürlich genau den Teilchencharakter der Elektronen. Erst wenn man

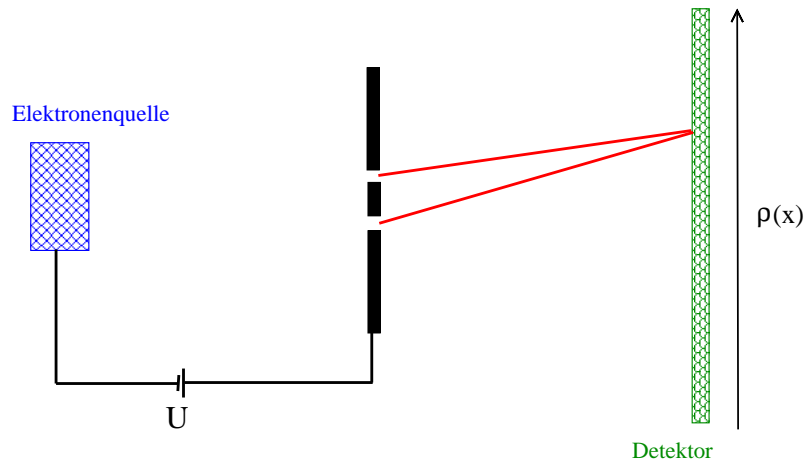


Abbildung 4.4: Schematischer Aufbau eines Doppelspaltexperiments für Elektronen

dieses Experiment mit jeweils einem Elektron viele Male wiederholt hat, ergibt sich für die aufsummierte Verteilung der Elektronen am Detektorschirm die Intensitätsverteilung des Interferenzmusters.

Das Interferenzmuster, d.h. das Betragsquadrat der interferierenden Materiewelle mit den 2 Teilstrahlen, die durch beide Spalte propagieren, repräsentiert die Wahrscheinlichkeitsverteilung der Auftrefforte für die einzelnen Elektronen. Bezeichnet  $r_1$  die Wegstrecke vom Spalt 1 zum Auftreffort  $x$  und  $r_2$  die entsprechende Wegstrecke vom Spalt 2, so ist diese Wahrscheinlichkeit gegeben durch

$$\rho(x) \sim \left| e^{ikr_1} + e^{ikr_2} \right|^2 \quad (4.16)$$

wobei wir der Einfachheit halber angenommen haben, dass die Anteile der Elektronenstrahlen durch die beiden Spalte identisch sind. Die Wellenzahl  $k$  ist entsprechend der de Broglie Wellenlänge in (4.15) gegeben durch

$$k = \frac{2\pi}{\lambda} = \frac{|\vec{p}|}{\hbar} = \frac{\sqrt{2m_e E}}{\hbar} .$$

Beachtenswert ist dabei, dass also auch bei jeweils nur 1 Elektron in der Apparatur, die Welleneigenschaft auftritt. Das Elektron interferiert mit sich selbst und man kann nicht sagen, durch welchen Spalt das Elektron geflogen ist. Erst durch den Nachweis am Detektor verliert das Elektron diesen Wellencharakter. Durch die Messung des Auftrefforts geht die Welleneigenschaft verloren. Man sagt auch:

Die Wellenfunktion des Elektrons kollabiert durch die Messung zu einem Punktteilchen am Ort  $x$ .

Die Wahrscheinlichkeitsverteilung für die einzelnen Zielorte  $x$  ist dabei durch  $\rho(x)$  aus (4.16) gegeben.

Nun kann man den Aufbau so umgestalten, dass es prinzipiell auch möglich ist zu bestimmen, ob ein Elektron durch den oberen oder den unteren Spalt geflogen ist. In diesem

Fall sollte also die Wellenfunktion bereits in der Spaltebene kollabieren und ein Elektron entweder durch Spalt 1 oder durch Spalt 2 treten.

Bei einem solchen Aufbau ergibt sich in der Tat eine Wahrscheinlichkeitsverteilung als Summe der Prozess mit jeweils einem Spalt, also

$$\rho(x) \sim |e^{ikr_1}|^2 + |e^{ikr_2}|^2. \quad (4.17)$$

Dabei ist es natürlich irrelevant, ob man die Messung wirklich durchführt. Allein die Möglichkeit einer solchen Messung zertört die Interferenz, zerstört die Kohärenz der beiden Teilstrahlen.

### 4.2.5 Stationäre Zustände

Man stellt im Experiment fest, dass Atome und Moleküle bestimmte, für das Atom (Molekül) charakteristische Resonanzfrequenzen besitzen, bei denen sie elektromagnetische Strahlung absorbieren. Offenbar kann man also Atome durch die Einstrahlung einer bestimmten Frequenz  $\nu$  aus dem Grundzustand in einen angeregten Zustand versetzen, der relativ zum Ausgangszustand der Energie  $E_i$  die Energie

$$E_f = E_i + h\nu,$$

besitzt. Durch die Absorption eines Photons der Energie  $h\nu$  wird das Atom aus seinem Grundzustand in einen angeregten Zustand versetzt. Natürlich ist auch der Umkehrprozess möglich: Das angeregte Atom gibt seine Energie wieder in Form eines Photons mit gleicher Energie  $h\nu$  ab und geht dabei wieder in den Grundzustand zurück. Dies erklärt, dass die Resonanzfrequenzen im Absorptionsspektrum denen im Emissionsspektrum entsprechen.

Das Erstaunliche an diesem Phänomen ist die Tatsache, dass es nur bestimmte, diskrete Resonanzfrequenzen gibt, was ja gleichbedeutend damit ist, dass es auch nur diskrete Werte für die Anregungsenergien eines solchen Atoms

$$E_f - E_i = h\nu$$

gibt. Dieser Zusammenhang wurde zuerst vom dänischen Physiker Niels Bohr formuliert und führte ihn zum Bohrschen Atommodell, bei dem die Elektronen den positive geladenen Atomkern umkreisen. Dabei sind aber nur bestimmte Bahnen erlaubt, die dann genau zu den diskreten Energien der Anregungszustände gehören.

Im Rahmen der Klassischen Physik stellt sich ein System aus einer positiven Kernladung, die durch ein Elektron umkreist wird, als eine zeitlich oszillierender Dipol dar, der kontinuierlich Ladung abstrahlen sollte, wobei dann die Umlaufamplitude immer kleiner wird: Das Elektron stürzt in den Atomkern. Das Bohrsche Modell entspricht also stationären Zuständen bei diskreten Energien des Systems.

Neben der Diskretisierung des Energiespektrums von Atomen, gibt es auch andere Beobachtungsgrößen, wie z.B. der Drehimpuls von Teilchen in Zentralfeldern, die nur in bestimmten "Quantenschritten" verändert werden können. Man spricht auch von Quantensprüngen. Im Gegensatz zur Umgangssprache, wo ja ein Quantensprung eine besonders massive Änderung bezeichnen soll, sind aber die Quantensprünge in der Quantenmechanik die kleinsten überhaupt möglichen Änderungen einer Observablen.

### 4.3 Zusammenfassung

In diesem Abschnitt werden die experimentellen Hinweise auf die Quantenmechanik zusammengefasst und einige erste Konsequenzen diskutiert. Dabei werden wir uns als Beispiel immer wieder am Doppelspaltexperiment orientieren, bei dem z.B. einzelne Elektronen oder aber auch die Photonen einer elektromagnetischen Welle einen Doppelspalt durchlaufen und anschließend auf einem Schirm mit ihrem genauen Auftreffpunkt detektiert werden.

- Sowohl die elektromagnetischen Wellen als aber auch Teilchen auf der atomaren Größenskala werden durch eine Wellenfunktion beschrieben. Dies ist im Allgemeinen eine komplexwertige Funktion  $\Phi(\vec{r}, t)$ , die von einem Ortsvektor und der Zeit  $t$  abhängt. Als Spezialfall verweisen wir auf die ebene Welle, bei der diese Wellenfunktion die Form

$$\Phi(\vec{r}, t) = \exp\left(i\left(\vec{k}\vec{r} - \omega t\right)\right),$$

annimmt. Das Betragsquadrat dieser Wellenfunktion

$$\rho(\vec{r}, t) = \Phi^* \Phi, \quad (4.18)$$

ergibt die Wahrscheinlichkeitsdichte, das Teilchen oder, im Falle des elektromagnetischen Feldes, ein Photon zur Zeit  $t$  am Ort  $\vec{r}$  zu finden. Integriert man diese Wahrscheinlichkeitsdichte über ein bestimmtes Volumen  $\Delta V$  und ein Zeitintervall  $\Delta t$ , so erhält man die Wahrscheinlichkeit, das Teilchen (oder Photon) in diesem Zeitintervall im Volumen  $\Delta V$  vorzufinden. Die Quantenmechanik liefert uns Werkzeuge, die Wellenfunktion  $\Phi(\vec{r}, t)$  zu berechnen. Daraus können wir dann über (4.18) die ugehörige Wahrscheinlichkeitsdichte bestimmen. Damit liefert uns aber die Theorie nur eine Aussage über Wahrscheinlichkeiten.

Führt man das zugehörige Experiment, z.B. den Durchlauf von Teilchen durch einen Doppelspalt, durch, so ergibt sich ein eindeutiges Ergebnis. Im Fall des Doppelspaltexperimentes mit einem einzelnen Elektron wird man das Elektron an einer einzigen Stelle im Detektor beobachten. Es ist also nicht so, dass das Elektron seine Teilchenidentität verliert und sich entsprechend der Wahrscheinlichkeitsverteilung auflöst. Wiederholt man das Experiment mit einem weiteren Elektron, so wird dieses Elektron in der Regel an einer anderen Stelle landen. Erst wenn man das Experiment hinreichend oft wiederholt ergibt sich eine Verteilung der Punkte, an denen Elektronen beobachtet werden, die der theoretischen Vorhersage für die Wahrscheinlichkeitsdichte entspricht. Die Quantenmechanik liefert also Aussagen über die statistische Verteilung von einer ganzen Messreihe. Im Allgemeinen kann sie keine Aussagen über den Ausgang eines einzelnen Messvorganges machen.

Dies ist kein Defizit der Theorie. Nach unserem heutigen Verständnis gibt es keine Theorie, die präzise Vorhersagen machen kann, die über diese statistische Aussage hinausgeht. Die Quantenmechanik liefert also eine optimal mögliche Beschreibung der Vorgänge.

- Interferenzphänomene sind ein Spezifikum von Wellen. Gibt es also, wie im Beispiel des Doppelspaltexperimentes zwei Wege, durch die das Teilchen hindurchlaufen

könnte, so ergibt sich die Gesamtwellenfunktion als Summe von zwei Funktionen

$$\Phi(\vec{r}, t) = \Phi_1(\vec{r}, t) + \Phi_2(\vec{r}, t), \quad (4.19)$$

man spricht auch von einer **Superposition** dieser beiden Wellenfunktionen  $\Phi_1$  und  $\Phi_2$ . Dabei steht in unserem Beispiel  $\Phi_1$  für die Prozesse, bei denen das Teilchen durch den Spalt 1, und  $\Phi_2$  für solche Prozesse, bei denen das Teilchen durch den Spalt 2 läuft. Kann man nun experimentell nicht auflösen, ob das Teilchen durch den Spalt 1 oder 2 gelaufen ist, so ergibt sich die Wahrscheinlichkeitsdichte entsprechend (4.18) als Betragsquadrat der Gesamtwellenfunktion in (4.19) zu

$$\begin{aligned} \rho(\vec{r}, t) &= (\Phi_1 + \Phi_2)^* (\Phi_1 + \Phi_2) \\ &= \Phi_1^* \Phi_1 + \Phi_2^* \Phi_2 + \underbrace{\Phi_1^* \Phi_2 + \Phi_2^* \Phi_1}_{\text{Interferenzterme}}. \end{aligned} \quad (4.20)$$

Die speziell gekennzeichneten Interferenzterme erzeugen die Strukturen, die auch im Experiment als typische Interferenzstreifen beobachtet werden.

Führt man das Experiment aber nun so aus, dass man für jedes Teilchen, das den Doppelspalt durchläuft, genau angeben kann, ob es durch den Spalt 1 oder durch den Spalt 2 hindurchgegangen ist, so entspricht dieses Experiment natürlich genau dem, bei dem bei allen Teilchen, die durch einen Spalt hindurchgegangen sind, den anderen zugedeckt hat. Die gesamte Verteilung ergibt sich also als Summe der Verteilung von zwei Experimenten am jeweiligen Einzelspalt

$$\begin{aligned} \rho(\vec{r}, t) &= \rho_1 + \rho_2 \\ &= \Phi_1^* \Phi_1 + \Phi_2^* \Phi_2. \end{aligned} \quad (4.21)$$

In diesem Fall werden die beiden Wellenfunktionen inkohärent überlagert und es kommt nicht zu den charakteristischen Interferenzstreifen. Dazu werden die Wellenfunktionen der beiden Telexperimente quadriert und dann die Wahrscheinlichkeiten addiert. Dies entspricht dem Durchgang von inkohärenten elektromagnetischen Wellen durch einen Doppelspalt.

Dabei ist es natürlich nicht entscheidend, ob die einzelnen Teilchen wirklich bei dem Durchgang beobachtet werden. Die Kohärenz wird dadurch aufgehoben, dass prinzipiell mit dem vorgegebenen Experimentaufbau beobachtet werden könnte, durch welchen Spalt das Teilchen hindurchgegangen ist. Man kann sich ja z.B. vorstellen, dass das Experiment so durchgeführt wird, dass die Information durch welchen Spalt das Teilchen gelaufen ist, in einem Speichermedium registriert wird. Die Frage ob nun bei den beobachteten Auftreffpunkten am Beobachtungsschirm ein Interferenzmuster entsprechend (4.20) auftritt oder (4.21) nicht, kann ja nicht davon abhängen ob die auf dem Medium gespeicherte Information darüber welchen Spalt die einzelnen Teilchen durchlaufen haben, ausgewertet wird oder nicht.

Dies Überlegungen mache deutlich, dass der experimentelle Aufbau das Ergebnis des Experimentes beeinflusst. Ob man am Doppelspalt eine Interferenz beobachtet oder nicht hängt davon ab, ob man prinzipiell feststellen kann durch welchen Spalt die einzelnen Teilchen oder Phononen laufen. Der Experimentaufbau beeinflusst also das Ergebnis. Dadurch wird natürlich eines der Grundprinzipien der Naturwissenschaften, dass experimentelle Beobachtungen objektiv also unabhängig von der

Beobachtung ablaufen relativiert. Häufig ist es nicht leicht festzustellen, wie weit eine Kohärenz durch Beobachtungsmöglichkeit gestört wird.

- Nach den bisher gemachten Bemerkungen könnte der Eindruck entstehen, dass die Quantenmechanik im Vergleich zur Klassischen Mechanik eine “schwächere Theorie” darstellt. Man könnte den Eindruck gewinnen, dass wir in einzelnen exotischen Bereichen einfach noch nicht genug wissen und deshalb nur sehr vage Aussagen machen können. In Wahrheit ist es aber eher anders herum: Die Quantenmechanik ist die allgemein gültige Theorie. In einem noch näher zu bestimmenden Grenzfall reduziert sich die Quantenmechanik auf die Klassische Mechanik.

Dies ist also genau so wie bei dem Verhältnis der Speziellen Relativitätstheorie zur Newtonschen Mechanik: Die Relativitätstheorie besitzt die allgemeiner Gültigkeit und reduziert sich im Falle kleiner Geschwindigkeiten auf die Newtonsche Mechanik.

So kann man natürlich auch nicht erwarten, dass wir die Quantenmechanik aus der Klassischen Mechanik herleiten können. Vielmehr müssen wir versuchen, die Gesetze der Klassischen Mechanik als eine Näherung aus der Quantenmechanik herzuleiten, wobei diese Näherung bei den Verhältnissen, bei denen die Beschreibung der Klassischen Mechanik ausreicht, offensichtlich sehr gut ist.

- Warum reduziert sich die Quantenmechanik auf die Klassische Mechanik, wenn man die Bewegung von makroskopischen Körpern betrachtet? Muss ich, wenn ich durch eine Doppeltür (also einen Doppelspalt) gehe, befürchten, dass ich nur in solche Richtungen gehen kann, die mit dem Interferenzmuster an einem Doppelspalt verträglich sind? Muss ich also auch meine Bewegung im Rahmen der Quantenmechanik über die Bestimmung von Wellenfunktionen beschreiben? Die Antwort auf diese Frage ist natürlich ein klares “Nein”, in gewisser Weise ein doppeltes Nein:

Einmal sind Interferenzphänomene, die mit der Wellenfunktion der Quantenmechanik zu tun haben, nur dann beobachtbar, wenn man Messungen vornimmt mit einer Auflösung in der Ortsbestimmung  $\Delta x$ , die etwa vergleichbar ist mit der de-Broglie Wellenlänge des betrachteten Teilchens. Zur Abschätzung dieser Wellenlängen betrachten wir zwei unterschiedliche Bewegungen von Punktteilchen:

1. Die Bewegung eines Elektrons, das durch eine Beschleunigungsspannung  $U$  von 1 Volt aus der Ruhelage in seine Endgeschwindigkeit gebracht worden ist. Es gilt also die Energiebilanz

$$eU = E = \frac{p^2}{2m} = 1 \text{ eV}, \quad (4.22)$$

wobei  $e$  der Betrag der Ladung des Elektrons ist,  $m$  seine Masse und  $E$  die Energie, die zu Beginn der Beschleunigung in Form der potenziellen Energie  $eU$  vorliegt, am Ende in Form der kinetischen Energie mit  $p$  dem Betrag des Impulses. Für den Impuls gilt also

$$p = \sqrt{2mE} = \hbar k, \quad (4.23)$$

mit dem Planckschen Wirkungsquantum

$$\hbar = \frac{h}{2\pi} = 1.05 \dots 10^{-34} \text{ Js}, \quad (4.24)$$

also in Einheiten Joule (J) mal Sekunde (s). Die Wellenzahl  $k$  ist mit der Wellenlänge  $\lambda$  verknüpft über

$$k = \frac{2\pi}{\lambda}.$$

Für die Umrechnung ist es häufig nützlich, nicht den Zahlenwert für  $\hbar$  sondern den für das Produkt aus dem Planckschen Wirkungsquantum und der Lichtgeschwindigkeit  $c$

$$\bar{c} = 197.31 \dots \text{ MeV fm} \quad (4.25)$$

Berücksichtigt man nämlich, dass die Ruhenergie eines Elektrons gegeben ist durch

$$mc^2 = 0.511 \text{ MeV},$$

ergibt sich aus (4.23) und (4.22) für die Wellenzahl

$$k = \frac{\sqrt{2mc^2E}}{\hbar c} = \frac{1}{197000 \text{ fm}}$$

die Wellenlänge  $\lambda$  ist also von der Größenordnung von 200000 fm oder  $2 \cdot 10^{-10}$  Meter. Wird die Energie  $E$  vergrößert wird die Wellenzahl  $k$  größer und die zugehörige Wellenlänge kleiner. Um ein sehr gutes Auflösungsvermögen zu bekommen benötigt man deshalb Teilchenbeschleuniger mit großer Energie.

2. Als zweites Beispiel betrachten wir eine kleine Bleikugel mit einem Gewicht von 1 Gramm, was gleichbedeutend ist mit

$$mc^2 = 5.5 \cdot 10^{26} \text{ MeV}.$$

Lässt man diese Bleikugel aus der Ruhelage im Schwerfeld der Erde um 1 Meter fallen, so besitzt sie eine Energie von etwa  $10^{-3}$  J oder  $6.242 \cdot 10^9$  MeV. Daraus ergibt sich mit (4.23) eine Wellenlänge  $\lambda$  von etwa  $10^{-29}$  cm. Wellenphänomene würden also bei der Bewegung dieser Bleikugel erst beobachtet, wenn man die Position bis auf  $\Delta x$  von  $10^{-29}$  cm genau bestimmen würde, eine Messgenauigkeit, die unerreichbar ist. Ähnlich winzige de-Broglie Wellenlängen ergeben sich für alle makroskopischen Bewegungen

Der erste Grund dafür, dass die Bewegung makroskopischer Teilchen keine Wellenphänomene erkennen lässt, sind also die extrem kleinen Wellenlängen, die mit dieser Bewegung verknüpft sind. Der zweite Grund für die fehlende Quanteninterferenz liegt darin, dass makroskopische Bewegung einer ständigen Beobachtungsmöglichkeit unterliegen. Damit wird die für die Interferenz notwendige Kohärenz zertsört.

Insbesondere die Tatsachen, dass die Quantenmechanik nur Aussagen über Wahrscheinlichkeiten liefert (siehe den ersten Punkt in unserer Zusammenstellung) und der Einfluss des Experimentaufbaus auf das Ergebnis der Untersuchung haben viele Wissenschaftler zu Beginn des 20. Jahrhunderts in eine erkenntnistheoretische Krise geführt. Ein deutliches Beispiel für diese Verunsicherung kommt in der Äußerung von Albert Einstein zum Ausdruck:

Alle meine Versuche, die theoretischen Grundlagen der Physik dieser neuen Art von Wissen anzupassen, haben völlig versagt. Es war als ob mir der Boden unter den Füßen weggezogen würde, mit keinem Fundament irgendwo in Sicht, auf dem man hätte bauen können. (zitiert aus R.A. Schilpp: Albert Einstein. Philosoph - Scientist; Evanston Illinois 1945)

Wie sehr die “Revolution der Quantenmechanik” die Physiker erschüttert hat wird auch belegt durch ein Zitat aus Niels Bohr: “Atomtheorie und Naturbeschreibung” (Springer Verlag Berlin 1931)

Die große Erweiterung unserer Erfahrung in jüngster Zeit hat die Unzulänglichkeit unserer einfachen mechanischen Begriffe ans Licht gebracht und als Folge davon die Fundamente erschüttert, auf denen die übliche Interpretation der Beobachtungen basierte.

Wir werden in den folgenden Abschnitten genauer darstellen, welche Eigenheiten der Quantenmechanik das Weltbild der Naturwissenschaften so revolutioniert haben. Es sei aber schon an dieser Stelle betont, dass die Vorhersagen der Quantenmechanik absolut zuverlässig sind. Es gibt zur Zeit keine Evidenz dafür, dass die quantenmechanische Beschreibung in Frage gestellt werden muss.



# Kapitel 5

## Wellenmechanik eindimensionaler Systeme

In diesem Kapitel sollen einige Grundbegriffe der Quantenmechanik eingeführt und an einfachen Beispielen erläutert werden. Bei diesen Beispielen werden wir uns meistens auf die Beschreibung von einem Teilchen der Masse  $m$  und seiner Bewegung in einer Raumrichtung, charakterisiert durch die Koordinate  $x$ , beschränken.

### 5.1 Wellenpakete und Erwartungswerte

Wie wir bereits in der Diskussion in der Einleitung gesehen haben, wird der Aufenthaltsort und die Bewegung eines Teilchens (Masse  $m$ , eine Raumrichtung  $x$ ) in der Quantenmechanik beschrieben durch eine Wahrscheinlichkeitsamplitude,  $\Psi(x, t)$ , die wir häufig auch als Wellenfunktion bezeichnen werden. Besitzt ein solches Teilchen den Impuls  $p_0 = \hbar k_0$  und die Energie  $E_0 = \hbar\omega$ , so hat diese Wellenfunktion die Form einer ebenen Welle (siehe auch die Diskussion der Teilchen am Doppelspalt in der Einleitung)

$$\Psi(x, t) = Ae^{i(k_0x - \omega t)}. \quad (5.1)$$

Die Wahrscheinlichkeitsdichte,  $\rho(x, t)$ , das Teilchen zur Zeit  $t$  am Ort  $x$  zu finden, berechnet sich dann als das Betragsquadrat dieser Wellenfunktion

$$\rho(x, t) = \Psi^* \Psi = |A|^2. \quad (5.2)$$

Wenn also die Amplitude  $A$  in der ebenen Welle (5.1) eine Konstante ist, so ist auch diese Wahrscheinlichkeitsdichte eine Konstante unabhängig von Ort und Zeit. Wir wollen zunächst einmal versuchen, eine Wellenfunktion zu finden, mit der das Teilchen zur Zeit  $t = 0$  in der Nähe des Koordinatenursprungs  $x = 0$  lokalisiert wird. Eine einfache Möglichkeit ist es, eine  $x$ -abhängige Amplitudenfunktion  $A(x)$  anzunehmen, sodass die Wahrscheinlichkeitsdichte  $\rho(x)$  die Form einer Gaußverteilung annimmt

$$\rho(x) = |A(x)|^2 = N^2 e^{-\frac{x^2}{a^2}}. \quad (5.3)$$

Diese Wahrscheinlichkeitsverteilung ist in Abb. 5.1 dargestellt. Die Normierungskonstante  $N$  soll dabei so gewählt werden, dass die Wahrscheinlichkeit das Teilchen irgendwo zu

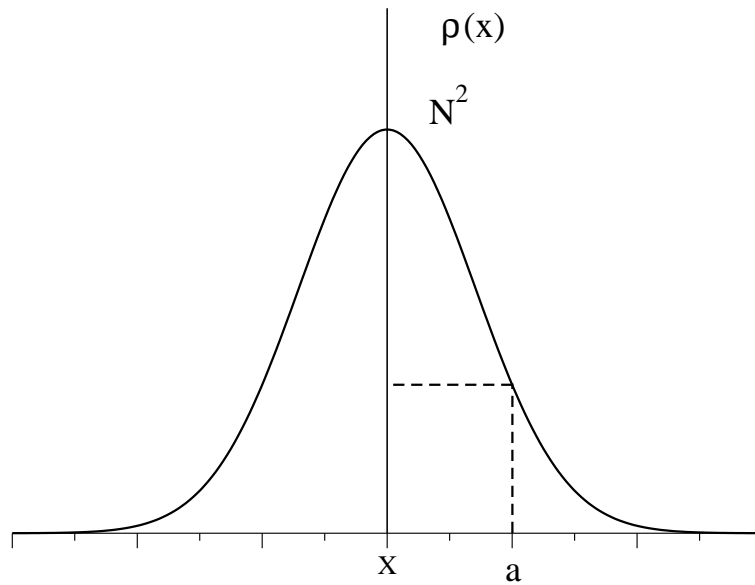


Abbildung 5.1: Wahrscheinlichkeitsverteilung einer Gaußfunktion nach (5.3)

finden, d.h. die Wahrscheinlichkeitsdichte integriert über alle möglichen Aufenthaltsorte  $x$ , den Wert 1 ergibt

$$\begin{aligned}
 1 &= \int_{-\infty}^{\infty} dx \rho(x) \\
 &= N^2 \int_{-\infty}^{\infty} dx e^{-\frac{x^2}{a^2}} \\
 &= N^2 a \sqrt{\pi}
 \end{aligned} \tag{5.4}$$

Die Wellenfunktion die zu dieser Wahrscheinlichkeitsverteilung führt hat demnach die Form

$$\Psi(x) = \frac{1}{\sqrt{a} \sqrt{\pi}} e^{-\frac{x^2}{2a^2}} e^{ik_0 x} \tag{5.5}$$

Was bedeutet die so definierte Wellenfunktion beziehungsweise die dazugehörige Wahrscheinlichkeitsverteilung  $\rho(x)$ ? Wir beschreiben mit dieser das System eines Punktteilchens, das sich irgendwo auf der  $x$ -Achse aufhalten kann. Führen wir eine Messung der  $x$ -Koordinate durch, so wird jede Messung einen Wert für diese  $x$ -Koordinate liefern. Betrachten wir dann die Häufigkeit, mit der wir einen Messwert für  $x$  in einem bestimmten Intervall der Länge  $dx$ :  $x \in [x_0 - dx/2, x_0 + dx/2]$  so sollte diese Häufigkeitsverteilung  $\rho(x_0)$  der Wahrscheinlichkeitsverteilung entsprechen, die durch das Betragsquadrat der Wellenfunktion gegeben ist. Das Ergebnis jeder Einzelmessung ist dabei nicht vorhersagbar, die Häufigkeitsverteilung der gesamten Messreihe hingegen wird durch die Wellenfunktion und die daraus bestimmte Wahrscheinlichkeitsverteilung gegeben.

Für das Beispiel der Verteilung  $\rho(x)$ , die in (5.3) definiert und in Abb. 5.1 dargestellt ist bedeutet dies, dass wir Messergebnisse für die Aufenthaltsorte erwarten, die in der Nähe von  $x = 0$  liegen. Die Wahrscheinlichkeit ist am größten für den Wert  $x = 0$ , nämlich gerade  $N^2$ . Diese Wahrscheinlichkeitsdichte fällt zu positiven und negativen Werten von  $x$  ab und zwar so, dass bei  $x = \pm a$ , die Wahrscheinlichkeit nur noch  $N^2/e$  beträgt. Der Parameter  $a$  ist also ein Maß für die Streubreite der Messergebnisse.

Liefert eine Messung einer Observablen eine Häufigkeitsverteilung von Ergebnissen, wie etwa die gerade diskutierte Messung der x-Koodinate, so ist natürlich die wichtigste Information zunächst einmal der Mittelwert dieser Verteilung. Am Beispiel der Messung der x-Koordinate würden wir diesen Mittelwert berechnen nach

$$\begin{aligned}
 \bar{x} &= \int_{-\infty}^{\infty} dx x \rho(x) \\
 &= \int_{-\infty}^{\infty} dx \Psi^*(x) x \Psi(x) \\
 &= \frac{1}{a\sqrt{\pi}} \int_{-\infty}^{\infty} dx x e^{-\frac{x^2}{a^2}} \\
 &= 0
 \end{aligned} \tag{5.6}$$

Natürlich kann man mit der Wahrscheinlichdichte  $\rho(x)$  auch die Mittelwerte für andere Observable, die Funktionen  $f(x)$  sind bestimmen. Wir definieren deshalb allgemein den Mittelwert oder Erwartungswert einer Observablen mit:

**Definition:**

Der **Erwartungswert** einer Observablen  $f(x)$  bezeichnet den Mittelwert dieser Observablen bezogen auf eine Wahrscheinlichkeitsverteilung  $\rho(x) = \Psi^*\Psi$ . Dieser Erwartungswert berechnet sich also gemäß:

$$\langle f \rangle = \langle \Psi | f | \Psi \rangle = \int_{-\infty}^{\infty} dx \Psi^*(x) f(x) \Psi(x) \tag{5.7}$$

Bei dieser Definitionsgleichung ist natürlich die Reihenfolge der Faktoren des Integranden  $\Psi^*$ ,  $f$  und  $\Psi$  irrelevant. Wir werden jedoch unten sehen, dass diese Schreibweise von links nach rechts auch für Observable geeignet ist, die von Ort und Impuls abhängen und wollen deshalb bereits hier diese Konvention für die Reihenfolge wählen.

Als Beispiel für die Berechnung einer solchen Observablen kann man sich vorstellen, dass das Teilchen, dessen Aufenthaltswahrscheinlichkeit durch die Wellenfunktion  $\Psi$  beschrieben wird, sich in einem Potential  $V = f(x)$  befindet. Der Erwartungswert von  $V = f$  liefert also in diesem Fall den Mittelwert der potentiellen Energie, die dieses Teilchen besitzt.

Neben dem Mittelwert ist aber natürlich auch von Interesse, wie stark Einzelmessungen typisch von diesem Mittelwert abweichen. Wir definieren also eine Abweichungsfunktion

$$\Delta(f(x)) = f(x) - \langle f \rangle$$

die für jede Messung angibt, wie weit der Ergebniswert dieser Einzelmessung  $f(x)$  vom Erwartungswert  $\langle f \rangle$  abweicht. Würde man einfach den Erwartungswert für diese Funktion  $\Delta(f(x))$  berechnen, so ergäbe sich der Wert 0: es zeigen gleich viele Messungen eine positive Abweichung  $\Delta f$  wie eine entsprechende negative. Ein Maß für die mittlere Abweichung ist aber das mittlere Schwankungsquadrat, das wir definieren

**Definition:**

Das **Mittlere Schwankungsquadrat**  $(\Delta f)^2$  für eine Observable  $f$  bezogen auf eine Wellenfunktion  $\Psi$  berechnet sich als Erwartungswert

$$(\Delta f)^2 = \langle \Psi | (f - \langle f \rangle)^2 | \Psi \rangle \quad (5.8)$$

Diese Definition des mittleren Schwankungsquadrat lässt sich auch leicht umschreiben auf die Form

$$\begin{aligned} (\Delta f)^2 &= \langle (f^2 - 2f \langle f \rangle + \langle f \rangle^2) \rangle \\ &= \underbrace{\int dx \Psi^* f(x)^2 \Psi}_{=\langle f^2 \rangle} - 2 \langle f \rangle \underbrace{\int dx \Psi^* f(x) \Psi}_{=\langle f \rangle} + \langle f \rangle^2 \underbrace{\int dx \Psi^* \Psi}_{=1} \\ &= \langle f^2 \rangle - \langle f \rangle^2 \end{aligned} \quad (5.9)$$

Als Anwendungsbeispiel wollen wir das mittlere Schwankungsquadrat der Observable  $f(x) = x$  berechnen für die in (5.5) definierte Wellenfunktion. Dazu benötigen wir den Wert der Integrale

$$\int_{-\infty}^{\infty} dx x^k e^{-\frac{x^2}{a^2}} = \begin{cases} 0 & , \text{ für } k \text{ ungerade natürliche Zahl} \\ a^{k+1} \Gamma\left(\frac{k+1}{2}\right) & , \text{ für } k \text{ gerade natürliche Zahl} \end{cases} \quad (5.10)$$

Zur Vervollständigung benötigen wir noch den Wert der  $\Gamma$  Funktion für halbzahlige Argumente. Diese sind gegeben durch die rekursiven Gleichungen

$$\begin{aligned} \Gamma\left(\frac{1}{2}\right) &= \sqrt{\pi} \\ \Gamma(k+1) &= k\Gamma(k) \end{aligned} \quad (5.11)$$

Diese Gleichungen beinhalten auch das Ergebnis für das Normierungsintegral (5.4) ( $k = 0$ ) und den bereits in (5.6) berechneten Erwartungswert  $\langle x \rangle = 0$  (Fall  $k = 1$ ). Zur Berechnung des mittleren Schwankungsquadrates benötigen wir nur noch den Erwartungswert

$$\begin{aligned} \langle x^2 \rangle &= \frac{1}{a\sqrt{\pi}} \int_{-\infty}^{\infty} dx x^2 e^{-\frac{x^2}{a^2}} \\ &= \frac{1}{2a\sqrt{\pi}} a^3 \sqrt{\pi} \end{aligned}$$

Mit (5.9) ergibt sich also in diesem Fall

$$(\Delta x)^2 = \langle x^2 \rangle - \langle x \rangle^2 = \frac{1}{2} a^2. \quad (5.12)$$

Der bereits oben diskutierte Halbwertsparameter  $a$  in der Gaußverteilung quadriert entspricht also dem mittleren Schwankungsquadrat.

Die Wellenfunktion  $\Psi(x)$  aus (5.5), die wir konstruiert haben, um ein lokalisiertes Teilchen zu beschreiben, entspricht aber nicht mehr der Wellenfunktion eines Teilchens mit scharf

definiertem Impuls  $p = \hbar k_0$  aus (5.1), da ja nun die Amplitude  $A$  von der Ortskoordinate  $x$  abhängt. Wir werden sehen, dass die Wellenfunktion des lokalisierten Teilchens (5.5) sich darstellen lässt als eine Summe oder Superposition von ebenen Wellen mit verschiedenen Impulsen  $p = \hbar k$ . Allgemein gilt nämlich der folgende Satz der **Fouriertransformation**

**Satz:**

Eine Funktion  $\Psi(x)$  kann dargestellt werden als

$$\Psi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk C(k) e^{ikx} \quad (5.13)$$

dabei kann die Gewichtsfunktion  $C(k)$  berechnet werden nach

$$C(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx \Psi(x) e^{-ikx} \quad (5.14)$$

Zum Beweis dieser Aussage:

$$\begin{aligned} \Psi(x) &= \int_{-\infty}^{\infty} dx' \delta(x - x') \Psi(x') \\ &= \int_{-\infty}^{\infty} dx' \Psi(x') \underbrace{\frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ik(x-x')}}_{=\delta(x-x')} \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk e^{ikx} \underbrace{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx' \Psi(x') e^{-ikx'}}_{=C(k)} \end{aligned}$$

Wir sehen also aus der Darstellung (5.13), dass die Wellenfunktion als eine Summe, beziehungsweise ein Integral von Wellenfunktionen mit verschiedenen Impulsen  $p = \hbar k$  dargestellt wird. Die Funktion  $C(k)$  stellt dabei die Wahrscheinlichkeitsamplitude für die jeweiligen Impulse dar. Wir können also ein Teilchen nur dann mit einem scharf definierten Impuls beschreiben, wenn es durch die ebene Welle mit der entsprechenden Wellenzahl beschrieben wird. In diesem Fall ist die Wahrscheinlichkeitsdichte als Funktion der Ortskoordinate aber eine Konstante. Das Teilchen ist in keiner Weise lokalisiert und hält sich an allen Orten mit gleicher Wahrscheinlichkeit auf. Versuchen wir aber das Teilchen durch eine ortsabhängige Amplitude  $A(x)$  zu lokalisieren, so stellt sich die Wellenfunktion nicht mehr durch eine ebene Welle mit einer, scharf definierten Impuls dar, sondern sie ist eine Summe von ebenen Wellen mit verschiedenen Impulskomponenten.

In der klassischen Mechanik kann der Ort und der Impuls eines Teilchens jederzeit genau bestimmt werden. Beschreibt man aber die Bewegung eines Teilchens in der Quantenmechanik mit Hilfe einer Wellenfunktion, so können Ort und Impuls immer nur in Form einer Wahrscheinlichkeitsverteilung angegeben werden, wobei man für eine präzisere Festlegung des Ortes den Preis zahlen muss, dass der Impuls ungenauer bestimmt ist, während eine genauere Festlegung der Impulsverteilung zu einer Unschärfe bei der Bestimmung des Ortes führt. Wir werden später sehen, dass dies nicht etwa die Folge einer ungenauen Beschreibung durch die Wellenfunktionen ist. Vielmehr gehört es zu den Grundprinzipien

der Quantenmechanik, dass zueinander konjugierte Variable, wie z.B. Ort und Impuls, nie gleichzeitig scharf festgelegt oder gemessen werden können.

Als **Beispiel** soll nun die Impulsverteilung  $C(k)$  für die Wellenfunktion aus (5.5) berechnet werden. Wir setzen dazu (5.5) in (5.14) ein und erhalten

$$\begin{aligned} C(k) &= \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{a\sqrt{\pi}}} \int_{-\infty}^{\infty} dx e^{-\frac{x^2}{2a^2} + i(k_0 - k)x} \\ &= \frac{1}{\sqrt{2\pi a\sqrt{\pi}}} e^{-\frac{a^2}{2}(k_0 - k)^2} \int_{-\infty}^{\infty} dx e^{-\frac{x^2}{2a^2} + i(k_0 - k)x} e^{\frac{a^2}{2}(k_0 - k)^2} \\ &= \frac{1}{\sqrt{2\pi a\sqrt{\pi}}} e^{-\frac{a^2}{2}(k_0 - k)^2} \int_{-\infty}^{\infty} dx e^{-\left(\frac{x}{\sqrt{2a}} - i\frac{a}{\sqrt{2}}(k_0 - k)\right)^2} \end{aligned}$$

Zur weiteren Berechnung substituieren wir nun

$$z = \frac{x}{\sqrt{2a}} - i\frac{a}{\sqrt{2}}(k_0 - k)$$

sodass  $dx = \sqrt{2a} dz$ . Damit berechnet sich

$$\begin{aligned} C(k) &= \frac{1}{\sqrt{2\pi a\sqrt{\pi}}} e^{-\frac{a^2}{2}(k_0 - k)^2} \int_{-\infty}^{\infty} dz \sqrt{2a} e^{-z^2} \\ &= \frac{\sqrt{2a}\sqrt{\pi}}{\sqrt{2\pi a\sqrt{\pi}}} e^{-\frac{a^2}{2}(k_0 - k)^2} . \end{aligned}$$

Bei dem Übergang zur letzten Zeile haben wir (5.10) ausgenutzt. Mit dieser Wahrscheinlichkeitsamplitude  $C(k)$  können wir dann die Wahrscheinlichkeitsverteilung  $\omega(k)$  definieren, das betrachtete Teilchen mit dem Impuls  $p = \hbar k$  vorzufinden:

$$\omega(k) = C^*(k)C(k) \quad (5.15)$$

Diese Definition der Wahrscheinlichkeitsverteilung für die Impulse erfüllt die notwendigen Voraussetzungen für eine Wahrscheinlichkeitsverteilung:

- $\omega(k)$  ist positiv definit:  $\omega(k) \geq 0$  für alle  $k$
- $\omega(k)$  ist normiert in der Form

$$\int_{-\infty}^{\infty} \omega(k) dk = 1 .$$

Die Wahrscheinlichkeitsverteilung der Impulse, beziehungsweise der entsprechenden Wellenzahlen  $k$  ist also in unserem Beispiel gegeben durch:

$$\omega(k) = \frac{a}{\sqrt{\pi}} e^{-a^2(k_0 - k)^2} .$$

Dies ist also wie die Wahrscheinlichkeitsverteilung  $\rho(x)$  im Ortsraum (5.3) eine Gaussverteilung, die ein Maximum besitzt bei der Wellenzahl  $k = k_0$  und eine Breite von  $2/a$ .

Dies bedeutet, dass die Wahrscheinlichkeitsverteilung  $\omega(k)$  bei  $k = k_0 \pm 1/a$  jeweils auf den Wert  $\omega(k_0)/e$  also dem  $e$ 'ten Teil des Maximalwertes abgefallen ist.

Mit dieser Wahrscheinlichkeitsverteilung  $\omega(k)$  für die Wellenzahlen kann man Erwartungswerte für dynamische Variable berechnen, die ausschließlich von der Wellenzahl beziehungsweise dem entsprechenden Impuls abhängen. Nehmen wir einmal an, wir wollen den Erwartungswert für eine Reibungskraft die proportional zum Impuls ist,  $f(k) = \alpha k$ , berechnen. Dieser Erwartungswert ergibt sich dann als

$$\begin{aligned} \langle C|f(k)|C \rangle &= \int_{-\infty}^{\infty} dk \omega(k) f(k) & (5.16) \\ &= \frac{a}{\sqrt{\pi}} \int_{-\infty}^{\infty} dk e^{-a^2(k_0-k)^2} \alpha k \\ &= \alpha \frac{a}{\sqrt{\pi}} \int_{-\infty}^{\infty} dq e^{-a^2 q^2} (q + k_0) \\ &= \alpha k_0. & (5.17) \end{aligned}$$

Zur Berechnung dieses Erwartungswertes haben wir die Variablensubstitution  $q = k_0 - k$  durchgeführt und die Ergebnisse von (5.10) benutzt. Für  $\alpha = \hbar$  liefert dieses Beispiel den Erwartungswert für den Impuls  $\langle p \rangle = \hbar k_0$ . Natürlich können wir auch hier das mittlere Schwankungsquadrat für die Impulsverteilung berechnen

$$\begin{aligned} (\Delta p)^2 &= \langle C|(p - \langle p \rangle)^2|C \rangle \\ &= \hbar^2 \frac{a}{\sqrt{\pi}} \int_{-\infty}^{\infty} dk e^{-a^2(k_0-k)^2} (k - k_0)^2 \\ &= \frac{\hbar^2}{2a^2} & (5.18) \end{aligned}$$

Vergleichen wir dieses mittlere Schwankungsquadrat der Impulse mit dem mittleren Schwankungsquadrat berechnet für die zugehörige Verteilung im Ortsraum (5.12), so ergibt sich

$$\Delta x \Delta p = \frac{1}{2} \hbar. \quad (5.19)$$

Dieses Produkt ist also unabhängig vom Parameter  $a$ . Für große Werte von  $a$  ist die Verteilung im Ortsraum nur schwach lokalisiert, während die Impulsverteilung ein kleines mittleres Schwankungsquadrat aufweist. Bei kleinen Werten von  $a$  ist die Position des betrachteten Teilchens im Ortsraum relativ scharf definiert, während die Verteilung im Impulsraum entsprechend unscharf ist. Wir werden später sehen, dass dieser Zusammenhang zwischen Orts- und Impulsverteilung keine spezielle Eigenschaft des hier betrachteten Beispiels ist. Die Gleichung (5.19) stellt vielmehr einen Spezialfall der Heisenbergschen Unschärferelation dar, die wir allgemein im Kapitel 9 behandeln werden.

Mit den bisher zur Verfügung stehenden Werkzeugen können wir Erwartungswerte für dynamische Variable berechnen, die ausschließlich von der Ortskoordinate des betrachteten Teilchens abhängen indem wir die entsprechenden Funktionen  $f(x)$  gewichtet mit der Wahrscheinlichkeitsdichte  $\rho(x)$  integrieren. Zur Berechnung von dynamischen Variablen, die ausschließlich von der Impulskoordinate abhängen, müssen wir zunächst aus der Wellenfunktion mit Hilfe der Fouriertransformation (5.14) die Wahrscheinlichkeitsamplitude  $C(k)$  bestimmen und können dann über (5.16) den entsprechenden Erwartungswert

berechnen. Wir wollen jetzt zeigen, dass man den Erwartungswert für den Impuls und Funktionen des Impulses aber auch direkt aus der Wellenfunktion  $\Psi(x)$ , also ohne den Umweg über die Fouriertransformation berechnen kann. Es gilt nämlich der folgende

**Satz:**

Der Erwartungswert des Impulses kann direkt aus der Wellenfunktion  $\Psi(x)$  berechnet werden durch:

$$\begin{aligned} \langle p \rangle &= \langle \Psi | \frac{\hbar}{i} \frac{d}{dx} | \Psi \rangle \\ &= \int_{-\infty}^{\infty} dx (\Psi(x))^* \frac{\hbar}{i} \frac{d}{dx} \Psi(x). \end{aligned} \quad (5.20)$$

Zum Beweis dieses Satzes ersetzen wir die Wellenfunktionen  $\Psi$  und  $\Psi^*$  in (5.20) durch die entsprechenden Darstellungen der Fouriertransformationen (5.13). Damit können wir die rechte Seite von (5.20) umschreiben in

$$\int_{-\infty}^{\infty} dx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dq C^*(q) e^{-iqx} \frac{\hbar}{i} \frac{d}{dx} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk C(k) e^{ikx}$$

Die Ableitung der Funktion  $e^{ikx}$  nach  $x$  ist einfach auszuführen und man erhält durch einfache Umschreibung

$$\int_{-\infty}^{\infty} dq C^*(q) \int_{-\infty}^{\infty} dk C(k) \hbar k \underbrace{\frac{1}{2\pi} \int_{-\infty}^{\infty} dx e^{i(k-q)x}}_{=\delta(k-q)}.$$

Damit ergibt sich aber direkt für die rechte Seite von (5.20)

$$\int_{-\infty}^{\infty} dq C^*(q) \int_{-\infty}^{\infty} dk C(k) \hbar k \delta(k-q) = \int_{-\infty}^{\infty} dk C^*(k) \hbar k C(k) = \langle p \rangle$$

womit der Satz bewiesen ist.

Wir definieren also für die Berechnung von Impulserwartungswerten direkt aus der Wellenfunktion  $\Psi(x)$

**Definition**

Der **Impulsoperator** für die Anwendung auf Wellenfunktionen  $\Psi(x)$  ist definiert durch

$$\hat{p} = \frac{\hbar}{i} \frac{d}{dx} \quad (5.21)$$

Erwartungswerte des Impulses lassen sich dann berechnen, indem man diesen Impulsoperator auf die Wellenfunktion  $\Psi$  anwendet und entsprechend (5.20) ausrechnet. Ganz analog ergeben sich auch Erwartungswerte für Potenzen des Impulses  $p^n$  dadurch, dass man den Operator  $\hat{p}$  n-fach anwendet. Wir sind also auch in der Lage Erwartungswerte



für dynamische Variable, die als Funktionen  $g(p)$  definiert sind zu berechnen, indem wir ausführen

$$\langle g \rangle = \int_{-\infty}^{\infty} dx \Psi^*(x) g\left(\frac{\hbar}{i} \frac{d}{dx}\right) \Psi(x).$$

Probleme haben wir aber immer noch bei der Berechnung von Erwartungswerten für dynamische Variable, die sowohl die Ortskoordinate  $x$  wie auch die Impulskoordinate  $p$  enthalten. Betrachten wir als Beispiel den Fall, dass wir das Produkt aus Orts- und Impulskoordinate berechnen wollen. In der klassischen Mechanik ist es natürlich gleich in welcher Reihenfolge wir diese beiden Koordinaten multiplizieren, es gilt ja

$$px = xp.$$

Wollen wir aber nun mit den bisher gelernten Regeln den entsprechenden Erwartungswert berechnen, so gilt

$$\begin{aligned} \langle px \rangle &= \int_{-\infty}^{\infty} dx \Psi^*(x) \frac{\hbar}{i} \frac{d}{dx} x \Psi(x) \\ &= \int_{-\infty}^{\infty} dx \Psi^*(x) \frac{\hbar}{i} \Psi(x) + \int_{-\infty}^{\infty} dx \Psi^*(x) x \frac{\hbar}{i} \frac{d}{dx} \Psi(x) \\ &= \int_{-\infty}^{\infty} dx \Psi^*(x) \frac{\hbar}{i} \Psi(x) + \langle xp \rangle \\ &\neq \langle xp \rangle \end{aligned}$$

Diesen allgemeinen Fall, die Darstellung von dynamischen Variablen, die Orts- und Impulskoordinaten enthalten werden wir später aufgreifen.

### Zusammenfassung

- Ein System wie ein Partikel wird durch eine komplexwertige Wellenfunktion  $\Psi(x, t)$ , beziehungsweise die Fouriertransformierte  $C(k, t)$  beschrieben.
- Aus diesen Funktionen kann die Wahrscheinlichkeitsverteilung, das Teilchen zur Zeit  $t$  am Ort  $x$  zu finden  $\rho(x, t) = \Psi^* \Psi$  berechnet werden. Entsprechend ergibt sich die Wahrscheinlichkeitsverteilung für den Impuls  $p = \hbar k$  als  $\omega(k, t) = C^* C$ .
- Mit diesen Verteilungen können Mittelwerte für die Messung von dynamischen Variablen, die Erwartungswerte berechnet werden mit

$$\begin{aligned} \langle f(x) \rangle &= \int dx \Psi^*(x, t) f(x) \Psi(x, t) \\ \langle g(k) \rangle &= \int dx C^*(k, t) g(k) C(k, t) \end{aligned}$$

- Erwartungswerte für dynamische Variable, die vom Impuls abhängen, können auch direkt aus der Wellenfunktion  $\Psi(x)$  berechnet werden mit Hilfe des Impulsoperators

$$\langle g(p) \rangle = \int dx \Psi^*(x, t) g\left(\frac{\hbar}{i} \frac{d}{dx}\right) \Psi(x, t)$$

Probleme ergeben sich bei dynamischen Variablen, die von Ort und Impuls abhängen.

- Das mittlere Schwankungsquadrat ist ein Maß für die typische Abweichung einer Einzelmessung für eine dynamische Variable  $f$  vom Erwartungswert. Es berechnet sich nach

$$(\Delta f)^2 = \langle (f - \langle f \rangle)^2 \rangle = \langle f^2 \rangle - \langle f \rangle^2$$

## 5.2 Schrödinger Gleichung und Ehrenfest'sches Theorem

In den vorhergehenden Abschnitten haben wir gesehen, wie das System etwa eines Punktteilchens in einer Dimension zu einer fest vorgegebenen Zeit beschrieben wird durch eine Wellenfunktion  $\Psi(x, t)$ . In der Physik interessiert natürlich besonders, wie sich der Zustand des Systems mit der Zeit ändert. Diese zeitliche Entwicklung wird durch die Schrödinger Gleichung beschrieben

$$\hat{H}\Psi(x, t) = i\hbar \frac{d}{dt}\Psi(x, t). \quad (5.22)$$

Dabei steht  $\hat{H}$  für den **Hamiltonoperator**. Zur Definition des Hamiltonoperators erinnern wir uns an die klassische Mechanik, wo die **Hamiltonfunktion**  $H(p, x)$  als eine Funktion der generalisierten Koordinaten, in dem hier betrachteten Beispiel der Bewegung in einer Dimension ist das die Ortskoordinate  $x$ , und den zugehörigen Impuls  $p$  definiert wurde. Nehmen wir als Beispiel die Bewegung eines Teilchens der Masse  $m$  in einem konservativen Kraftfeld, definiert durch das Potential  $V(x)$ , so ist diese Hamiltonfunktion gleich der Energie also der Summe aus kinetischer Energie und potentieller Energie

$$H(p, x) = \frac{p^2}{2m} + V(x)$$

Der Hamiltonoperator, genauer gesagt der Hamiltonoperator in der Ortsdarstellung (andere Darstellungen werden wir später behandeln), ist dann definiert durch die Vorschrift, dass in dieser Hamiltonfunktion die Impulskoordinate  $p$  durch den Impulsoperator aus Gleichung (5.21) ersetzt wird

$$p \implies \hat{p} = \frac{\hbar}{i} \frac{d}{dx} \quad (5.23)$$

In der Schrödingergleichung (5.22) führt also jeder Impulsoperator zu einer Ableitung der Wellenfunktion  $\Psi(x, t)$  nach der Ortskoordinate multipliziert mit dem Faktor  $-i\hbar$ . Andererseits wird jede Koordinate  $x$  in der Hamiltonfunktion durch den Ortsoperator

$$x \implies \hat{x} = x \quad (5.24)$$

Jede Ortskoordinate  $x$  in der Hamiltonfunktion führt also im Hamiltonoperator zu einer Multiplikation mit eben dieser Koordinate. Zur Konkretisierung wollen wir nun annehmen, dass  $V(x) = \alpha x^2$ , das Potential eines harmonischen Oszillators ist. In diesem Fall ist also der Hamiltonoperator gegeben durch

$$\begin{aligned} \hat{H} &= \frac{\hat{p}^2}{2m} + \alpha \hat{x}^2 \\ &= -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \alpha x^2 \end{aligned}$$

Setzen wir diesen Hamiltonoperator in die Schrödingergleichung (5.22) ein,

$$\left\{ -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \alpha x^2 \right\} \Psi(x, t) = i\hbar \frac{d}{dt}\Psi(x, t) \quad (5.25)$$

so sehen wir, dass diese Gleichung eine partielle Differentialgleichung zur Bestimmung der Wellenfunktion  $\Psi(x, t)$  darstellt. Kennen wir die Lösung dieser Gleichung zu einer bestimmten Startzeit  $t_0$ , so liefert uns diese Differentialgleichung auch in eindeutiger Weise die Wellenfunktion für alle Zeiten  $t$ . Die Schrödingergleichung liefert also in der Tat ein Ergebnis für die Entwicklung des betrachteten Systems mit der Zeit. Sie übernimmt also die Funktion, die in der Klassischen Mechanik durch die Bewegungsgleichungen (z.B. Hamiltonsche Bewegungsgleichungen) wahrgenommen wird. Kennt man in der Klassischen Mechanik, die Werte für die generalisierten Koordinaten und Impulse zu einer Zeit  $t_0$ , dann liefert die Lösung der Hamiltonschen Bewegungsgleichungen ein eindeutiges Ergebnis für die Koordinaten und Impulse für alle Zeiten. Die zeitliche Entwicklung des Systems ist also eindeutig beschrieben. Entsprechend liefert die Quantenmechanik für eine vorgegebene Wellenfunktion zur Zeit  $t_0$  ein eindeutiges Ergebnis für die Wellenfunktion zu allen Zeiten.

Die Schrödingergleichung repräsentiert eine der Grundprinzipien der Quantenmechanik. Da die Quantenmechanik eine Theorie sein soll, die umfassender ist als die Klassische Mechanik, können wir nicht erwarten, dass wir diese Grundgleichung aus der Klassischen Mechanik herleiten können. Vielmehr sollten sich die Gesetze der Klassischen Mechanik als ein Spezialfall aus der Quantenmechanik ergeben. Bevor wir diesen Nachweis erbringen, wollen wir uns zunächst einmal klar machen, dass die Schrödingergleichung vernünftige Ergebnisse liefert und plausibel ist.

Dazu betrachten wir zunächst einmal den Fall des Teilchens, das sich frei bewegt. Dies bedeutet, dass wir in (5.25) die Konstante  $\alpha$  in der potentiellen Energie zu Null setzen. Wir können nun verifizieren, dass die Wellenfunktion (5.1) für ein freies Teilchen

$$\Psi(x, t) = Ae^{i(k_0x - \omega t)}$$

genau dann eine Lösung dieser Schrödingergleichung ist, wenn gilt

$$\frac{\hbar^2 k_0^2}{2m} = \hbar\omega$$

Berücksichtigen wir, dass die Wellenzahl  $k_0$  mit dem Impuls des Teilchens  $p_0$  über  $p_0 = \hbar k_0$  verknüpft ist, und dass für die Energie  $E_0 = \hbar\omega$  gilt, so stellt diese Gleichung die Beziehung zwischen Energie und Impuls eines freien Teilchens her. Die Wellenfunktion eines freien Teilchens ist also eine Lösung der Schrödingergleichung (5.25) für den Grenzfal  $\alpha = 0$ .

Es ist aber auch plausibel, dass der Hamiltonoperator eine zeitliche Änderung der Wellenfunktion beschreibt so wie es in der Schrödingergleichung (5.22) zum Ausdruck kommt. In der Klassischen Mechanik betrachtet man kanonische Transformation der generalisierten Koordinaten und Impulse, die den Phasenraum aufspannen. Solche kanonischen Transformationen sind charakterisiert durch Erzeugende Funktionen. Dabei zeigt sich dass der Impuls die Erzeugende Funktion für eine infinitesimale Transformation der zugehörigen Koordinate ist. Bei dem Übergang zur Quantenmechanik wurde die Impulskoordinate "übersetzt" in den Impulsoperator (5.23), die Ableitung nach der zugehörigen Koordinate multipliziert mit  $-i\hbar$ . Analog ist die Hamiltonfunktion in der Klassischen Mechanik die Erzeugende Funktion für einen infinitesimale Transformation der Zeit. Entsprechend ist der Hamiltonoperator der Ableitung nach der Zeit gleichzusetzen, wie das in der Schrödingergleichung geschieht. Ein Unterschied besteht also allein im Vorzeichen des Vorfaktors  $i\hbar$ . Den Zusammenhang zwischen diesen Operatoren der Quantenmechanik und den Erzeugenden Funktionen für Infinitesimale Transformation von Koordinaten und Zeit, werden wir zu einem späteren Zeitpunkt noch vertiefen.

Es soll nun gezeigt werden in welcher Form die Gesetze der Klassischen Mechanik in der Quantenmechanik realisiert sind. Dies wird ausgedrückt durch das **Ehrenfest'sche Theorem**

**Satz:** *Sei die Wellenfunktion  $\Psi$  eine Lösung der Schrödingergleichung. Für die Erwartungswerte von dynamischen Variablen, berechnet mit dieser Wellenfunktion, gelten die Gesetze der Klassischen Mechanik*

An dieser Stelle soll das Ehrenfestsche Theorem zunächst nur bewiesen werden für die Bewegung eines Teilchens in einer Raumrichtung unter dem Einfluss eines konservativen Kraftfeldes. In der Klassischen Mechanik wird diese Bewegung beschrieben durch die Newtonsche Bewegungsgleichung:

$$\frac{dp}{dt} = K = -\frac{dV}{dx}$$

Zum Beweis des Ehrenfest'schen Theorems muss man also zeigen:

$$\frac{d}{dt} \langle \Psi | \hat{p} | \Psi \rangle = - \langle \Psi | \frac{dV}{dx} | \Psi \rangle \quad (5.26)$$

Zunächst betrachten wir die linke Seite dieser Gleichung:

$$\begin{aligned} \frac{d}{dt} \langle \Psi | \hat{p} | \Psi \rangle &= \frac{d}{dt} \int dx \Psi^* \frac{\hbar}{i} \frac{d}{dx} \Psi \\ &= \int dx \frac{d\Psi^*}{dt} \frac{\hbar}{i} \frac{d}{dx} \Psi + \int dx \Psi^* \frac{\hbar}{i} \frac{d}{dx} \frac{d\Psi}{dt} \end{aligned}$$

Durch Umformen der Schrödingergleichung (5.22) erhält man die gesuchten Ableitungen der Wellenfunktion  $\Psi$  nach der Zeit

$$\begin{aligned} \frac{d\Psi}{dt} &= -\frac{i}{\hbar} \hat{H} \Psi \\ \frac{d\Psi^*}{dt} &= \frac{i}{\hbar} \hat{H} \Psi^* \end{aligned} \quad (5.27)$$

die Gleichung für  $\Psi^*$  ergibt sich durch komplex konjugieren der ersten Gleichung. Damit erhält man

$$\begin{aligned} \frac{d}{dt} \langle \Psi | \hat{p} | \Psi \rangle &= \int dx (\hat{H} \Psi^*) \frac{d\Psi}{dx} - \int dx \Psi^* \frac{d(\hat{H} \Psi)}{dx} \\ &= \underbrace{\int dx \left( -\frac{\hbar^2}{2m} \frac{d^2 \Psi^*}{dx^2} \right) \frac{d\Psi}{dx}}_{(1a)} + \underbrace{\int dx V \Psi^* \frac{d\Psi}{dx}}_{(1b)} \\ &\quad - \underbrace{\int dx \Psi^* \frac{d}{dx} \left( -\frac{\hbar^2}{2m} \frac{d^2 \Psi}{dx^2} \right)}_{(2a)} - \underbrace{\int dx \Psi^* \frac{dV \Psi}{dx}}_{(2b)} \end{aligned} \quad (5.28)$$

Betrachten wir zunächst einmal den Term in dieser Gleichung, der mit der Bezeichnung (1a) gekennzeichnet ist. Dieser lässt sich umformen mit Hilfe der partiellen Integration zu

$$\begin{aligned} (1a) &= -\frac{\hbar^2}{2m} \int_{-\infty}^{\infty} dx \frac{d^2 \Psi^*}{dx^2} \frac{d\Psi}{dx} \\ &= -\frac{\hbar^2}{2m} \left[ \frac{d\Psi^*}{dx} \frac{d\Psi}{dx} \right]_{-\infty}^{\infty} + \frac{\hbar^2}{2m} \int_{-\infty}^{\infty} dx \frac{d\Psi^*}{dx} \frac{d^2 \Psi}{dx^2}. \end{aligned}$$

Der Oberflächenterm dieser partiellen Integration, das ist der erste Term in der untersten Zeile, ist identisch Null, da man davon ausgehen kann, dass die Wellenfunktion wie auch deren Ableitung für  $x \rightarrow \pm\infty$  verschwinden. Das verbleibende Integral wird dann noch einmal partiell integriert, so dass

$$\begin{aligned} (1a) &= \frac{\hbar^2}{2m} \left[ \Psi^* \frac{d^2\Psi}{dx^2} \right]_{-\infty}^{\infty} - \frac{\hbar^2}{2m} \int_{-\infty}^{\infty} dx \Psi^* \frac{d^3\Psi}{dx^3} \\ &= -(1b). \end{aligned}$$

Man erkennt also, dass der Beitrag des mit (1a) bezeichneten Terms in Gleichung (5.28) kompensiert wird durch den Beitrag des Terms (1b). Es verbleiben also nur noch die Terme (2a) und (2b), die wir nach Anwendung der Produktregeln in (2b) zusammenfassen können zu

$$\begin{aligned} \frac{d}{dt} \langle \Psi | \hat{p} | \Psi \rangle &= \int dx \left\{ \Psi^* V \frac{d\Psi}{dx} - \Psi^* \frac{dV}{dx} \Psi - \Psi^* V \frac{d\Psi}{dx} \right\} \\ &= - \langle \Psi | \frac{dV}{dx} | \Psi \rangle \end{aligned}$$

Damit ist das Ehrenfest'sche Theorem (5.26) bewiesen.

Betrachtet man also nur in der Quantenmechanik nur die Mittelwerte der berechneten oder gemessenen Wahrscheinlichkeitsverteilung für die Ortskoordinate des betrachteten Teilchens, so stellt man fest, dass dieser Mittelwert sich mit der Zeit genauso verändert, wie dies von der Klassischen Mechanik vorhergesagt wird. Da aber für den Aufenthaltsort des Teilchen immer nur eine Wahrscheinlichkeitsverteilung angegeben werden kann, werden Messungen der Ortskoordinate zu verschiedenen Zeiten des Bewegungsablaufes auch immer wieder Ergebnisse liefern, bei denen die Ortskoordinate von dem Mittelwert abweicht. Eine Klassische Bewegung liegt genau dann vor, wenn die typischen Abweichungen der Einzelmessungen vom Mittelwert, also das mittlere Schwankungsquadrat für die Bestimmung der Ortskoordinate verschwindend klein im Vergleich zur Auflösung der Messapparatur ist. Dies ist der Klassische Grenzfall der Quantenmechanik, der uns bei den makroskopischen Prozessen in unserem Umfeld normalerweise begegnet.

## 5.3 Wahrscheinlichkeitsstrom

Es ist ein charakteristisches Merkmal der Quantenmechanik, dass die Position etwa eines Massenpunktes nur durch eine Wahrscheinlichkeitsdichte  $\rho(x, t)$  beschrieben werden kann. Diese Wahrscheinlichkeitsdichte erinnert an die Ladungsdichte der Elektrodynamik. Es ist ein wichtiges Postulat der Elektrodynamik, dass Ladungen nicht verschwinden können. Kommt es zu einer lokalen Änderung der Ladungsdichte, so muss diese Änderung mit einem elektrischen Strom verknüpft sein. Diese Verknüpfung zwischen einer zeitlichen Änderung der Ladungsdichte und der entsprechenden Stromdichte wird durch die Kontinuitätsgleichung der Elektrodynamik hergestellt.

Analog zur Erhaltung der Ladung in der Elektrodynamik muss man natürlich auch in der Quantenmechanik eine Erhaltung der Teilchen erwarten. Ändert sich die Wahrscheinlichkeitsdichte zur Beschreibung der Aufenthaltsorte mit der Zeit, so können die betrachteten Teilchen nicht im Nichts verschwinden. Auch hier muss jede lokale Änderung der Wahrscheinlichkeitsdichte mit einem Teilchenstrom verbunden sein. Dieser Teilchenstrom wird durch eine Wahrscheinlichkeitsstromdichte beschrieben und man erwartet eine Kontinuitätsgleichung, die ganz analog zur Elektrodynamik den Zusammenhang zwischen Dichte und Stromdichte herstellt und die Erhaltung der Teilchenzahl gewährleistet.

In einem ersten Schritt wollen wir zunächst wieder die Bewegung eines Teilchens der Masse  $m$  in einem Potential  $V$  in einer Raumrichtung  $x$  betrachten. Die Wahrscheinlichkeitsdichte  $\rho(x, t)$  ist definiert als das Betragsquadrat der Wellenfunktion  $\Psi(x, t)$  (5.2)

$$\rho(x, t) = \Psi^*(x, t)\Psi(x, t)$$

Die Ableitung dieser Wahrscheinlichkeitsdichte ergibt sich unter Verwendung der Produktregel aus den Ableitungen der Wellenfunktionen  $\Psi$  und  $\Psi^*$  nach der Zeit. Setzt man für diese Zeitableitungen die Ausdrücke (5.27), die sich ja direkt aus der Schrödingergleichung (5.22) ergaben, so erhält man

$$\begin{aligned} \frac{d\rho(x, t)}{dt} &= \frac{d\Psi^*(x, t)}{dt}\Psi(x, t) + \Psi^*(x, t)\frac{d\Psi(x, t)}{dt} \\ &= -\frac{1}{i\hbar}(\hat{H}\Psi^*)\Psi + \Psi^*\frac{1}{i\hbar}(\hat{H}\Psi) \\ &= -\frac{1}{i\hbar}\left(-\frac{\hbar^2}{2m}\frac{d^2\Psi^*}{dx^2} + V\Psi^*\right)\Psi + \Psi^*\frac{1}{i\hbar}\left(-\frac{\hbar^2}{2m}\frac{d^2\Psi}{dx^2} + V\Psi\right) \\ &= \frac{\hbar}{2mi}\frac{d}{dx}\left[\Psi\frac{d\Psi^*}{dx} - \Psi^*\frac{d\Psi}{dx}\right] \end{aligned} \quad (5.29)$$

Diesen Fall der Bewegung in einer Raumrichtung  $x$  können wir jetzt leicht verallgemeinern auf die Bewegung im dreidimensionalen Raum, beschrieben durch die drei kartesischen Koordinaten  $x, y, z$  beziehungsweise den entsprechenden Ortsvektor  $\vec{r}$ . In diesem Fall hängen also die Wahrscheinlichkeitsdichte  $\rho$ , die Wellenfunktionen  $\Psi$  und das Potential  $V$  von den drei Koordinaten ab und der Hamiltonoperator enthält neben dem Term für die kinetische Energie in  $x$ -Richtung auch entsprechende Terme für die  $y$ - und  $z$ -Richtung

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} - \frac{\hbar^2}{2m}\frac{d^2}{dy^2} - \frac{\hbar^2}{2m}\frac{d^2}{dz^2} + V(\vec{r})$$

$$= -\frac{\hbar^2}{2m}\Delta + V(\vec{r}),$$

wobei in der 2. Zeile die Ableitungen nach  $x$ ,  $y$  und  $z$  ersetzt wurden durch den Laplace Operator

$$\begin{aligned}\Delta &= \text{divGrad} = \vec{\nabla} \cdot \vec{\nabla} \\ &= \frac{d^2}{dx^2} + \frac{d^2}{dy^2} + \frac{d^2}{dz^2} \quad \text{in karthes. Koordinaten}\end{aligned}\quad (5.30)$$

Mit diesen Verallgemeinerungen auf drei Dimensionen ergibt sich entsprechend (5.29) für

$$\begin{aligned}\frac{d\rho(\vec{r}, t)}{dt} &= \frac{\hbar}{2mi} \left\{ \frac{d}{dx} \left[ \Psi \frac{d\Psi^*}{dx} - \Psi^* \frac{d\Psi}{dx} \right] + \frac{d}{dy} \left[ \Psi \frac{d\Psi^*}{dy} - \Psi^* \frac{d\Psi}{dy} \right] + \frac{d}{dz} \left[ \Psi \frac{d\Psi^*}{dz} - \Psi^* \frac{d\Psi}{dz} \right] \right\} \\ &= -\frac{\hbar}{2mi} \begin{pmatrix} \frac{d}{dx} \\ \frac{d}{dy} \\ \frac{d}{dz} \end{pmatrix} \cdot \begin{pmatrix} \Psi^* \frac{d\Psi}{dx} - \Psi \frac{d\Psi^*}{dx} \\ \Psi^* \frac{d\Psi}{dy} - \Psi \frac{d\Psi^*}{dy} \\ \Psi^* \frac{d\Psi}{dz} - \Psi \frac{d\Psi^*}{dz} \end{pmatrix} \\ &= -\vec{\nabla} \cdot \vec{j} = -\text{div} \vec{j}.\end{aligned}\quad (5.31)$$

Wir haben also hier eine Kontinuitätsgleichung der Form

$$\frac{d\rho(\vec{r}, t)}{dt} + \text{div} \vec{j} = 0 \quad (5.32)$$

für die Wahrscheinlichkeitsdichte  $\rho = \Psi^* \Psi$  hergeleitet mit der Wahrscheinlichkeitsstromdichte

$$\vec{j}(\vec{r}, t) = \frac{\hbar}{2mi} \left[ \Psi^* \vec{\nabla} \Psi - \Psi \vec{\nabla} \Psi^* \right]. \quad (5.33)$$

In der Diskussion dieser Kontinuitätsgleichung soll in einem ersten Schritt noch einmal (vergleiche entsprechende Diskussion in der Elektrodynamik) die Bedeutung dieser Gleichung erörtert werden. Dazu betrachtet man ein endliches Volumen  $\mathcal{V}$  und integriert die Kontinuitätsgleichung über dieses Volumen

$$\begin{aligned}0 &= \int_{\mathcal{V}} d^3r \frac{d\rho(\vec{r}, t)}{dt} + \int_{\mathcal{V}} d^3r \text{div} \vec{j}(\vec{r}, t) \\ &= \frac{dN(\mathcal{V})}{dt} + \oint_{S(\mathcal{V})} \vec{j} \cdot \vec{df}\end{aligned}$$

Beim Übergang zur zweiten Zeile wurde das Integral der Dichte  $\rho$  über das Volumen  $\mathcal{V}$  durch die entsprechende Teilchenzahl  $N(\mathcal{V})$  bezeichnet. Bei dem zweiten Term wurde der Gauss'sche Integralsatz benutzt, nach dem das Volumenintegral über  $\text{div} \vec{j}$  umgeschrieben wird in ein Integral über die Oberfläche des Volumens  $S(\mathcal{V})$  mit dem Integranden  $\vec{j}$ . Die Differentialform für dieses Oberflächenintegral  $\vec{df}$  bezeichnet dabei einen Vektor, der senkrecht zur Oberfläche steht und nach aussen weist. Damit wird also z.B. jede Verringerung der Teilchenzahl  $N$  in der integrierten Kontinuitätsgleichung kompensiert durch einen positiven Beitrag durch das Oberflächenintegral: der gesamte Abfluss der Stromdichte  $\vec{j}$  durch die Oberfläche ist soviel grösser als der entsprechende Zufluss, dass die Summe der beiden Terme insgesamt Null ergibt.



Zur Interpretation der physikalischen Bedeutung der Wahrscheinlichkeitsstromdichte schreiben wir die Definition der Stromdichte (5.33) mit Hilfe des Impulsoperators, das ist für drei Raumdimensionen der Vektoroperator

$$\hat{\vec{p}} = \frac{\hbar}{i} \vec{\nabla}, \quad (5.34)$$

um zu

$$\vec{j} = \frac{1}{2m} \left[ \Psi^* (\hat{\vec{p}} \Psi) + (\hat{\vec{p}} \Psi^*) \Psi \right] \quad (5.35)$$

Daraus erkennt man, dass die Wahrscheinlichkeitsstromdichte gleich dem Integranden zur Berechnung des Erwartungswertes für die Geschwindigkeit des Massenpunktes ( $\vec{v} = \vec{p}/m$ ) ist. Die zwei Beiträge sind dabei erforderlich, damit dieser Ausdruck reelle Werte ergibt. Die Wahrscheinlichkeitsstromdichte kann also als Geschwindigkeitsdichte interpretiert werden.

Als Beispiel für die Berechnung der Stromdichte wollen wir die Wellenfunktion einer ebenen Welle

$$\Psi(\vec{r}, t) = A e^{i(\vec{k} \cdot \vec{r} - \omega t)}, \quad (5.36)$$

betrachten. Zu einer fest vorgegebenen Zeit  $t = t_0 = 0$  ist der Wert dieser Funktion  $\Psi(\vec{r}, t_0)$  für alle Ortsvektoren  $\vec{r}$ , für die das Skalarprodukt mit dem Wellenzahlvektor  $\vec{k} \cdot \vec{r}$  den gleichen Wert ergibt, gleich. Die Punkte mit diesen Ortsvektoren liegen auf einer Ebene senkrecht zum Wellenvektor  $\vec{k}$  (siehe Abbildung 5.2). Diese Ebene bildet eine Wellenfront der ebenen Welle. Zu einem späteren Zeitpunkt  $t_1 > t_0$  ist diese Wellenfront in Richtung des Wellenzahlvektors,  $\vec{k}$ , gewandert mit einer Geschwindigkeit  $v$ , die sich aus der Forderung nach der Konstanz der Phase in (5.36) ergibt (siehe auch Abbildung 5.2):

$$\begin{aligned} |\vec{k}| r_0 - \omega t_0 &= |\vec{k}| r_1 - \omega t_1 \\ v &= \frac{r_1 - r_0}{t_1 - t_0} = \frac{\omega}{|\vec{k}|}. \end{aligned}$$

Zur Berechnung der Stromdichte sei nun angenommen, dass der Wellenzahlvektor  $\vec{k} = (k, 0, 0)$  also parallel zur  $x$ -Achse sei. In diesem Fall berechnet sich die Wahrscheinlichkeitsstromdichte  $\vec{j}$  für die Wellenfunktion  $\Psi$  aus (5.36) mit (5.33)

$$\begin{aligned} \vec{j}(\vec{r}, t) &= \frac{\hbar}{2mi} |A|^2 \left[ e^{-i(kx - \omega t)} \begin{pmatrix} ik \\ 0 \\ 0 \end{pmatrix} e^{i(kx - \omega t)} - e^{i(kx - \omega t)} \begin{pmatrix} -ik \\ 0 \\ 0 \end{pmatrix} e^{-i(kx - \omega t)} \right] \\ &= |A|^2 \frac{\hbar k}{m} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \end{aligned} \quad (5.37)$$

Die Stromdichte ist also in diesem Fall ein konstanter Vektor, unabhängig von Ort und Zeit, in Richtung der Wellenzahl, und damit auch in Richtung des Impulsvektors, mit einem Betrag, der bis auf die Normierungskonstante  $|A|^2$  gerade gleich dem Betrag des Impulses ( $\hbar k$ ) dividiert durch die Masse  $m$  ist.

Erwähnenswert ist ausserdem, dass die Wellenfunktion in diesem Beispiel eine Wahrscheinlichkeitsdichte  $\rho$  definiert, die unabhängig von der Zeit ist. Die Ableitung von  $\rho$

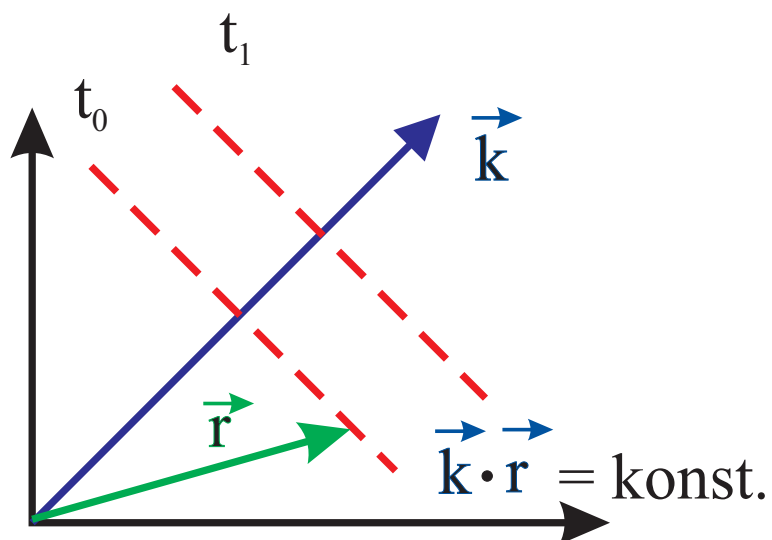


Abbildung 5.2: Ausbreitung einer ebenen Welle (5.36)

nach der Zeit ist also gleich Null, was wegen der Kontinuitätsgleichung (5.32) dann aber auch bedeutet, dass

$$\operatorname{div} \vec{j} = 0. \quad (5.38)$$

Analog zu den elektrischen Strömen in der Magnetostatik, spricht man auch in diesem Fall von stationären Strömen, Ströme die ungleich null sein können aber zu keiner Veränderung der Ladungs- beziehungsweise Wahrscheinlichkeitsverteilung führen.

## 5.4 Stationäre Lösungen der Schrödingergleichung

In diesem Abschnitt und ebenso in den meisten der folgenden Abschnitte wollen wir uns mit Systemen beschäftigen, die durch eine Hamiltonfunktion beschrieben werden, die nicht explizit von der Zeit  $t$  abhängt. Für die Bewegung von Massenpunkten in einem Potenzial bedeutet dies, dass das Potenzial  $V$  sich nicht mit der Zeit ändert. In diesem Fall können wir spezielle Lösungen, die sogenannten stationären Lösungen der Schrödingergleichung (5.22) finden. Es gilt nämlich

**Satz:** *Ist die Hamiltonfunktion, beziehungsweise der entsprechende Hamiltonoperator  $\hat{H}$  nicht explizit zeitabhängig, so besitzt die Schrödingergleichung **stationäre Lösungen** der Form:*

$$\Psi(x, t) = \Phi_E(x) f_E(t) = \Phi_E(x) e^{-iEt/\hbar}. \quad (5.39)$$

Dabei ist der ortsabhängige Faktor eine Lösung der **stationären Schrödingergleichung**

$$\hat{H}\Phi_E(x) = E\Phi_E(x) \quad (5.40)$$

und die Konstante  $E$  bezeichnet den Erwartungswert des Hamiltonoperators  $\hat{H}$  berechnet mit den normierten Funktionen  $\Phi_E(x)$ .  $E$  ist also im Fall von konservativen Kraftfeldern der Wert für die Energie.

Zum Beweis dieses Satzes verifizieren wir, dass die stationäre Lösung (5.39) eine Lösung der Schrödingergleichung ist, und setzen diese dazu ein:

$$\begin{aligned} \hat{H}\Phi_E(x)f_E(t) &= i\hbar \frac{\Phi_E(x)f_E(t)}{dt} \\ \frac{1}{\Phi_E(x)f_E(t)} \Big| & f_E(t) \hat{H}\Phi_E(x) = \Phi_E(x) i\hbar \frac{df_E(t)}{dt}. \end{aligned}$$

Beim Übergang zur zweiten Zeile wurde ausgenutzt, dass der Hamiltonoperator  $\hat{H}$  nur auf den Ortsanteil der Wellenfunktion wirkt. Ausserdem wird in dieser Zeile angedeutet, dass wir im nächsten Schritt die beiden Seiten der Gleichung jeweils von links mit der Inversen der Wellenfunktion multiplizieren wollen. Dies führt zu

$$\frac{1}{\Phi_E(x)} \hat{H}\Phi_E(x) = \frac{1}{f_E(t)} i\hbar \frac{df_E(t)}{dt}.$$

Ist der Hamiltonoperator zeitunabhängig, so hängt die linke Seite dieser Gleichung allenfalls von der Ortskoordinate  $x$  ab, ist aber unabhängig von der Zeit  $t$ . Die rechte Seite ist aber unabhängig von  $x$ . Die Gleichung kann also nur dann erfüllt sein, wenn beide Seiten eine Konstante  $E$  ergeben, die weder von  $x$  noch von  $t$  abhängt. Es gilt also

$$\frac{1}{\Phi_E(x)} \hat{H}\Phi_E(x) = E \Rightarrow \hat{H}\Phi_E(x) = E\Phi_E(x) \quad (5.41)$$

$$\frac{1}{f_E(t)} i\hbar \frac{df_E(t)}{dt} = E \Rightarrow i\hbar \frac{df_E(t)}{dt} = E f_E(t) \quad (5.42)$$

Die erste dieser beiden Gleichungen, die erfüllt sein müssen, damit der Produktansatz (5.39) eine Lösung der Schrödingergleichung ist, ist gerade die stationäre Schrödingergleichung (5.40). Die zweite Gleichung (5.42) ist eine Differentialgleichung zur Bestimmung von  $f_E(t)$ . Durch Einsetzen verifiziert man leicht, dass

$$f_E(t) = e^{-iEt/\hbar}$$

die (bis auf einen konstanten Faktor) eindeutige Lösung dieser Gleichung ist. Um die Bedeutung der Zahl  $E$  zu bestätigen, berechnen wir nun noch den Erwartungswert des Hamiltonoperators

$$\begin{aligned} \langle \Psi(x, t) | \hat{H} | \Psi(x, t) \rangle &= \int dx \Phi_E^*(x) e^{iEt/\hbar} \underbrace{\hat{H} \Phi_E(x)}_{=E\Phi_E(x)} e^{-iEt/\hbar} \\ &= E \int dx \Phi_E^*(x) \Phi_E(x) \\ &= E \end{aligned}$$

Die letzte Gleichung gilt genau dann, wenn die ortsabhängige Wellenfunktion auf 1 normiert ist. Damit ist also der Beweis des Satzes über die stationäre Lösung der Schrödingergleichung erbracht.

Für diese stationäre Lösung gilt, dass der Erwartungswert des Hamiltonoperators den Wert  $E$  für alle Zeiten  $t$  liefert. Es gilt aber auch

$$\begin{aligned} \hat{H}^2 \Phi_E(x) &= \hat{H} [E \Phi_E(x)] \\ &= E [\hat{H} \Phi_E(x)] \\ &= E^2 \Phi_E(x). \end{aligned} \tag{5.43}$$

Dies bedeutet, dass der Erwartungswert des Operators  $\hat{H}^2$  in diesem Fall den Wert  $E^2$  ergibt. Damit ist das mittlere Schwankungsquadrat (5.8) für die Messung oder Berechnung der Energie

$$(\Delta E)^2 = \langle \Psi | \hat{H}^2 | \Psi \rangle - \langle \Psi | \hat{H} | \Psi \rangle^2 = 0. \tag{5.44}$$

Die Energie eines System, das durch eine stationäre Lösung beschrieben wird, ist also nicht nur konstant als Funktion der Zeit sondern auch scharf bestimmt.

Betrachten wir nun als ein erstes Beispiel für die stationäre Lösung der Schrödingergleichung den Fall, dass sich ein Teilchen der Mass  $m$  in einem konstanten Potenzial  $V$  bewegt. Die zeitunabhängige Schrödingergleichung hat dann die Form

$$\begin{aligned} \left\{ -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V \right\} \Phi(x) &= E \Phi(x) \\ \text{beziehungsweise} \quad \frac{d^2 \Phi(x)}{dx^2} &= \underbrace{\frac{\hbar^2}{2m} (V - E)}_{=\alpha^2} \Phi(x) \end{aligned} \tag{5.45}$$

Für dies Differentialgleichung zweiter Ordnung gibt es zwei linear unabhängige Lösungen

$$\Phi_1(x) = e^{\alpha x} \quad \text{und} \quad \Phi_2(x) = e^{-\alpha x} \tag{5.46}$$

Je nachdem welches Vorzeichen die Konstante  $\alpha^2$ , die wir in (5.45) definiert haben, besitzt, sind diese Lösungen ganz unterschiedlich. Wir unterscheiden:

**V > E:** Dieser Fall ist in der klassischen Mechanik ausgeschlossen. Da sich die Energie  $E$  als Summe aus potenzieller Energie  $V$  und der positiv definiten kinetische Energie ergibt, muss  $E$  grösser als  $V$  sein. Auch in der Quantenmechanik gibt es in diesem Fall keine akzeptable Lösung, wenigstens nicht wenn  $V$  im gesamten Intervall  $-\infty \leq x \leq \infty$  definiert ist. In diesem Fall ist  $\alpha^2$  positiv,  $\alpha$  also reell. Je nach Vorzeichen von  $\alpha$  divergiert aber dann die Lösung (5.46) entweder für  $x \rightarrow \infty$  oder für  $x \rightarrow -\infty$ . Das bedeutet, dass die Aufenthaltswahrscheinlichkeit für diesen Wert von  $x$  divergiert, beziehungsweise, dass die Wellenfunktionen nicht normierbar sind. Deshalb können wir diese Lösungen als unphysikalisch verwerfen.

**V < E** In diesem Fall ist  $\alpha^2$  negativ,  $\alpha$  also rein imaginär. Dies bedeutet, dass die Lösungen (5.46) die Form annehmen

$$\Phi_{1/2}(x) = \frac{1}{\sqrt{2\pi}} e^{\pm ikx}$$

mit  $k = \sqrt{\frac{2m}{\hbar^2}(E - V)}$  (5.47)

Diese Lösungen entsprechen gerade den ebenen Wellen, sind also identisch mit den Lösungen des freien Teilchens ( $V = 0$ ). Allerdings wird die Wellenzahl  $k$  durch ein attraktives Potenzial ( $V < 0$ ) vergrößert beziehungsweise verkleinert durch ein repulsives Potenzial. Anstelle dieser komplexwertigen Wellenfunktion aus (5.47) können wir natürlich auch die beiden Linearkombinationen

$$\begin{aligned} \tilde{\Phi}_1(x) &= \frac{1}{\sqrt{2}} [\Phi_1(x) + \Phi_2(x)] = \frac{1}{\sqrt{4\pi}} [e^{ikx} + e^{-ikx}] = \frac{1}{\sqrt{\pi}} \cos kx \\ \tilde{\Phi}_2(x) &= \frac{1}{\sqrt{2i}} [\Phi_1(x) - \Phi_2(x)] = \frac{1}{i\sqrt{4\pi}} [e^{ikx} - e^{-ikx}] = \frac{1}{\sqrt{\pi}} \sin kx \end{aligned} \quad (5.48)$$

als Satz von linear unabhängigen Lösungen betrachten.

Als weiteres Beispiel wollen wir nun für das Potenzial annehmen

$$V(x) = \begin{cases} 0 & \text{für } 0 \leq x \leq d \\ V & \text{sonst} \end{cases} \quad (5.49)$$

das auch in Figur 5.3 dargestellt ist. Speziell wollen wir eine Lösung für ein Teilchen mit einer Energie  $E < V$  suchen, das also im Potenzialtopf zwischen  $x = 0$  und  $x = d$  gebunden ist. Zur Lösung dieses Problem betrachten wir zunächst isoliert voneinander die 3 Bereiche, die in der Abbildung 5.3 skizziert sind:

$$\begin{aligned} \text{Bereich I:} & \quad -\infty \leq x < 0 \\ \text{Bereich II:} & \quad 0 \leq x \leq d \\ \text{Bereich III:} & \quad d < x \leq \infty \end{aligned}$$

In diesen drei Teilbereichen ist das Potenzial jeweils konstant, wir können also die Ergebnisse des gerade diskutierten Falles  $V = \text{konst}$  übernehmen. Für den Bereich I ( $E < V$ ) erhalten wir also als allgemeine Lösung eine Linearkombination der zwei Lösungen aus (5.46)

$$\Phi_{\text{I}}(x) = a_{\text{I}} e^{\alpha x} + b_{\text{I}} e^{-\alpha x} \quad (5.50)$$

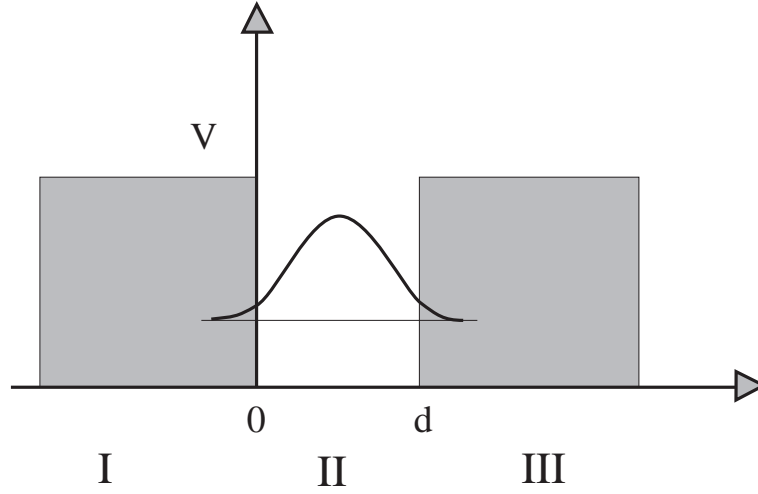


Abbildung 5.3: Darstellung des Kastenpotenzials aus (5.49). Die Abbildung enthält auch die schematische Darstellung einer Wellenfunktion für  $E < V$

mit

$$\alpha = \sqrt{\frac{2m}{\hbar^2} |V - E|} \quad (5.51)$$

Dieser Ansatz gilt im Bereich I also für  $-\infty \leq x < 0$ . Folglich muss  $b_{\text{I}} = 0$  sein, da sonst die Funktion für  $x \rightarrow -\infty$  divergieren würde. Analog gilt im Bereich III

$$\Phi_{\text{III}}(x) = a_{\text{III}} e^{\alpha x} + b_{\text{III}} e^{-\alpha x} \quad (5.52)$$

mit dem gleichen Wert für  $\alpha$ . Diese Wellenfunktion soll aber im Bereich  $d < x \leq \infty$  gelten; also muss  $a_{\text{III}} = 0$  sein. Im Bereich II, hier ist  $E > V(x) = 0$ , setzen wir als allgemeine Lösung der stationären Schrödingergleichung eine Linearkombination der Lösungen (5.48) an:

$$\Phi_{\text{II}}(x) = a_{\text{II}} \cos kx + b_{\text{II}} \sin kx \quad \text{mit} \quad k = \sqrt{\frac{2m}{\hbar^2} E} \quad (5.53)$$

Zur Lösung des Problems müssen die 4 Konstanten  $a_{\text{I}}$ ,  $a_{\text{II}}$ ,  $b_{\text{II}}$  und  $b_{\text{III}}$  bestimmt werden. Dazu vergegenwärtigen wir uns, dass es zur Lösung der Schrödingergleichung im gesamten Raum nicht ausreicht, diese isoliert in den einzelnen Teilbereichen zu lösen. Die Schrödingergleichung ist eine Differentialgleichung zweiter Ordnung im ganzen Raum, die Wellenfunktion und deren erste Ableitung sollten also an Übergängen zwischen 2 Bereichen stetig sein. Wir erhalten also 4 Bestimmungsgleichungen:

$$\Phi_{\text{I}}(x=0) = \Phi_{\text{II}}(0) \quad \text{sowie} \quad \Phi_{\text{II}}(x=d) = \Phi_{\text{III}}(d) \quad (5.54)$$

$$\frac{d\Phi_{\text{I}}}{dx}(x=0) = \frac{d\Phi_{\text{II}}}{dx}(0) \quad \text{sowie} \quad \frac{d\Phi_{\text{II}}}{dx}(x=d) = \frac{d\Phi_{\text{III}}}{dx}(d) \quad (5.55)$$

Ausserdem soll die Wellenfunktion auch noch normiert sein:

$$1 = \int_{-\infty}^0 dx \Phi_{\text{I}}^* \Phi_{\text{I}}(x) + \int_0^d dx \Phi_{\text{II}}^* \Phi_{\text{II}}(x) + \int_d^{\infty} dx \Phi_{\text{III}}^* \Phi_{\text{III}}(x) \quad (5.56)$$

Wir haben also 5 Gleichungen (5.54 - 5.56) zur Bestimmung von den 4 Unbekannten. Es ist klar, dass dieses Problem im allgemeinen Fall keine Lösung besitzt. Wir können nur bei einigen wohldefinierten Werten für die Energie  $E$  und damit für  $\alpha$  in (5.51) bzw.  $k$  in (5.53) eine Lösung erwarten.

Ein Beispiel für eine solche Wellenfunktion in den Bereichen I - III mit stetigen Übergängen an den Schnittstellen ist in der Abbildung 5.3 skizziert. Man sieht an dieser Skizze, dass die Wellenfunktion auch in den Bereichen I und III nicht vollständig verschwindet, sie fällt exponentiell ab. Das Teilchen hat also auch in den Bereichen, in denen es sich nach den Gesetzen der Klassischen Mechanik nicht aufhalten kann, eine gewisse Aufenthaltswahrscheinlichkeit. Die Eindringtiefe ist proportional  $1/\alpha$  ist also grösser, je kleiner die Energiedifferenz  $V - E$ .

Es gibt also für Teilchen, die in einem Potenzialtopf gebunden sind, nur Lösungen bei einzelnen diskreten Energien: Zwar gibt es bei jeder Energie 2 linear unabhängige Lösungen der Schrödingergleichung, da diese ja eine Differentialgleichung zweiter Ordnung ist. Für ein lokalisiertes Teilchen muss die Wellenfunktion aber Randbedingungen für  $x \rightarrow \infty$  und  $x \rightarrow -\infty$  erfüllen. Hinzu kommt die Forderung nach der Normierung der Wellenfunktion. Eine Lösung dieses Problems ist nur bei einzelnen Energien möglich.

Dies kann man sich am oben diskutierten Beispiel (5.49) im Grenzfall  $V \rightarrow \infty$  einfach veranschaulichen. In diesem Fall gilt  $\alpha \Rightarrow \infty$ , die Eindringtiefe der gebundenen Teilchen in die klassisch verbotenen Bereiche I und III wird 0 und die Wellenfunktionen in diesen Bereichen verschwinden identisch:  $\Phi_{\text{I}} = \Phi_{\text{III}} = 0$ . Damit ergeben sich auch für die Wellenfunktion im Bereich II die Randbedingungen

$$\Phi_{\text{II}}(0) = \Phi_{\text{II}}(d) = 0.$$

Damit ergibt sich im Ansatz (5.53)

$$a_{\text{II}} = 0 \quad \text{und} \quad kd = n\pi \iff k = \frac{n\pi}{d}$$

Durch diese Bedingung an  $k$  bekommen wir in diesem Fall über (5.53) auch direkt die Werte für Energien, für die eine solche Lösung gefunden wird. Es sind dies die diskreten Energiewerte

$$E_n = \frac{\hbar^2}{2m} k^2 = \frac{\hbar^2 \pi^2}{2m d^2} n^2 \quad (5.57)$$

### 5.4.1 Tunneleffekt

Als ein weiteres Beispiel für die Behandlung von quantenmechanischen Effekten mit stückweisen konstanten Potenzialen, wollen wir betrachten, wie sich ein Teilchen an einer Potenzialschwelle verhält. In diesem Fall betrachten wir ein Potenzial der Form

$$V(x) = \begin{cases} V & \text{für } 0 \leq x \leq d \\ 0 & \text{sonst} \end{cases} \quad (5.58)$$

und ein Teilchen der Masse  $m$  mit einer Energie  $E$ , die kleiner sein soll als der Wert der Potenzialschwelle  $V$  (siehe Abb. 5.4). Im Rahmen der klassischen Mechanik kann sich ein

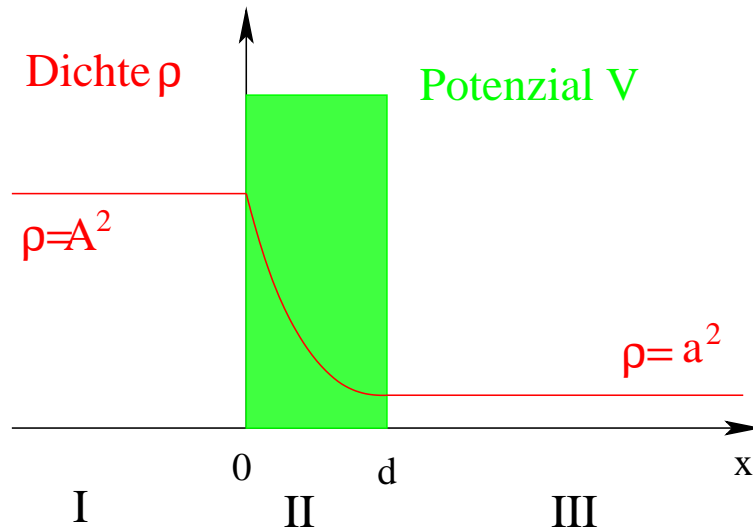


Abbildung 5.4: Darstellung der Potenzialschwelle (5.58). Die Abbildung enthält auch die schematische Darstellung der Wahrscheinlichkeitsdichten in den 3 Bereichen I - III

solches Teilchen deshalb entweder nur auf der linken Seite der Schwelle, also im Bereich I oder auf der rechten Seite der Schwelle also im Bereich III aufhalten. Nehmen wir also an, dass das Teilchen sich im linken Bereich I aufhält. In der Quantenmechanik würde ein solches Teilchen mit der Energie

$$E = \frac{p^2}{2m}$$

und einem zugehörigen Impuls  $p = \hbar k$  beschrieben durch eine ebene Welle

$$\Phi_{\text{I}}(x) = A e^{ikx},$$

und hätte also überall in diesem Bereich die Wahrscheinlichkeitsdichte

$$\rho_{\text{I}} = A^2$$

Im Bereich II, also im Bereich der Schwelle hat die Wellenfunktion die Form

$$\Phi_{\text{I}}(x) = A e^{-\alpha x} \quad \text{mit} \quad \alpha = \sqrt{\frac{\hbar^2}{2m}(V - E)}.$$

Damit können wir im Bereich III eine Wellenfunktion der Form

$$\Phi_{\text{III}}(x) = a e^{ik(x-d)} \quad \text{mit} \quad a = A e^{-\alpha d},$$

stetig an die Funktion im Bereich II anpassen. Damit ergibt sich also auch für den Bereich III, in dem sich ja das Teilchen klassisch nicht aufhalten darf eine von Null verschiedene Wahrscheinlichkeitsdichte. Man sagt, das Teilchen ist durch die Potenzialbarriere **getunnelt**. Die **Tunnelwahrscheinlichkeit** oder der entsprechende **Transmissionskoeffizient** ergibt sich zu

$$T = \frac{\rho_{\text{III}}}{\rho_{\text{I}}} = \frac{a^2}{A^2} = e^{-2\alpha d}. \quad (5.59)$$

Dieser Transmissionskoeffizient fällt also exponentiell mit der Breite der Schwelle  $d$  ab.



### 5.4.2 Beschreibung von bewegten Teilchen

Die Wahrscheinlichkeitsdichte, berechnet für eine stationäre Lösung der Schrödingergleichung, ist unabhängig von der Zeit

$$\rho(x, t) = \Psi_E^*(x, t)\Psi_E(x, t) = \Phi_E^*(x)\Phi_E(x)$$

daher rührt auch der Name: **stationäre** Lösung. Da sich die Dichte mit der Zeit nicht ändert, bleibt der Erwartungswert für den Aufenthaltsort des Teilchens, wie auch alle anderen Erwartungswerte zeitlich konstant. Dies bedeutet aber nicht, dass die Teilchen einen Impuls oder Geschwindigkeit 0 besitzen. Der Erwartungswert auch für den Impuls ist eine Konstante und berechnet sich z.B. für eine ebene Welle

$$\begin{aligned} \langle \Psi_E | \hat{p} | \Psi_E \rangle &= \langle \Phi_E | \hat{p} | \Phi_E \rangle \\ &= \hbar k_0 \quad \text{für } \Phi(x) = \frac{1}{\sqrt{2\pi}} e^{ik_0 x} \end{aligned}$$

wie wir ihn im Rahmen der Elektrodynamik aus der Magnetostatik kennen.

Die stationären Lösungen der Schrödingergleichung beschreiben also Systeme, bei denen die Erwartungswerte sich mit der Zeit nicht ändern. Nach dem Ehrenfestschen Theorem entspricht aber die Entwicklung des Erwartungswertes mit der Zeit der Änderung dieser Observablen wie sie in der Klassischen Mechanik beschrieben wird. Solche stationären Lösungen entsprechen also Klassischen Systemen, bei denen sich nichts bewegt. Wie können wir nun solche stationären Lösungen benutzen, um bewegte Systeme zu beschreiben? Dazu betrachten wir eine Linearkombination zweier stationärer Lösungen  $\Psi_i$  zu verschiedenen Energien  $E_i$

$$\Psi(x, t) = \alpha\Psi_1(x, t) + \beta\Psi_2(x, t) \quad (5.60)$$

mit komplexwertigen Koeffizienten  $\alpha$  und  $\beta$ .

**Es gilt:** Die Linearkombination (5.60) ist ebenfalls eine Lösung der Schrödingergleichung. Diese Lösung ist nicht stationär und führt zu einer Wahrscheinlichkeitsdichte  $\rho(x, t)$  und zu Erwartungswerten, die sich mit der Zeit ändern.

Zum Beweis wenden wir den Hamiltonoperator auf die in (5.60) definierte Wellenfunktion an

$$\begin{aligned} \hat{H}\Psi(x, t) &= \alpha\hat{H}\Psi_1(x, t) + \beta\hat{H}\Psi_2(x, t) \\ &= \alpha i\hbar \frac{d}{dt}\Psi_1(x, t) + \beta i\hbar \frac{d}{dt}\Psi_2(x, t) \\ &= i\hbar \frac{d}{dt}(\alpha\Psi_1(x, t) + \beta\Psi_2(x, t)) \\ &= i\hbar \frac{d}{dt}\Psi(x, t) \end{aligned}$$

In der ersten Zeile dieser Gleichungen wurde ausgenutzt, dass die Konstanten  $\alpha$  und  $\beta$  unabhängig von der Koordinate  $x$  sind, die Ableitungsoperatoren, die in  $\hat{H}$  stecken, also

nicht auf diese Konstanten angewandt werden müssen. Bei dem Übergang zur zweiten Zeile wird ausgenutzt, dass  $\Psi_1$  und  $\Psi_2$  Lösungen der Schrödingergleichung sind. Bei dem Übergang zur dritten Zeile geht ein, dass  $\alpha$  und  $\beta$  unabhängig von der Zeit  $t$  sind. Insgesamt zeigt also diese Rechnung, dass  $\Psi$  eine Lösung der Schrödingergleichung ist. Der zweite Teil der obigen Behauptung wird durch direktes Ausrechnen verifiziert:

$$\begin{aligned} \rho &= \left( \alpha \Phi_1(x) e^{-E_1/\hbar t} + \beta \Phi_2(x) e^{-E_2/\hbar t} \right)^* \left( \alpha \Phi_1(x) e^{-E_1/\hbar t} + \beta \Phi_2(x) e^{-E_2/\hbar t} \right) \\ &= \alpha^* \alpha \Phi_1^*(x) \Phi_1(x) + \beta^* \beta \Phi_2^*(x) \Phi_2(x) + \alpha \beta^* \Phi_2^*(x) \Phi_1(x) e^{i \frac{E_2 - E_1}{\hbar} t} \\ &\quad + \alpha^* \beta \Phi_1^*(x) \Phi_2(x) e^{i \frac{E_1 - E_2}{\hbar} t} \end{aligned}$$

Zwar sind die ersten beiden Terme unabhängig von der Zeit, aber in Interferenztermen zwischen den beiden Wellenfunktionen  $\Phi_1$  und  $\Phi_2$  ergibt sich eine Zeitabhängigkeit. Dadurch ist also die Wahrscheinlichkeitsdichte, berechnet für die Überlagerung von zwei stationären Lösungen (5.60), abhängig von der Zeit.

Natürlich gelten die gerade bewiesenen Aussagen auch für Linearkombinationen von drei und mehr stationären Lösungen. Insbesondere ist das Wellenpaket

$$\Psi(x, t) = \int dk C(k) e^{ikx} e^{-iE_k/\hbar t} \quad \text{mit} \quad E_k = \frac{\hbar^2 k^2}{2m} \quad (5.61)$$

eine Überlagerung von stationären Lösungen der Schrödingergleichung eines freien Teilchens. Dieses Wellenpaket ist also selbst eine Lösung der Schrödingergleichung für ein freies Teilchen. Wie wir in der Diskussion des Abschnittes 8.1 gesehen haben, kann  $C(k)$  so gewählt werden, dass die Wahrscheinlichkeitsverteilung zu dieser Wellenfunktion lokalisiert ist. Da diese Wahrscheinlichkeitsverteilung sich mit der Zeit ändert, "bewegt" sich diese lokalisierte Wahrscheinlichkeitsverteilung. Wir können also mit (5.61) die Bewegung eines freien Teilchens beschreiben. Allerdings können wir dem Wellenpaket (5.61) keine feste Energie zuordnen. Es handelt sich vielmehr um eine Überlagerung von Zuständen mit verschiedenen Energien  $E_k$ . Wie wir sehen werden, ist diese Energieverteilung, beziehungsweise die Unschärfe des Energiewertes, der Preis, den man in der Quantenmechanik "bezahlen" muss, wenn man eine Bewegung beschreiben will.

## 5.5 Der Harmonische Oszillator I

Als erstes nichttriviales Beispiel für die Lösung der stationären Schrödingergleichung wollen wir die Bewegung eines Teilchens im Potenzial des *Harmonischen Oszillators*

$$V(x) = \frac{1}{2}\kappa x^2 = \frac{1}{2}m\omega^2 x^2 \quad (5.62)$$

behandeln. Dieses Beispiel des Harmonischen Oszillators ist in der Mechanik wie in der Quantenmechanik sicher auch deshalb sehr beliebt, da man die Bewegungsgleichungen für ein Teilchen der Masse  $m$  in einem Oszillatorpotenzial recht einfach lösen kann. In der Klassischen Mechanik ergibt sich für die Position des Teilchens  $x$  als Funktion der Zeit

$$x(t) = A \cos \omega t + B \sin \omega t, \quad (5.63)$$

also eine Überlagerung von einer Cosinus und Sinusschwingung mit der Winkelgeschwindigkeit

$$\omega = \sqrt{\frac{\kappa}{m}}$$

(siehe auch (5.62)). Die Konstanten  $A$  und  $B$  in (5.63) lassen sich dann durch die Anfangsbedingungen, also den Ort und die Geschwindigkeit des Teilchens zur Zeit  $t = 0$ , eindeutig festlegen. Das Harmonische Oszillatorpotenzial wird aber nicht nur aus didaktischen Gründen so intensiv behandelt. In vielen Fällen, wo es etwa um eine kleine Auslenkung eines Systems aus seiner Ruhelage geht, liefert das Potenzial des Harmonischen Oszillators eine gute Näherung für die Lösung des exakten Problems. Als ein Beispiel hierfür wollen wir kurz die potentielle Energie eines Moleküls aus zwei Atomen als Funktion des Abstandes der beiden Atome  $r$  betrachten. Dieses Potenzial hat etwa den Verlauf wie er in Figur 5.5 skizziert ist. Das Potenzial ist attraktiv für mittlere Abstände mit  $V \rightarrow 0$  für  $r \rightarrow \infty$ . Wegen der positiven Ladungen der Atomkerne muß Energie aufgebracht werden, um die Atome sehr nahe zusammenzubringen, das Potenzial wird also für kleine  $r$  positiv. Läßt man das Molekül "in Ruhe", so werden die Atome den optimalen Abstand  $r_0$  einnehmen, bei dem das Potenzial das Minimum aufweist. Stört man dieses Molekül ein wenig, indem man es etwa mit geeigneten elektromagnetischen Wellen bestrahlt, so kann man das Molekül aus diese Ruhelage bringen und es zu kleinen Schwingungen um die Gleichgewichtslage anregen. Für solch kleine Auslenkungen aus der Ruhelage ist es häufig ausreichend, das Potenzial  $V(r)$  durch die Taylorentwicklung um den Ruhepunkt  $r_0$  anzunähern

$$\begin{aligned} V(r) &\approx V(r_0) + \left. \frac{dV}{dr} \right|_{r=r_0} (r - r_0) + \frac{1}{2} \left. \frac{d^2V}{dr^2} \right|_{r=r_0} (r - r_0)^2 + \dots \\ &\approx \frac{1}{2} \kappa x^2 + \text{Konst.} \end{aligned}$$

Da das Potenzial  $V(r)$  an der Stelle  $r = r_0$  ein Minimum besitzt, wird an dieser Stelle die Ableitung identisch 0, sodass neben der Konstanten  $V(r_0)$ , die wir durch Umnormierung der Energie zu 0 machen können, in dieser Näherung nur der quadratische Term proportional  $(r - r_0)^2$  auftritt. Benennen wir diese Auslenkung aus der Ruhelage  $(r - r_0)$  um in  $x$ , so entspricht die Näherung also gerade dem Oszillatorpotenzial aus (5.62).

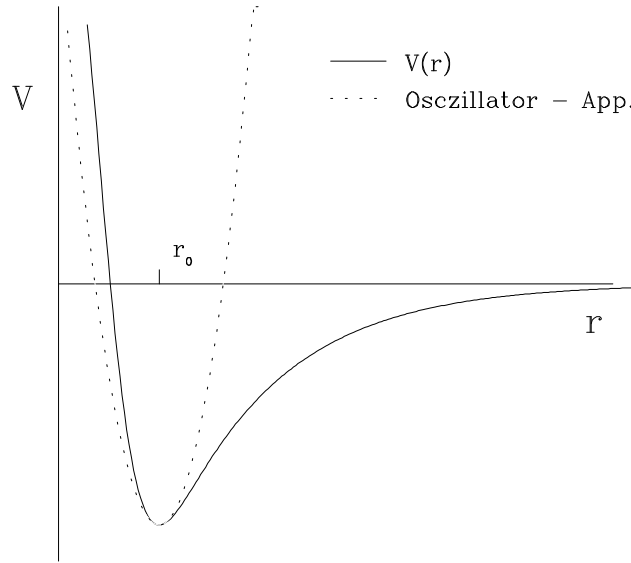


Abbildung 5.5: Schematische Darstellung der potentiellen Energie  $V(r)$  für ein Molekül als Funktion des Abstandes  $r$  zwischen den Atomen. Vergleich mit der Oszillatornäherung für dieses Potenzial.

Für die Bewegung eines Teilchens  $m$  in einem Harmonischen Oszillatorpotenzial der Form (5.62) ergibt sich also analog zu (5.45) die stationäre Schrödingergleichung

$$\left\{ -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2} m \omega^2 x^2 \right\} \Phi(x) = E \Phi(x) \quad (5.64)$$

Zur Lösung dieser Differentialgleichung führen wir eine geeignetere Variable ein mit

$$q = \frac{x}{b} \quad (5.65)$$

und der Oszillatorlänge

$$b := \sqrt{\frac{\hbar}{m\omega}} = \sqrt{\frac{\hbar c^2}{m c^2 \omega}} \quad (5.66)$$

Die zweite Gleichung nutzen wir zur Abschätzung der Dimension von  $b$ . Da die Lichtgeschwindigkeit  $c$  die Dimension [Länge / Zeit], das Produkt  $\hbar c$  [Energie  $\times$  Länge], die Winkelgeschwindigkeit  $\omega$  [1 / Zeit] und  $m c^2$  einer Energie entspricht, ergibt sich für die Oszillatorlänge  $b$  also die Dimension einer Länge und die Variable  $q$  ist dimensionslos. Substituiert man nun die Koordinate  $x$  in (5.64) durch  $q$  ergibt sich die Differentialgleichung

$$-\frac{\hbar\omega}{2} \frac{d^2\Phi}{dq^2} + \frac{\hbar\omega}{2} q^2 \Phi = E \Phi$$

Aus dem Vergleich der linken mit der rechten Seite dieser Gleichung ersieht man, dass  $\hbar\omega$  die Dimension einer Energie  $E$  besitzt. Dividiert man diese Gleichung durch  $\hbar\omega/2$  so ergibt sich die Differentialgleichung zur Bestimmung von  $\Phi$

$$\frac{d^2\Phi(q)}{dq^2} = q^2 \Phi(q) - \epsilon \Phi(q) \quad \text{mit } \epsilon := \frac{2E}{\hbar\omega} \quad (5.67)$$

Zur Lösung dieser Differentialgleichung betrachten wir diese zunächst einmal im asymptotischen Bereich  $q^2 \rightarrow \infty$ . In diesem Grenzfall ist auf der rechten Seite von (5.67) die Konstante  $\epsilon$  vernachlässigbar gegen  $q^2$  und es ergibt sich für  $\Phi(q)$  in diesem asymptotischen Bereich die Lösung

$$\Phi q \underset{q^2 \rightarrow \infty}{\rightleftharpoons} \alpha e^{\pm q^2/2} \quad (5.68)$$

Von diesen beiden Lösungen für das asymptotische Verhalten müssen wir die Lösung  $\Phi = \alpha \exp(+q^2/2)$  verwerfen. In diesem Fall würde sich ja für die Wahrscheinlichkeitsdichte  $\rho = \Phi^* \Phi$  ein divergierender Wert im Grenzfall  $q$  beziehungsweise  $x = \pm\infty$ , was natürlich dem zu beschreibenden Objekt, Teilchen gebunden in einem Oszillatorpotenzial, widerspricht. Es sind also die physikalischen Anforderungen, die diese mathematisch mögliche Lösung ausschliessen. Damit ergibt sich für  $\Phi(q)$  der Ansatz

$$\Phi(q) = h(q)e^{-q^2/2} \quad (5.69)$$

Die unbekannte Funktion  $h(q)$  darf keine Pole besitzen, damit die Wahrscheinlichkeitsdichte  $\rho$  an keiner Stelle divergiert. Somit kann  $h(q)$  also durch eine Taylorreihe entwickelt nach Potenzen von  $q$  dargestellt werden. Die Funktion  $h(q)$  darf außerdem für große Werte von  $|q|$  nicht so dominant werden, dass notwendige asymptotische Verhalten  $\Phi(q) \sim \exp(-q^2/2)$  im Bereich  $q^2 \rightarrow \infty$  gestört wird. Daraus ergibt sich dass die Taylorreihe für  $h(q)$  nur Terme mit endlichen Potenzen  $q^n$  enthalten darf<sup>1</sup>, sodass  $h(q)$  also durch ein Polynome vom Grade  $n$

$$h(q) = \sum_{\nu=0}^n a_{\nu} q^{\nu} \quad (5.70)$$

dargestellt wird. Setzt man nun den Ansatz (5.69) in die Differentialgleichung (5.67) ein, so ergibt sich

$$\begin{aligned} \frac{d^2 \Phi}{dq^2} &= \left[ \frac{d^2 h}{dq^2} - 2q \frac{dh}{dq} + q^2 h - h \right] e^{-q^2/2} \\ &= \left[ q^2 h - \epsilon h \right] e^{-q^2/2} \end{aligned}$$

Nach Multiplikation dieser Gleichung mit  $\exp(+q^2/2)$  erhält man

$$\frac{d^2 h(q)}{dq^2} - 2q \frac{dh(q)}{dq} + (\epsilon - 1)h(q) = 0.$$

Setzt man in diese Bestimmungsgleichung für  $h(q)$  den Polynomansatz (5.70) ein so ergibt sich

$$\sum_{\nu=2}^n \nu(\nu-1)a_{\nu} q^{\nu-2} - \sum_{\mu=1}^n 2\mu a_{\mu} q^{\mu-1} + (\epsilon-1) \sum_{\kappa=0}^n a_{\kappa} q^{\kappa} = 0$$

Die Summationsindices  $\nu, \mu$  und  $\kappa$  werden nun so umbenannt, dass die Exponenten von  $q$  sich entsprechen

$$\sum_{\nu=0}^n q^{\nu} [(\nu+2)(\nu+1)a_{\nu+2} - 2\nu a_{\nu} + (\epsilon-1)a_{\nu}] = 0$$

<sup>1</sup>Eine etwas sorgfältigere Analyse dieses asymptotischen Verhaltens findet sich z.B. im Kapitel 5 des Buches "Quantenphysik" von S.Gasiorowicz

Der Ausdruck auf der rechten Seite ist genau dann identisch 0 für alle Werte von  $q$ , wenn die Ausdrücke in den eckigen Klammern für alle  $\nu$  identisch 0 sind. Dies ergibt eine Rekursionsformel für die Koeffizienten  $a_\nu$

$$a_{\nu+2} = a_\nu \frac{2\nu - \epsilon + 1}{(\nu + 2)(\nu + 1)} \quad (5.71)$$

Damit diese Rekursionsformel Koeffizienten ungleich null nur liefert für Indices  $\nu \leq n$ , mit  $n$  dem Grad des Polynoms  $h$  (siehe (5.70)) ergibt sich

$$a_{n+2} = 0 = a_n \frac{2n - \epsilon + 1}{(n + 2)(n + 1)} \quad \text{bei } a_n \neq 0$$

Damit also die Schrödingergleichung der Form (5.69) mit einem Polynom vom endlichen Grade  $n$  existiert muss also gelten

$$2n - \epsilon + 1 = 0$$

beziehungsweise ergibt sich mit (5.67 für die Energie dieser stationären Lösung

$$E_n = \frac{\hbar\omega}{2}\epsilon = \frac{\hbar\omega}{2}(2n + 1) \quad (5.72)$$

Wir erhalten also auch für ein Teilchen, dass im Potenzial eines Harmonischen Oszillators gebunden ist stationäre Lösungen der Schrödingergleichung nur für die diskreten Energien, die (5.72) erfüllen.

Im Folgenden sollen zunächst die Eigenschaften der stationären Lösungen des Harmonischen Oszillatorproblems für einige niedrige Werte von  $n$  diskutiert werden.

**n=0** Die niedrigste Energie, die (5.72) genügt ergibt sich für  $n = 0$ :

$$E_0 = \frac{\hbar\omega}{2} \quad (5.73)$$

Die zugehörige Wellenfunktion ergibt sich aus (5.69), wobei für  $h(q)$  ein Polynom vom Grade  $n = 0$ , also eine Konstante einzusetzen ist:

$$\Phi_0 = a_0 e^{-\frac{q^2}{2}} = a_0 e^{-\frac{x^2}{2b^2}} \quad (5.74)$$

(vergleiche die Definition von  $q$  in (5.65)). Die Konstante  $a_0$  ergibt sich aus der Normierungsbedingung

$$\begin{aligned} 1 &= \int_{-\infty}^{+\infty} dx \Phi_0^*(x) \Phi_0(x) \\ &= a_0^2 \int_{-\infty}^{+\infty} dx e^{-\frac{x^2}{b^2}} \\ &= a_0^2 b \sqrt{\pi} \end{aligned}$$

Die resultierende Wellenfunktion  $\Phi_0(x)$  ist in Abbildung 5.6 dargestellt. Die Ordinate für die Darstellung der Wellenfunktion ist dabei so gewählt, dass sie die Parabel

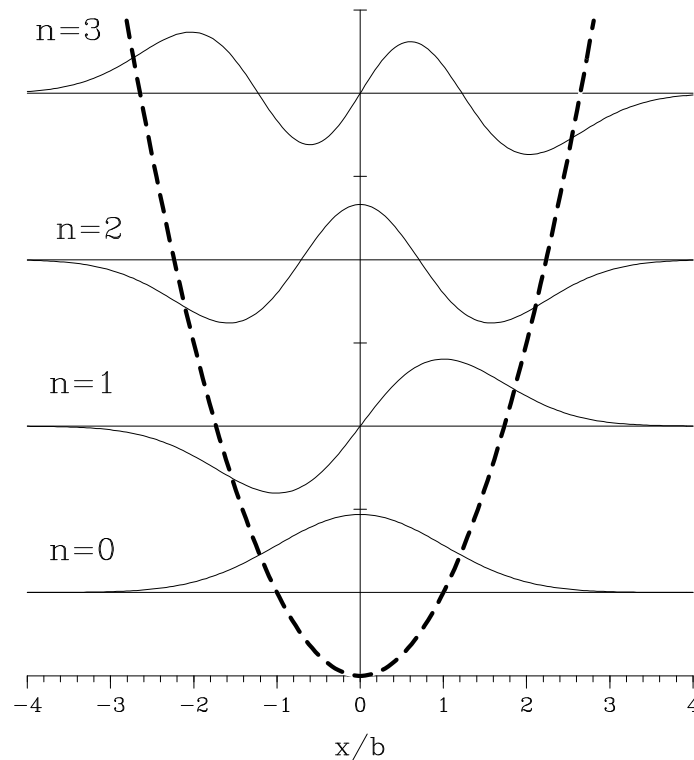


Abbildung 5.6: Die Wellenfunktionen des Harmonischen Oszillators für  $n=0,1,2,3$

des Harmonischen Oszillatorpotenzials an der Stelle  $x_0$  schneidet, an der die Energie  $E_0$  gerade der potenziellen Energie entspricht

$$E_0 = \frac{\hbar\omega}{2} = \frac{1}{2}m\omega^2 x_0^2$$

Klassisch ist ein Aufenthalt des Masseteilchens verboten an Stellen mit  $|x| > x_0$  da dies eine negative kinetische Energie erfordern würde. Wir beobachten aber auch hier, dass die Wahrscheinlichkeitsamplitude  $\Phi_0(x)$  durchaus von 0 verschieden ist in diesem klassisch verbotenen Bereich. Ähnliches gilt natürlich auch für die im folgenden diskutierten Zustände mit  $n > 0$ .

**n=1** Die Energie für das Masseteilchen im ersten angeregten Zustand des Oszillatorpotenzials ergibt sich aus (5.72) für  $n = 1$  mit  $E_1 = 3/2 \hbar\omega$ . Entsprechend höher liegt auch die Ordinate zur Darstellung der Wellenfunktion  $\Phi_1(x)$  in Figur 5.6. Als Ansatz für die Wellenfunktion schreiben wir

$$\Phi_1(x) = \left[ a_1 \frac{x}{b} + a_0 \right] e^{-\frac{x^2}{2b^2}}$$

Die Rekursionsformel (5.71) für die Koeffizienten  $a_\nu$  liefert bei der Energie  $E_1$  für  $a_3$  und damit auch für  $a_5, a_7$  usw. den Wert 0. Wäre  $a_0$  von Null verschieden, so würden sich auch für alle anderen Entwicklungskoeffizienten zu graden Potenzen von  $x$  von Null verschiedene Werte ergeben. Dies steht aber im Widerspruch zum Ansatz, dass  $h(q)$  in (5.69) ein Polynom endlichen Grades ist. Also muss  $a_0 = 0$

sein. Dieses Beispiel lässt sich verallgemeinern: Das Polynom  $h(q)$  kann entweder nur gerade Potenzen in  $q$  beziehungsweise  $x$  oder nur ungerade Potenzen besitzen, da sich bei vorgegebener Energie jeweils nur eine Abbruchbedingung für die Koeffizienten  $a_\nu$  ergibt. Der einzig noch unbekannte Koeffizient  $a_1$  ergibt sich wieder aus der Normierungsbedingung für die Wellenfunktion und man erhält

$$\Phi_1(x) = \frac{1}{\sqrt{2b\sqrt{\pi}}} 2 \frac{x}{b} e^{-\frac{x^2}{2b^2}} \quad (5.75)$$

also eine Wellenfunktion mit genau einer Nullstelle (abgesehen vom asymptotischen Verhalten bei  $|x| \rightarrow \infty$ ) bei  $x = 0$ . (siehe auch Abbildung 5.6).

**n=2** Für  $n=2$  ergibt sich die Energie  $E_2 = 5/2 \hbar \omega$  und das Polynom  $h(q)$  im Ansatz für die Wellenfunktion (5.69) sollte nur Terme mit geraden Potenzen von  $q$  enthalten bis maximal  $\nu = 2$ . Wir können also schreiben

$$\Phi_2(x) = a_0 \left[ 1 + \frac{a_2}{a_0} q^2 \right] e^{-\frac{q^2}{2}}.$$

Die Rekursionsformel (5.71) liefert das Verhältnis

$$\frac{a_2}{a_0} = \frac{0 - 5 + 1}{2} = -2$$

und die Konstante  $a_0$  wird wieder durch die Normierung der Wellenfunktion festgelegt. Dies ergibt

$$\Phi_2(x) = \frac{2}{\sqrt{8b\sqrt{\pi}}} \left[ 1 - 2 \frac{x^2}{b^2} \right] e^{-\frac{x^2}{2b^2}}, \quad (5.76)$$

also eine Wellenfunktion mit 2 Nullstellen für endliche Werte von  $x$ . Entsprechend können auch die Wellenfunktionen für größere Werte von  $n$  konstruiert werden.

An dieser Stelle soll noch einmal die Tatsache aufgegriffen werden, dass das Polynom  $h(q)$  im Ansatz für die Wellenfunktion (5.69) entweder nur gerade oder nur ungerade Potenzen von  $q$  enthält. Wie man auch aus der Abbildung 5.6 entnehmen kann führt dies ja dazu, dass die Wellenfunktionen für gerade Werte von  $n$  spiegelsymmetrisch um den Punkt  $x = 0$  sind. Für ungerade Werte von  $n$  kehrt sich bei einer Spiegelung am Koordinatenursprung genau das Vorzeichen der Wellenfunktion um. Es gilt also in jedem Fall die Symmetriebedingung

$$\Phi_n(-x) = (-)^n \Phi(x) \quad (5.77)$$

Man sagt: Die Oszillatorfunktionen  $\Phi_n$  haben positive (negative) Parität für gerade (ungerade) Zahlen  $n$ , d.h. bei Spiegelung der Koordinate am Koordinatenursprung  $x \rightarrow -x$  ändert sich (bis auf das Vorzeichen bei negativer Parität) die Wellenfunktion nicht. Wir wollen uns nun davon überzeugen, dass dieses Ergebnis kein Zufall ist sondern in der Symmetrie des Potentials begründet ist.

Dazu nehmen wir an, wir haben eine stationäre Lösung der Schrödingergleichung zur Energie  $E$ ,  $\varphi(x)$ , gefunden

$$\left[ -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2} m \omega^2 x^2 \right] \varphi(x) = E \varphi(x)$$



In dieser Gleichung können wir eine Koordinatentransformation  $x \rightarrow -x$  durchführen. Wir betrachten also bei dieser Transformation das ganze System im Spiegel. Nach dieser Transformation stellt sich die stationäre Schrödingergleichung dar in der Form

$$\left[ -\frac{\hbar^2}{2m} \frac{d^2}{d(-x)^2} + \frac{1}{2} m \omega^2 (-x)^2 \right] \varphi(-x) = E \varphi(-x)$$

beziehungsweise

$$\left[ -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2} m \omega^2 x^2 \right] \varphi(-x) = E \varphi(-x)$$

Wir sehen also, dass neben  $\varphi(x)$  auch  $\varphi(-x)$  eine Lösung zur gleichen Energie  $E$  ist. Dies gilt nicht nur für das Oszillatorpotenzial sondern für alle Potenziale, die sich durch eine Spiegeltransformation nicht ändern, für die also gilt  $V(-x) = V(x)$ . Wenn  $\varphi(x)$  und  $\varphi(-x)$  Lösungen sind, sind auch die Linearkombinationen

$$\begin{aligned} \Phi_1(x) &= \frac{1}{2} [\varphi(x) + \varphi(-x)] \\ \Phi_2(x) &= \frac{1}{2} [\varphi(x) - \varphi(-x)] \end{aligned}$$

Lösungen der Schrödingergleichung zur selben Energie. Diese Lösungen haben aber eine wohldefinierte Parität:  $\Phi_1(-x) = \Phi_1(x)$  und  $\Phi_2(-x) = -\Phi_2(x)$ . Wir sind also immer in der Lage stationäre Lösungen der Schrödingergleichung zu finden mit wohldefinierter Parität.

Bei Wellenfunktion positiver wie auch negativer Parität gilt in jedem Fall

$$\rho(x) = \Phi^*(x)\Phi(x) = \Phi^*(-x)\Phi(-x) = \rho(-x)$$

die Wahrscheinlichkeitsdichte ist symmetrisch um den Koordinatenursprung. Dies reflektiert natürlich wieder die Symmetrie des betrachteten Potenzials  $V(x) = V(-x)$ .

Bisher haben wir nur die stationären Lösungen der Schrödingergleichung betrachtet. Auch wenn wir den ortsabhängigen Anteil dieser stationären Lösungen  $\Phi_n(x)$  nach (5.39) ergänzen durch die entsprechende Zeitabhängigkeit

$$\psi_n(x, t) = \Phi_n(x) e^{-i \frac{E_n t}{\hbar}} = \Phi_n(x) e^{-i(n+\frac{1}{2})\omega t} \quad (5.78)$$

so ergibt sich wie ja für jede stationäre Lösung im vorhergehenden Abschnitt diskutiert haben, dass die daraus resultierende Wahrscheinlichkeitsdichte genau so wie alle Erwartungswerte zeitunabhängig sind. Insbesondere erhalten wir für den Erwartungswert der Ortskoordinate  $x$  des betrachteten Teilchens

$$\langle \psi_n(x, t) | x | \psi_n(x, t) \rangle = 0,$$

das Teilchen bewegt sich nicht sondern "ruht" im Zentrum des Potenzials. Um eine Bewegung des Systems zu beschreiben, müssen wir entsprechend (5.60) eine Linearkombination von 2 stationären Lösungen betrachten, also z.B.

$$\begin{aligned} \Psi(x, t) &= \frac{1}{\sqrt{2}} [\psi_0(x, t) + \psi_1(x, t)] \\ &= \frac{1}{\sqrt{2b\sqrt{\pi}}} \left[ e^{-\frac{x^2}{2b^2}} e^{-i\frac{\omega t}{2}} + \sqrt{2} \frac{x}{b} e^{-\frac{x^2}{2b^2}} e^{-i\frac{3\omega t}{2}} \right] \end{aligned} \quad (5.79)$$

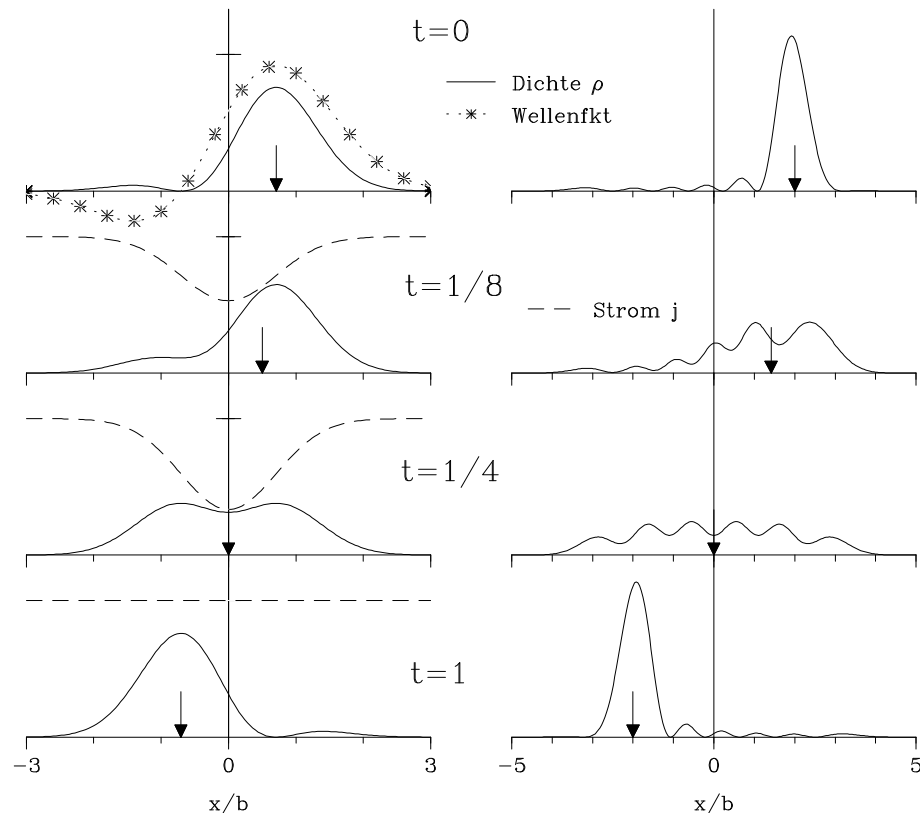


Abbildung 5.7: Wahrscheinlichkeitsdichten (durchgezogene Linie), Stromdichten (gestrichelte Linie) und Wellenfunktion für ein Teilchen im Potenzial des Harmonischen Oszillators zu verschiedenen Zeiten ( $t = 0$ ,  $t = 1/8T$ ,  $t = 1/4T$ ,  $t = 1/2T$  mit  $T$  der Schwingungsperiode des Oszillators). Im linken Teil sind die Ergebnisse für die Wellenfunktion (5.79) aufgetragen, im rechten Teil eine Wellenfunktion für ein Teilchen, das zur Zeit  $t = 0$  bei  $x = 2b$  lokalisiert ist. Weitere Erklärungen im Text

Zur Startzeit  $t = 0$  ist die Wellenfunktion rein reellwertig, positiv für  $x > 0$ , besitzt eine Nullstelle bei  $x/b = -1/\sqrt{2}$  und ist negativ für kleinere Werte von  $x$  (siehe punktierte Linie im oberen Teil der linken Teilfigur von Abbildung 5.7). Die aus dieser Wellenfunktion berechnete Wahrscheinlichkeitsdichte besitzt ein Hauptmaximum bei positiven  $x$  und ein kleines, in der Abbildung 5.7 kaum sichtbares Nebenmaximum bei negativen Werten von  $x$ . Der Erwartungswert für die Position  $x$  beträgt  $x = \sqrt{2}b$  und ist der Abbildung durch einen Pfeil angedeutet. Da die Wellenfunktion zur Zeit  $t = 0$  rein reellwertig ist, ist die Stromdichte (5.33)

$$j(x, t) = \frac{\hbar}{2mi} \left[ \Psi^* \frac{d}{dx} \Psi - \Psi \frac{d}{dx} \Psi^* \right]$$

identisch null. Da die Erwartungswerte ja den Observablen der Klassischen Mechanik entsprechen sollen, beschreiben wir also mit der Wellenfunktion (5.79) ein Teilchen im Potenzial des Harmonischen Oszillators, das sich zur Startzeit  $t = 0$  am Punkt  $x = \sqrt{2}b$  befindet mit der Geschwindigkeit (entsprechend der integrierten Stromdichte) von  $v = 0$ . Für  $t \neq 0$  wird die Wellenfunktion (5.79) komplexwertig. Dies bedeutet, dass auch die

Stromdichte  $j$  von Null verschiedene Werte annehmen kann. Dies ist in der linken Spalte von Abbildung (5.7) für  $t = 1/8T$ ,  $t = 1/4T$ ,  $t = 1/2T$  dargestellt wobei  $T$  für die Schwingungsperiode des Oszillators steht,  $T = 2\pi/\omega$ . Die Stromdichte geht für  $x \rightarrow \pm\infty$  gegen 0, und nimmt für die in unserem Beispiel angenommenen Zeiten ( $t \leq T/2$ ) stets negative Werte an. Dies reflektiert die Tatsache, dass bei der Beschreibung der Schwingung im Rahmen der Klassischen Mechanik, das Punktteilchen “nach links schwingt” also eine negative Geschwindigkeit besitzt. Parallel dazu ändert sich auch die Wahrscheinlichkeitsdichte, der Aufenthaltsort des Teilchens bewegt sich ebenfalls nach links. Berechnet man den Erwartungswert für die Koordinate  $x$

$$\langle x \rangle = \int_{-\infty}^{\infty} dx \Psi^*(x, t) x \Psi(x, t),$$

so erhält man die Werte von  $x$ , die jeweils durch den Pfeil in der Abbildung 5.7 herausgehoben sind. Diese Erwartungswerte verändern sich mit der Zeit wie wir es von der Klassischen Mechanik erwarten (5.63)

$$\langle x \rangle (t) = A \cos \omega t \quad \text{mit} \quad A = \sqrt{b}.$$

Damit wird nur das Ehrenfestsche Theorem (siehe Paragraph 2.2) verifiziert, nach dem sich ja die Erwartungswerte von dynamischen Variablen (hier die Ortskoordinate  $x$ ) so verhalten, wie es die Klassische Mechanik vorhersagt.

Die Wellenfunktion (5.79) ist nur ein Beispiel für die allgemeine Überlagerung von stationären Lösungen der Schrödingergleichung im Fall des Harmonischen Oszillators

$$\Psi(x, t) = \sum_{n=0}^N \alpha_n \Phi_n(x) e^{i(n+\frac{1}{2})\omega t}. \quad (5.80)$$

Durch die Wahl der Koeffizienten  $\alpha_n$  kann man die Startbedingungen, das heißt unter anderem die Position und die Geschwindigkeit des Massenpunktes zur Zeit  $t = 0$  festlegen. Als ein zweites Beispiel wählen wir

$$\alpha_n = \begin{cases} C \Phi_n(x = 2b) & \text{für } n = 0 \dots 6 \\ 0 & \text{sonst.} \end{cases} \quad (5.81)$$

Die Konstante  $C$  wird so bestimmt, dass die gesamte Wellenfunktion normiert ist. Wir werden später sehen, dass durch die Wahl  $\alpha_n = \Phi_n(x_0)$  gewährleistet wird, dass das Teilchen zur Zeit  $t = 0$  “möglichst gut” bei  $x = x_0$  lokalisiert ist. Je mehr stationäre Lösungen im Ansatz (5.80) berücksichtigt werden, je größer also  $N$ , desto ausgeprägter ist das Maximum der Dichteverteilung bei  $x = x_0$  ( $x_0 = 2b$ ) in unserem Beispiel. Da die Koeffizienten  $\alpha_n$  reell sind, ist außerdem gewährleistet, daß zur Zeit  $t = 0$  die Wahrscheinlichkeitsstromdichte identisch 0 ist. Wir beschreiben also auch in diesem Falle eine Oszillatorschwingung, bei der zur Startzeit das Teilchen maximale Auslenkung  $x = 2b$  und Geschwindigkeit  $v = 0$  besitzt. Die Änderung der Wahrscheinlichkeitsdichte und der Stromdichte  $j$  mit der Zeit  $t$  sind für dieses Beispiel in der rechten Spalte der Figur 5.7 dargestellt.

Der Ansatz für die Wellenfunktion (5.80) entspricht im Prinzip dem Wellenpaket von (5.13). Wurden dort verschiedene ebene Wellen, also Lösungen der freien Schrödingergleichung zur Gesamtwellenfunktion überlagert, so sind es hier, in (5.80) verschiedene Lösungen der stationären Schrödingergleichung des Harmonischen Oszillators. Man spricht deshalb auch von einem Wellenpaket aus Harmonischen Oszillatorfunktionen. Die Beispiele in Figur 5.7 zeigen, wie sich dieses Wellenpaket periodisch mit der Zeit ändert.

### Zusammenfassung

- Im Fall des Harmonischen Oszillatorpotenzials  $V(x) = 1/2m\omega^2x^2$  existieren stationäre Lösungen nur für diskrete Energien

$$E_n = \hbar\omega \left( n + \frac{1}{2} \right) \quad \text{für } n = 0, 1, \dots$$

- Die niedrigst mögliche Energie  $E_0 = \hbar\omega/2$ , die sogenannte Nullpunktsenergie, ist also größer als 0. In der Klassischen Mechanik gibt es natürlich die Lösung des Teilchens, das im Minimum des Oszillatorpotenzials  $x = 0$  ruht und somit die Energie 0 besitzt. Eine solche Lösung ist im Widerspruch zu Grundprinzipien der Quantenmechanik: Soll das Teilchen die potentielle Energie 0 besitzen, so muss nicht nur der Erwartungswert  $\langle x \rangle$  sondern auch  $\langle x^2 \rangle = 0$  sein. Damit ist das mittlere Schwankungsquadrat  $\Delta^2x = \langle x^2 \rangle - \langle x \rangle^2$  für die Ortskoordinate 0. Bei der Diskussion der ebenen Wellen haben wir bereits gesehen, dass eine solche scharfe Lokalisierung verknüpft ist mit einer Impulsunschärfe. Wir werden im nächsten Kapitel die Heisenbergsche Unschärferelation kennenlernen, wonach allgemein eine genaue Lokalisierung von  $x$  mit unendlichen Impulsunschärfe  $\Delta^2p = \langle p^2 \rangle = \infty$  verknüpft ist. In diesem Fall wäre die kinetische Energie  $\langle p^2 \rangle / 2m$  also unendlich groß. Die Nullpunktsenergie reflektiert also nur die Tatsache, dass in der Quantenmechanik Ort und Impuls nicht gleichzeitig identisch 0 sein können.
- Die Wellenfunktionen der stationären Lösungen der Schrödingergleichung hängen nur ab von dem Verhältnis  $x/b = q$  mit der Oszillatorlänge  $b$  aus (5.66). Sie lassen sich schreiben als Produkt einer Normierungskonstanten, einem Polynom vom Grade  $n$ , dem Hermite Polynom  $H_n$  und einer Exponentialfunktion

$$\Phi_n(x) = \frac{1}{\sqrt{2^n n! \sqrt{\pi} b}} H_n(q) e^{-\frac{q^2}{2}} \quad (5.82)$$

Dabei besitzt das Hermite Polynom, je nach dem Grade  $n$  nur Terme mit geraden oder ungeraden Potenzen von  $q$ . Entsprechend hat die Funktion  $\Phi_n(x)$  positive oder negative Parität.

- Zur Beschreibung eines Teilchens, das im Oszillatorpotenzial schwingt, muss ein Wellenpaket aus Oszillatorzuständen (5.80) betrachtet werden. Die Koeffizienten werden durch die Startbedingungen zur Zeit  $t = 0$  festgelegt. Zur Beschreibung einer Bewegung also eines Systems, dass sich mit der Zeit ändert, müssen Zustände mit verschiedenen Energien überlagert werden. Die Energie ist also in diesem Fall nicht mehr scharf definiert.

Zum Abschluss dieses Kapitels präsentieren wir ein Protokoll einer Maple Simulation, bei der die stationären Lösungen zum harmonischen Oszillator über die Hermite Polynome berechnet und normiert werden. Nach einer Darstellung der Funktion wird eine geeignete Linearkombination aus Oszillatorfunktionen für  $n = 0$  bis  $n = 5$  bestimmt. Dies Linearkombination liefert für  $t = 0$  eine Wahrscheinlichkeitsdichte, die bei  $x = 1.5$  ein Maximum aufweist. Die Animation unter Maple liefert eine Darstellung der Wahrscheinlichkeitsdichte als Funktion der Zeit.

```
> with(orthopoly);
```

```
[G, H, L, P, T, U]
```

**Kommentar:** Durch dieses Kommando stehen verschiedene Polynome zur Verfügung, unter anderem die Hermite Polynome. Durch den Aufruf  $H(n, x)$  wird das Hermite Polynom  $H_n$  an der Stelle  $x$  berechnet. Somit erhalten wir also die nicht normierten Oszillatorfunktionen (wir definieren die Darstellung von  $x$  so, dass die Oszillatorkonstante  $b = 1$ ) durch die Zuweisung:

```
> psiH0:=(x,n)->exp(-x^2/2)*H(n,x);
```

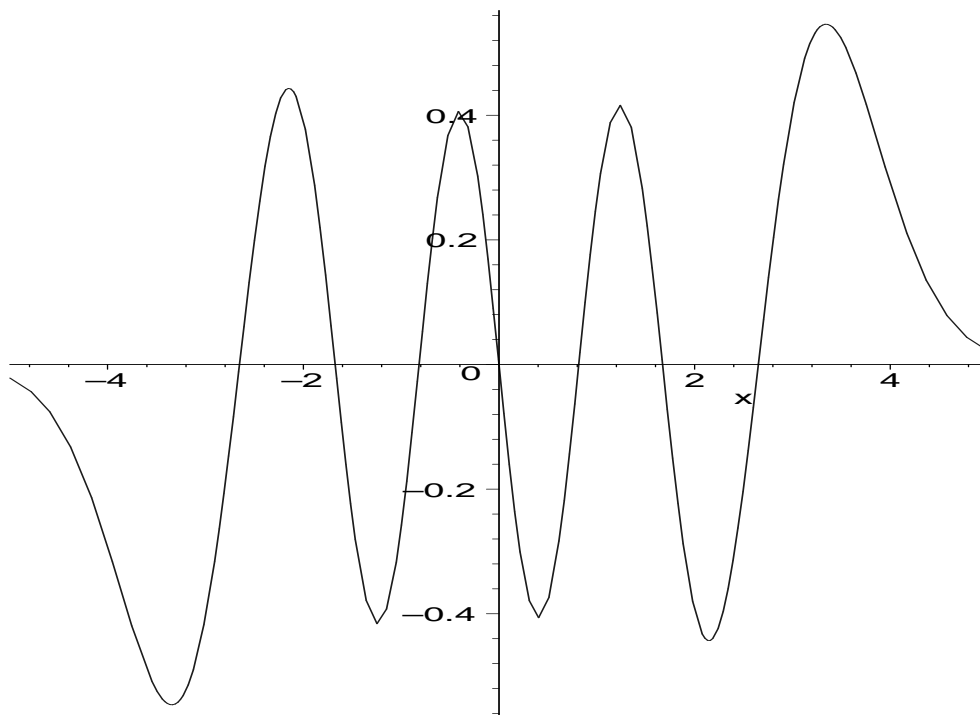
$$psiHO := (x, n) \rightarrow e^{(-1/2x^2)} H(n, x)$$

```
> psiH0n:=(x,n)->psiH0(x,n)/sqrt(int(psiH0(y,n)^2,y=-infinity..infinity
> ));
```

$$psiH0n := (x, n) \rightarrow \frac{psiHO(x, n)}{\sqrt{\int_{-\infty}^{\infty} psiHO(y, n)^2 dy}}$$

**Kommentar:** Darstellung der Oszillatorfunktion für  $n = 7$

```
> plot(psiH0n(x,7),x=-5..5,color=blue);
```

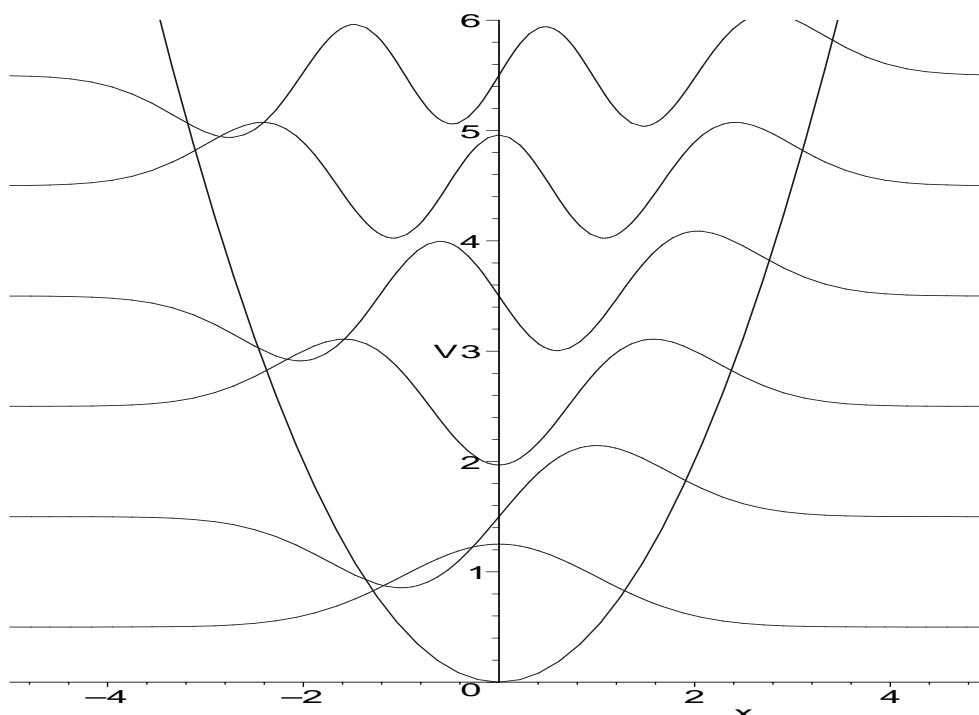


**Kommentar:** Darstellung von Oszillatorfunktionen für  $n = 0..5$

```
> wfpl:=seq(psiH0n(x,n)+(n+1/2),n=0..5);
```

$$\begin{aligned}
 wfpl &:= \frac{e^{(-1/2x^2)}}{\pi^{(1/4)}} + \frac{1}{2}, \frac{e^{(-1/2x^2)} x \sqrt{2}}{\pi^{(1/4)}} + \frac{3}{2}, \frac{1}{4} \frac{e^{(-1/2x^2)} (4x^2 - 2) \sqrt{2}}{\pi^{(1/4)}} + \frac{5}{2}, \\
 &\frac{1}{12} \frac{e^{(-1/2x^2)} (8x^3 - 12x) \sqrt{3}}{\pi^{(1/4)}} + \frac{7}{2}, \frac{1}{48} \frac{e^{(-1/2x^2)} (16x^4 - 48x^2 + 12) \sqrt{6}}{\pi^{(1/4)}} + \frac{9}{2}, \\
 &\frac{1}{240} \frac{e^{(-1/2x^2)} (32x^5 - 160x^3 + 120x) \sqrt{15}}{\pi^{(1/4)}} + \frac{11}{2}
 \end{aligned}$$

```
> plot({x^2/2,wfpl},x=-5..5,V=0..6,color=blue);
```

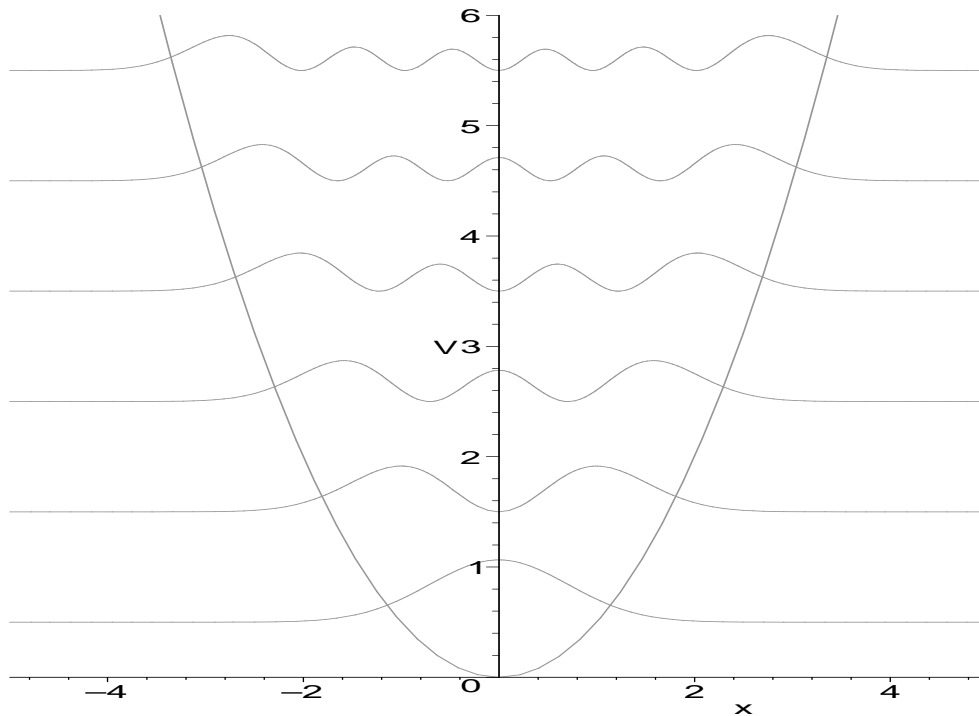


**Kommentar:** Darstellung von Wahrscheinlichkeitsdichten für die Oszillatorfunktionen für  $n = 0 \dots 5$

```
> wfpl:=seq(psiH0n(x,n)^2+(n+1/2),n=0..5);
```

$$\begin{aligned}
 wfpl &:= \frac{\%1}{\sqrt{\pi}} + \frac{1}{2}, 2 \frac{\%1 x^2}{\sqrt{\pi}} + \frac{3}{2}, \frac{1}{8} \frac{\%1 (4x^2 - 2)^2}{\sqrt{\pi}} + \frac{5}{2}, \frac{1}{48} \frac{\%1 (8x^3 - 12x)^2}{\sqrt{\pi}} + \frac{7}{2}, \\
 &\frac{1}{384} \frac{\%1 (16x^4 - 48x^2 + 12)^2}{\sqrt{\pi}} + \frac{9}{2}, \frac{1}{3840} \frac{\%1 (32x^5 - 160x^3 + 120x)^2}{\sqrt{\pi}} + \frac{11}{2} \\
 \%1 &:= (e^{(-1/2x^2)})^2
 \end{aligned}$$

```
> plot({x^2/2,wfpl},x=-5..5,V=0..6,color=green);
```



**Kommentar:** Berechnung von geeigneten Entwicklungskoeffizienten für eine Wellenfunktion mit einem Maximum der Wahrscheinlichkeitsdichte bei  $x = 1.5$ .

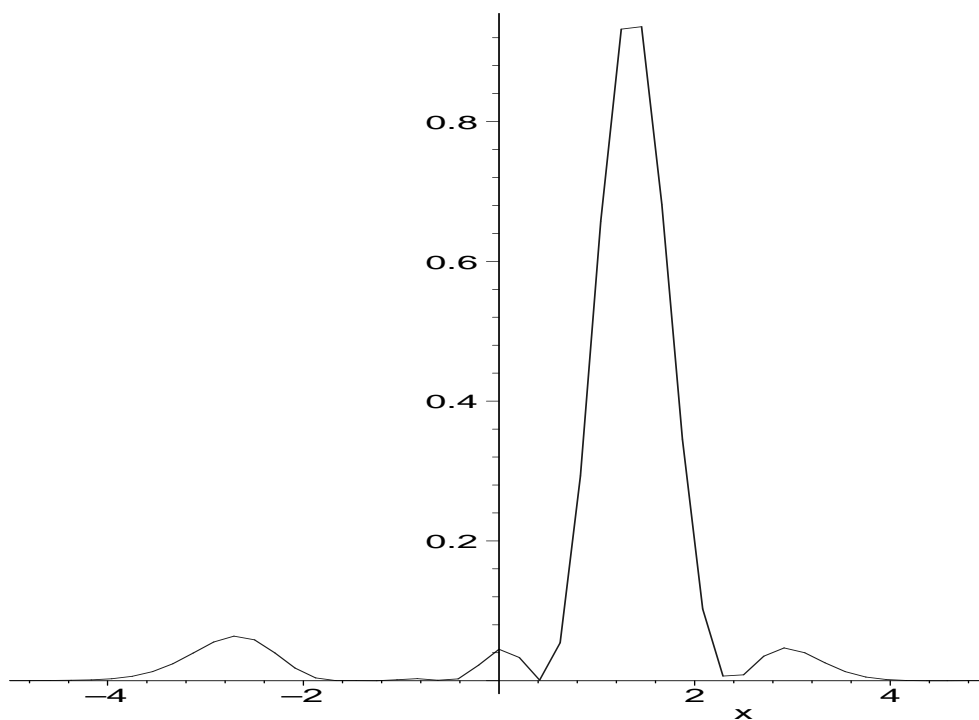
```

> a0:=evalf(psiH0n(1.5,0));
      a0 := .2438547613
> a1:=evalf(psiH0n(1.5,1));
      a1 := .5172940661
> a2:=evalf(psiH0n(1.5,2));
      a2 := .6035097438
> a3:=evalf(psiH0n(1.5,3));
      a3 := .3167766273
> a4:=evalf(psiH0n(1.5,4));
      a4 := -.1866624178
> a5:=evalf(psiH0n(1.5,5));
      a5 := -.4604171469
> Psi02:=(x,t)->a0*psiH0n(x,0)*exp(-I*(1/2)*t)+a1*psiH0n(x,1)*exp(-I*(3
> /2)*t)+a2*psiH0n(x,2)*exp(-I*(5/2)*t)+a3*psiH0n(3,1)*exp(-I*(7/2)*t)+a
> 4*psiH0n(x,4)*exp(-I*(9/2)*t)+a5*psiH0n(x,5)*exp(-I*(11/2)*t);
Psi02 := (x, t) -> a0 psiH0n(x, 0) e(-1/2 It) + a1 psiH0n(x, 1) e(-3/2 It)
+ a2 psiH0n(x, 2) e(-5/2 It) + a3 psiH0n(3, 1) e(-7/2 It) + a4 psiH0n(x, 4) e(-9/2 It)
+ a5 psiH0n(x, 5) e(-11/2 It)
> with(plots):

```

Warning, the name changecoords has been redefined

```
> animate(evalc(conjugate(Psi02(x,t))*Psi02(x,t)),x=-5..5,t=0..2*Pi,col  
> or=blue,frames=50);
```





# Kapitel 6

## Grundlagen der Quantenmechanik

Im letzten Kapitel wurden ausgehend von einer empirisch-phänomenologischen Basis einige Eigenschaften der Beschreibung physikalischer Systeme mit der Quantenmechanik aufgezeigt. In diesem Kapitel soll nun versucht werden, erste Grundlagen der Quantenmechanik zusammenzustellen und die mathematischen Werkzeuge zu erarbeiten. Wir werden sehen, dass die Darstellung eines Systems durch eine Wellenfunktion  $\Psi(x)$  oder der entsprechenden Impulsamplitude  $C(k)$  nur 2 Spezialfälle repräsentieren.

### 6.1 Zustände und Observable

In der Klassischen Mechanik beschreiben wir den Zustand eines physikalischen Systems zu einem Zeitpunkt  $t_0$  durch die Angabe der generalisierten Koordinaten  $q_i(t_0)$  und der dazugehörigen Impulse  $p_i(t_0)$ . Diese Angabe von Koordinaten und Impulsen definiert einen Punkt im  $2N$  dimensionalen Phasenraum, wobei  $N$  die Zahl der Freiheitsgrade des Systems bezeichnet. Wenn wir nun die Hamiltonfunktion  $H(p_i, q_i)$  kennen, so können wir die Hamilton'schen Bewegungsgleichungen

$$\begin{aligned}\frac{dq_i}{dt} &= \frac{\partial H}{\partial p_i} \\ \frac{dp_i}{dt} &= -\frac{\partial H}{\partial q_i}\end{aligned}$$

aufstellen und durch Integration dieser gekoppelten Differentialgleichungen auch die Koordinaten des Systemzustandes für  $t \neq t_0$  bestimmen. Die zeitliche Entwicklung des Systems wird also durch eine Trajektorie im Phasenraum beschrieben und diese Trajektorie ist bei gegebenen Startwerten  $q_i(t_0)$ ,  $p_i(t_0)$  eindeutig bestimmt. Kennen wir aber die Werte für diese generalisierten Koordinaten, so können wir auch die Werte für alle Beobachtungsgrößen, also alle dynamischen Variablen  $A(q_i, p_i, T)$ , bestimmen.

In der Quantenmechanik müssen wir den Zustand eines Systems anders charakterisieren. Bezeichnen wir einen solchen Zustand zunächst einmal ganz abstrakt durch

$$\text{Zustand des Systems} \iff |\Psi\rangle$$

so gibt es verschiedene Möglichkeiten diesen Zustand darzustellen:

$$|\Psi\rangle \Rightarrow \begin{cases} \Psi(x) := \langle x|\Psi\rangle & \text{Ortsdarstellung} \\ \tilde{\Psi}(k) = \langle k|\Psi\rangle & \text{Impulsdarstellung} \\ |\Psi\rangle = \sum_n c_n |\Phi_n\rangle & \text{Basis } \Phi_n \end{cases} \quad (6.1)$$

Im ersten Fall, der Ortsdarstellung, charakterisieren wir den Zustand durch die Wellenfunktion  $\Psi(x)$  (wir nehmen hier an, dass wir ein System haben, das nur durch eine Ortskoordinate  $x$  beschrieben wird, allgemein können wir  $x$  aber auch als Repräsentanten der generalisierten Koordinaten ansehen). Dies bedeutet, dass der Zustand  $|\Psi\rangle$  eindeutig durch eine Menge komplexer Zahlen  $\Psi(x)$  charakterisiert ist, wobei jeder Koordinate  $x$  genau eine komplexe Zahl  $\Psi(x)$  zugeordnet ist. Die Bezeichnung  $\Psi(x) = \langle x|\Psi\rangle$  verstehen wir zunächst einmal als eine synonyme Schreibweise. Die Bedeutung dieser Schreibweise wird erst dann klarer werden, wenn wir weiter unten den Klammersausdruck  $\langle a|b\rangle$  allgemeiner definieren.

Als Alternative zur Ortsdarstellung haben wir bereits im Abschnitt 2.1 die Impulsdarstellung kennengelernt. In dieser Darstellung ist also der Zustand des Systems beschrieben durch die Funktion  $\tilde{\Psi}(k)$ , in 2.1 auch  $C(k)$  genannt, also die Funktion, die jedem Impuls  $p = \hbar k$  eine komplexe Zahl  $\tilde{\Psi}(k) = C(k)$  zuweist. Wir werden gleich sehen, dass es neben diesen beiden uns bereits bekannten Möglichkeiten aber auch noch eine Vielzahl anderer Möglichkeiten gibt, den Zustand  $|\Psi\rangle$  durch komplexe Zahlen  $c_n$  zu charakterisieren, indem wir angeben, wie sich der betreffende Zustand als eine Linearkombination von bestimmten Zuständen  $|\Phi_n\rangle$  darstellen lässt. Auch eine solche Darstellung haben wir bereits kennengelernt, als wir z.B. in (5.80) einen Zustand des harmonischen Oszillators als Linearkombination von stationären Lösungen der Schrödingergleichung dargestellt haben.

Diese letzte Darstellung eines Zustandes  $|\Psi\rangle$  als Linearkombination beziehungsweise Überlagerung von anderen Zuständen  $|\Phi_n\rangle$  beruht auf dem empirischen Befund etwa des Experimentes am Doppelspalt (siehe Kapitel 1), bei dem die Zustände  $|\Phi_1\rangle$ : *Teilchen läuft durch Spalt 1* und  $|\Phi_2\rangle$ : *Teilchen läuft durch Spalt 2* zu einem gesamten Zustand oder Wellenfunktion überlagert wurden. Aus der mathematischen Sicht basiert diese Überlagerung auf der Eigenschaft dass die Schrödingergleichung eine lineare Differentialgleichung zur Bestimmung der Wellenfunktion ist. Sind nämlich bei einer solchen linearen Differentialgleichung die Wellenfunktionen  $\Phi_n$  jeweils eine Lösung, so gilt dies auch für die Linearkombination.

Dieses **Superpositionsprinzip** der Zustände in der Quantenmechanik legt es nahe, dass wir die möglichen Zustände eines Systemes als Elemente eines Vektorraumes auffassen. Genauer gesagt nennen wir  $|\Psi\rangle$  einen **Ket-Vektor** und betrachten diesen Zustand als ein Element eines Vektorraumes über den Körper der komplexen Zahlen  $\mathcal{C}$ . Dieser Vektorraum besitzt die Eigenschaften eines **Hilbertraumes**  $\mathcal{H}$ , die wir im folgenden noch ausführlicher darstellen wollen.

### 6.1.1 $\mathcal{H}$ ein linearer Vektorraum über $\mathcal{C}$

Zunächst bedeutet bedeutet dies ja:

- Für die Elemente  $|\Phi_i\rangle$  dieses Vektorraumes ist eine Addition und eine Multiplikation mit komplexen Zahlen  $\alpha, \beta$  definiert und es gilt, dass auch die damit gebildete

Linearkombination ein Element des Hilbertraumes ist:

$$|\Phi_1\rangle, |\Phi_2\rangle \in \mathcal{H}, \quad \alpha, \beta \in \mathcal{C} \quad \rightarrow \quad |\Psi\rangle = \alpha|\Phi_1\rangle + \beta|\Phi_2\rangle \in \mathcal{H} \quad (6.2)$$

Es gilt also das Superpositionsprinzip.

- Addition und Multiplikation sind kommutativ

$$\begin{aligned} |\Phi_1\rangle + |\Phi_2\rangle &= |\Phi_2\rangle + |\Phi_1\rangle = |\Phi_1 + \Phi_2\rangle \\ \alpha|\Phi_1\rangle &= |\Phi_1\rangle\alpha = |\alpha\Phi_1\rangle \end{aligned}$$

- Es gelten Assoziativgesetze

$$\begin{aligned} |\Phi_1\rangle + (|\Phi_2\rangle + |\Phi_3\rangle) &= (|\Phi_1\rangle + |\Phi_2\rangle) + |\Phi_3\rangle \\ (\alpha\beta)|\Phi\rangle &= \alpha(\beta|\Phi\rangle) \end{aligned}$$

- Es existiert ein Nullelement  $|0\rangle$  bezüglich der Addition

$$|\Phi\rangle + |0\rangle = |\Phi\rangle \quad \forall |\Phi\rangle \in \mathcal{H}$$

insbesondere gilt

$$0|\Phi\rangle = |0\rangle \quad \forall |\Phi\rangle \in \mathcal{H}$$

- Zu jedem Element  $|\Phi\rangle$  aus dem Vektorraum existiert ein bezüglich der Addition inverses Element  $|\Phi\rangle$  mit:

$$|\Phi\rangle + |-\Phi\rangle = |\Phi\rangle - |\Phi\rangle = |0\rangle$$

- Es gelten Distributivgesetze

$$\begin{aligned} \alpha(|\Phi_1\rangle + |\Phi_2\rangle) &= \alpha|\Phi_1\rangle + \alpha|\Phi_2\rangle \\ (\alpha + \beta)|\Phi\rangle &= \alpha|\Phi\rangle + \beta|\Phi\rangle \end{aligned}$$

- Die Elemente des Vektorraumes  $|\Phi_1\rangle \dots |\Phi_n\rangle$  heißen *linear unabhängig*, falls aus der Relation

$$\sum_{i=1}^n \alpha_i |\Phi_i\rangle = |0\rangle$$

zwingend folgt, dass  $\alpha_i = 0$  für  $i = 1 \dots n$ . Die Maximalzahl linear unabhängiger Elemente in  $\mathcal{H}$  bezeichnet man als die *Dimension des Vektorraumes*  $\mathcal{H}$ .

### 6.1.2 $\mathcal{H}$ ein unitärer Vektorraum

Dies bedeutet, dass im Hilbertraum  $\mathcal{H}$  ein Skalarprodukt definiert ist, d.h.: jedem Paar von Elementen des Hilbertraums  $\mathcal{H}$ ,  $|\Phi_k\rangle$  und  $|\Phi_l\rangle$  ist eine komplexe Zahl zugeordnet, die wir  $\langle \Phi_k | \Phi_l \rangle$  oder abgekürzt mit  $\langle k | l \rangle$  bezeichnen wollen. Stellen wir die Zustände der Quantenmechanik in der Ortsdarstellung durch Wellenfunktionen dar, so ist dieses Skalarprodukt definiert durch:

$$\langle \Phi_k | \Phi_l \rangle := \int_{-\infty}^{\infty} dx \Phi_k^*(x) \Phi_l(x). \quad (6.3)$$

Dabei deutet \* die komplexe Konjugation der Wellenfunktion  $\Phi_l(x)$  an. Man kann sich nun leicht davon überzeugen, dass dieses Skalarprodukt die folgenden Eigenschaften besitzt:

$$\begin{aligned} \langle k|l \rangle &= \langle l|k \rangle^* \\ \langle k|l_1 + l_2 \rangle &= \langle k|l_1 \rangle + \langle k|l_2 \rangle \\ \langle k|\beta l \rangle &= \beta \langle k|l \rangle = \langle \beta^* k|l \rangle \text{ für } \beta \in \mathcal{C} \\ \langle k|k \rangle &\geq 0 \end{aligned} \quad (6.4)$$

In der letzten Zeile gilt  $\langle k|k \rangle = 0$  nur für den Nullvektor  $|k \rangle = |0 \rangle$ .

Als ein Beispiel für die Definition eines solchen Skalarproduktes betrachten wir die komplexwertige Wellenfunktion für die Ortsdarstellung der Quantenmechanik, wie wir sie im vorhergehenden Kapitel benutzt haben. Sind  $|i \rangle \leftrightarrow \varphi_i(x)$  und  $|j \rangle \leftrightarrow \varphi_j(x)$  zwei solche komplexwertige Wellenfunktionen, so definieren wir das Skalarprodukt durch

$$\langle i|j \rangle = \int_{-\infty}^{\infty} dx \varphi_i^*(x) \varphi_j(x) \quad (6.5)$$

Voraussetzung dafür dass dieses Skalarprodukt definiert ist, ist also nur, dass die Funktionen quadratintegrierbar sind. Man spricht in diesem Fall vom Hilbertraum der quadratintegrierbaren Funktionen in einer Raumdimension.

Um den Hintergrund für die Schreibweise des Skalarproduktes mit  $\langle k|l \rangle$  zu verdeutlichen, wollen wir das Skalarprodukt mit diesen Eigenschaften noch einmal aus einer etwas anderen Sicht einführen. Wir ordnen dazu jedem Element  $|k \rangle$  des Hilbertraumes  $\mathcal{H}$  einen Vektor  $\langle k|$  in einem dualen Hilbertraum  $\mathcal{H}^*$  zu. Das Skalarprodukt  $\langle l|k \rangle$  ist dann das Produkt des Vektors  $|k \rangle \in \mathcal{H}$  mit dem zu  $|l \rangle$  dualen Vektor  $\langle l| \in \mathcal{H}^*$ . Man spricht auch von einem **ket** Vektor  $|k \rangle$  und einem **bra** Vektor  $\langle l|$ . Diese Nomenklatur wurde von Dirac eingeführt (man spricht deshalb auch von der *Dirac Notation*). Man kann sich diese Notation leicht merken, weil die Klammer (englisch bracket) des Skalarproduktes in diese Silben zerfällt

$$\begin{array}{ccc} \langle l| & & |k \rangle \\ \text{bra} & -c- & \text{ket} \end{array}$$

Die Rechenregeln für die bra-Vektoren im dualen Raum  $\mathcal{H}^*$  entsprechen weitgehend den oben aufgeführten Regeln im Hilbertraum  $\mathcal{H}$  mit der Ausnahme der Produktbildung, für die gilt

$$\langle \alpha \Phi| = \alpha^* \langle \Phi|.$$

Mit dem Skalarprodukt können wir auch definieren, welche Vektoren zueinander **orthogonal** sind: Zwei Vektoren oder Elemente des Hilbertraumes heißen orthogonal, wenn gilt:

$$|k \rangle, |l \rangle \text{ orthogonal} \Leftrightarrow \langle k|l \rangle = 0 \quad (6.6)$$

Aus den Eigenschaften des Skalarproduktes können wir direkt die **Schwarzsche Ungleichung**

$$|\langle \Phi|\Psi \rangle|^2 \leq \langle \Phi|\Phi \rangle \langle \Psi|\Psi \rangle \quad (6.7)$$

herleiten. In dieser Ungleichung gilt das Gleichzeichen = genau dann, wenn die beiden Vektoren  $|\Phi \rangle$  und  $|\Psi \rangle$  "parallel" sind, also

$$|\Phi \rangle = \lambda |\Psi \rangle. \quad (6.8)$$

Zum Beweis der Schwarzschen Ungleichung betrachten wir den Vektor  $|\Psi\rangle$  und addieren dazu eine "Null", indem wir einen Term  $\lambda|\Phi\rangle$  mit  $\lambda = \langle \Phi|\Psi\rangle / \langle \Phi|\Phi\rangle$  addieren und subtrahieren:

$$\begin{aligned} |\Psi\rangle &= |\Phi\rangle \underbrace{\frac{\langle \Phi|\Psi\rangle}{\langle \Phi|\Phi\rangle}}_{=\lambda} + \underbrace{|\Psi\rangle - |\Phi\rangle \frac{\langle \Phi|\Psi\rangle}{\langle \Phi|\Phi\rangle}}_{=|\xi\rangle} \\ &= \lambda|\Phi\rangle + |\xi\rangle \end{aligned} \quad (6.9)$$

In der ersten Zeile dieser Gleichung haben wir den Vektor  $|\xi\rangle$  definiert. Bildet man das Skalarprodukt dieses Vektors  $|\xi\rangle$  mit dem Vektor  $|\Phi\rangle$ , so erhält man

$$\begin{aligned} \langle \Phi|\xi\rangle &= \langle \Phi|\Psi\rangle - \langle \Phi|\Phi\rangle \frac{\langle \Phi|\Psi\rangle}{\langle \Phi|\Phi\rangle} \\ &= 0 \end{aligned}$$

der Vektor  $|\xi\rangle$  ist also orthogonal zu  $|\Phi\rangle$  und (6.9) liefert die Zerlegung des Vektors  $|\Psi\rangle$  in einen Anteil parallel zu  $|\Phi\rangle$  und einer Komponente  $|\xi\rangle$  senkrecht oder orthogonal zu  $|\Phi\rangle$ . Wir bilden nun das Skalarprodukt von  $|\Psi\rangle$  mit sich selbst und betrachten dabei die Darstellung von  $|\Psi\rangle$  aus (6.9)

$$\begin{aligned} \langle \Psi|\Psi\rangle &= \langle \lambda\Phi + \xi|\lambda\Phi + \xi\rangle \\ &= \lambda\lambda^* \langle \Phi|\Phi\rangle + \langle \xi|\xi\rangle \\ &\geq \lambda\lambda^* \langle \Phi|\Phi\rangle \\ &= \frac{\langle \Phi|\Psi\rangle \langle \Psi|\Phi\rangle}{\langle \Phi|\Phi\rangle} \end{aligned} \quad (6.10)$$

Multipliziert man diese Ungleichung mit der positiven Zahl  $\langle \Phi|\Phi\rangle$ , so erhält man aus der ersten bzw. letzten Zeile von (6.10) die Schwarzsche Ungleichung (6.8). Beim Übergang von der zweiten zur dritten Zeile in (6.10) ergibt sich die Ungleichung dadurch, dass die positiv definite Zahl (siehe (6.4)  $\langle \xi|\xi\rangle$  weggelassen wurde. Aus dieser Ungleichung wird also eine Gleichung genau dann, wenn  $|\xi\rangle = 0$ , und das bedeutet ja nach (6.9), dass  $|\Psi\rangle = \lambda|\Phi\rangle$  ist. Damit ist also die Schwarzsche Ungleichung vollständig bewiesen.

Durch das Skalarprodukt wird schließlich auch die Länge, beziehungsweise die Norm eines Vektors definiert

$$\text{Norm des Vektors } |k\rangle: \quad ||k|| = \sqrt{\langle k|k\rangle} \quad (6.11)$$

Zu den Eigenschaften eines Hilbertraumes gehört außer den bisher aufgeführten Eigenschaften eines unitären Vektorraumes, dass er **separabel** und **vollständig** ist. In diesem Rahmen soll nicht näher auf diese Eigenschaften eingegangen werden. Für uns ist hier aber wichtig, dass aus diesen Eigenschaften folgt, dass der Hilbertraum eine Basis aus orthonormalen Basiszuständen besitzt und dass die Dimension des Hilbertraumes höchstens abzählbar unendlich ist.

Sei nun

$$\{|\Phi_i\rangle, i = 1 \cdots N\} \quad (6.12)$$

eine solche Orthonormalbasis unseres Hilbertraumes. Es gilt also für das Skalarprodukt der Basiselemente

$$\langle \Phi_i|\Phi_j\rangle = \delta_{i,j}. \quad (6.13)$$

Da (6.12) eine Basis definiert, können wir jedes beliebige Element  $|\Psi\rangle$  des Hilbertraumes als Linearkombination dieser Basisvektoren darstellen

$$|\Psi\rangle = \sum_{i=1}^N \alpha_i |\Phi_i\rangle \quad (6.14)$$

Die komplexwertigen  $\alpha_i$  bezeichnet man als **Entwicklungskoeffizienten** des Vektors  $|\Psi\rangle$  in der Basis der  $|\Phi_i\rangle$ . Da (6.12) eine Orthonormalbasis definiert, sind die Entwicklungskoeffizienten eindeutig bestimmt und lassen sich mit dem Skalarprodukt

$$\alpha_i = \langle \Phi_i | \Psi \rangle$$

berechnen. Wir können also (6.14) auch umschreiben in

$$|\Psi\rangle = \sum_{i=1}^N \underbrace{|\Phi_i\rangle \langle \Phi_i|}_{=1} |\Psi\rangle$$

Da diese Gleichung für jeden Zustand  $|\Psi\rangle$  gilt, wird durch

$$\sum_{i=1}^N |\Phi_i\rangle \langle \Phi_i| = 1 \quad (6.15)$$

eine Darstellung des Einsoperators definiert. Man spricht von dieser Gleichung auch als der **Vollständigkeitsrelation**, da sie ja auf der Vollständigkeit des Hilbertraumes beruht.

Hat man sich also auf eine Basis geeinigt, so ist jedes Element des Hilbertraumes eindeutig durch die Entwicklungskoeffizienten, beziehungsweise durch den Spaltenvektor gebildet aus diesen Koeffizienten definiert

$$|\Psi\rangle \iff \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_N \end{pmatrix} \quad (6.16)$$

Durch diese Darstellung der Elemente mit den Entwicklungskoeffizienten lassen sich viele Rechnungen einfach durchführen. Seien z.B. zwei Elemente durch die Entwicklungskoeffizienten definiert nach

$$\begin{aligned} |\Psi_1\rangle &= \sum_{i=1}^N |\Phi_i\rangle \alpha_i \\ |\Psi_2\rangle &= \sum_{j=1}^N |\Phi_j\rangle \beta_j \end{aligned}$$

Nach den Rechenregeln für das Skalarprodukt (6.4) gilt nun

$$\begin{aligned} \langle \Psi_2 | \Psi_1 \rangle &= \sum_{i,j=1}^N \beta_j^* \alpha_i \underbrace{\langle \Phi_j | \Phi_i \rangle}_{=\delta_{i,j}} \\ &= \sum_{i=1}^N \beta_i^* \alpha_i \\ &= (\beta_1^* \cdots \beta_N^*) \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_N \end{pmatrix} \end{aligned}$$

Das Skalarprodukt berechnet sich also als das Produkt des komplex konjugierten Zeilenvektors, gebildet aus den Entwicklungskoeffizienten des Bra Vektors, mit dem Spaltenvektor, gebildet aus den Koeffizienten des Ket Vektors.

Natürlich kann man auch verschiedene Orthonormalbasen  $|\Phi_i\rangle$  und  $|\varphi_k\rangle$  für den selben Hilbertraum betrachten. Wie alle anderen Elemente können auch die  $|\varphi_k\rangle$  in der Basis der  $|\Phi_i\rangle$  entwickelt werden

$$|\varphi_k\rangle = \sum_{i=1}^N U_{ik} |\Phi_i\rangle \quad \text{mit} \quad U_{ik} = \langle \Phi_i | \varphi_k \rangle$$

Wir können uns die  $N \times N$  Entwicklungskoeffizienten  $U_{ik}$  in einer Matrix angeordnet vorstellen, wobei der erste Index  $i$  die Zeile und der zweite Index  $k$  die Spalte bezeichnen soll. Der erste Spaltenvektor dieser Matrix enthält also gerade die Entwicklungskoeffizienten für das Element  $|\varphi_1\rangle$ . Die Matrix  $U_{ik}$  bezeichnet man als Transformationsmatrix von der Basis  $|\Phi_i\rangle$  auf die Basis  $|\varphi_k\rangle$ . Natürlich gibt es auch die Umkehrtransformation

$$|\Phi_i\rangle = \sum_{k=1}^N X_{ki} |\varphi_k\rangle \quad \text{mit} \quad X_{ki} = \langle \varphi_k | \Phi_i \rangle$$

Es gilt also

$$X_{ki} = \langle \Phi_i | \varphi_k \rangle^* = U_{ik}^*$$

oder ausgedrückt in der Matrixschreibweise mit Hilfe der Transposition

$$X = U^{*t}$$

Schließlich können wir aber auch beide Transformationen anwenden in der Form

$$\begin{aligned} |\varphi_k\rangle &= \sum_{i=1}^N U_{ik} \left( \underbrace{\sum_{m=1}^N X_{mi} |\varphi_m\rangle}_{=|\Phi_i\rangle} \right) \\ &= \sum_{m=1}^N \sum_{i=1}^N \underbrace{X_{mi} U_{ik}}_{=\delta_{m,k}} |\varphi_m\rangle \end{aligned}$$

Das Kronecker  $\delta_{m,k}$  in der letzten Zeile ergibt sich, da dies die einzige Möglichkeit zur Darstellung von  $|\varphi_k\rangle$  liefert. In der Matrixdarstellung erhalten wir also

$$\sum_{i=1}^N X_{mi} U_{ik} = (XU)_{mk} = \delta_{m,k}$$

das Produkt  $XU$  liefert also die Einheitsmatrix. Dies bedeutet, dass die Matrix  $X$  für die Umkehrtransformation zu  $U$ , die inverse Matrix zur Matrix  $U$  darstellt bezüglich der Matrixmultiplikation

$$\begin{aligned} X_{ik} &= U_{ik}^{-1} = U_{ki}^* \\ X &= U^{-1} = U^{*t} \end{aligned} \quad (6.17)$$

Diese Beziehung kennzeichnet eine **Unitäre Transformation**. Sind alle Entwicklungskoeffizienten, d.h. die Elemente von  $U$  reell, so gilt

$$U^{-1} = U^t \quad (6.18)$$

und man erhält eine **Orthogonale Transformation**.

### 6.1.3 Orts- und Impulsdarstellung

Wir werden uns nun davon überzeugen, dass auch die Orts- und Impulsdarstellung der Quantenmechanik, die wir im vorhergehenden Kapitel diskutiert haben, Spezialfälle dieser Darstellungen von Elementen eines Hilbertraumes sind. In der Ortsdarstellung ist der Zustand  $|\Psi\rangle$  charakterisiert durch die komplexen Zahlen, die dem Wert der Wellenfunktion an den verschiedenen Positionen  $x$  entspricht. Wir haben ja auch schon die entsprechende Bezeichnung für diese Werte der Wellenfunktionen eingeführt

$$\langle x|\Psi\rangle = \Psi(x)$$

Wir wollen uns nun davon überzeugen, dass auch diese Zahlen  $\langle x|\Psi\rangle$  als Ergebnis eines Skalarproduktes des Elementes  $|\Psi\rangle$  mit Basisfunktionen  $|x\rangle$  betrachtet werden können. Dazu schreiben wir

$$\begin{aligned}\langle x|\Psi\rangle &= \int_{-\infty}^{\infty} dy \delta(x-y)\Psi(y) \\ &= \int_{-\infty}^{\infty} dy \langle x|y\rangle \langle y|\Psi\rangle .\end{aligned}\quad (6.19)$$

In dieser Gleichung ergibt sich die Wellenfunktion  $\Psi(x)$  als Skalarprodukt des Bra-Zustandes  $|\Psi\rangle$  mit dem Ket-Zustand  $\langle x|$ . Das Skalarprodukt ist dabei, wie ja in der Ortsdarstellung üblich durch Integration des Produktes der entsprechenden Wellenfunktionen über die Ortsvariable  $y$  ausgeführt. Die Dirac'sche  $\delta$  Funktion,  $\delta(x-y) = \langle x|y\rangle$ , ist die Ortsdarstellung für den Zustand  $\langle x|$ , der ein Teilchen lokalisiert am Punkte  $x$  beschreibt. Die Menge der Basiszustände  $|y\rangle$  bildet eine Orthonormalbasis mit einem kontinuierlichem Parameter  $y$ , der den jeweiligen Basiszustand charakterisiert. Die Orthonormalität der Basis wird ausgedrückt durch

$$\begin{aligned}\langle y|z\rangle &= \int_{-\infty}^{\infty} dx \langle y|x\rangle \langle x|z\rangle \\ &= \int_{-\infty}^{\infty} dx \delta(y-x)\delta(x-z) \\ &= \delta(y-z)\end{aligned}\quad (6.20)$$

Während also die Orthogonalität für eine diskretisierte Basis, d.h. eine Basis deren Elemente durch eine diskrete Indexmenge gekennzeichnet sind, durch die Orthogonalitätsrelation (6.13) ausgedrückt wird, erhalten wir für die kontinuierliche Basis eine Relation (6.20), bei der das Kronecker- $\delta$  durch die Dirac'sche  $\delta$ -Funktion ersetzt wird. Aus der Entwicklung (6.19) sehen wir außerdem, dass gilt:

$$\int_{-\infty}^{\infty} dy |y\rangle \langle y| = 1 \quad (6.21)$$

die Vollständigkeitsrelation (6.15) in einer diskreten Basis wird also ersetzt durch eine Vollständigkeitsrelation in der kontinuierlichen Basis, bei der die Summe über alle Basiszustände durch ein entsprechendes Integral ersetzt wird.

Diese Vollständigkeitsrelation hilft uns auch bei der Interpretation der Impulsdarstellung

$$\begin{aligned}\langle p|\Psi\rangle &= \int_{-\infty}^{\infty} dx \langle p|x\rangle \langle x|\Psi\rangle \\ &= \int_{-\infty}^{\infty} dx \left( \frac{1}{\sqrt{2\pi}} e^{ipx/\hbar} \right)^* \Psi(x) .\end{aligned}$$



Bei dem Übergang zur zweiten Zeile haben wir ausgenutzt, dass die Ortsdarstellung der Impulseigenfunktionen durch die ebenen Wellen gegeben ist

$$\langle x|p \rangle = \frac{1}{\sqrt{2\pi}} e^{ipx/\hbar}$$

und entsprechend (6.4)  $\langle p|x \rangle = \langle x|p \rangle^*$  gilt. Damit sehen wir also, dass sich die Impulsdarstellung eines Zustandes  $\langle p|\Psi \rangle$  aus der Ortsdarstellung des gleichen Zustandes  $\langle x|\Psi \rangle$  gerade durch die Fouriertransformation

$$\langle p|\Psi \rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} dx e^{-ipx/\hbar} \langle x|\Psi \rangle$$

ergibt. Dies ist konsistent mit unseren Überlegungen im Abschnitt 2.1 (siehe z.B. 5.14).

### 6.1.4 Dynamische Variable

In der klassischen Mechanik bezeichnet eine Beobachtungsgröße (Observable) oder auch Dynamische Variable eine Funktion der generalisierten Koordinaten und Impulse des Systems. In unserem eindimensionalen Standardbeispiel wäre also eine dynamische Variable gegeben als eine Funktion  $F(p, x)$ . Beispiele für solche dynamischen Variablen sind etwa die Hamiltonfunktion  $H$  oder auch die kartesischen Komponenten des Drehimpulses eines Teilchens.

Bei dem Übergang zur Quantenmechanik werden die Koordinaten des Phasenraumes der Mechanik  $x$  und  $p$  ersetzt durch Operatoren  $\hat{x}$  und  $\hat{p}$  und die durch diese generalisierten Koordinaten dargestellten dynamischen Variablen werden ersetzt durch den Operator  $\hat{F}$ , der sich aus der entsprechenden Funktion der Basisoperatoren ergibt:

$$\begin{aligned} x &\Rightarrow \hat{x} \\ p &\Rightarrow \hat{p} \\ F(p, x) &\Rightarrow \hat{F} = F(\hat{p}, \hat{x}) \end{aligned}$$

Die Darstellung dieser Operatoren hängt natürlich von der Darstellung oder Basis ab, in der Hilbertraum der quantenmechanischen Zustände gegeben ist. So ergibt sich z.B. für die **Ortsdarstellung**:

$$\begin{aligned} x &\Rightarrow \hat{x} = \text{Multiplikation mit } x \\ p &\Rightarrow \hat{p} = \frac{\hbar}{i} \frac{d}{dx} \end{aligned} \quad (6.22)$$

In der **Impulsdarstellung** haben die gleichen Operatoren die Gestalt:

$$\begin{aligned} x &\Rightarrow \hat{x} = -\frac{\hbar}{i} \frac{d}{dp} \\ p &\Rightarrow \hat{p} = \text{Multiplikation mit } p \end{aligned} \quad (6.23)$$

Die Darstellung der Operatoren in der Ortsdarstellung nach (6.22) ist uns bereits aus den Diskussionen im vorhergehenden Kapitel (siehe etwa die Definition 5.21) bekannt.

Zur Verifikation des Ortsoperators  $\hat{x}$  in der Impulsdarstellung (6.23), berechnen wir den Erwartungswert

$$\begin{aligned}
\langle \Psi | \hat{x} | \Psi \rangle &= \int dp \langle \Psi | p \rangle \left( -\frac{\hbar}{i} \frac{d}{dp} \right) \langle p | \Psi \rangle \\
&= \int dp \int dx' \int dx \langle \Psi | x' \rangle \langle x' | p \rangle \left( -\frac{\hbar}{i} \frac{d}{dp} \right) \langle p | x \rangle \langle x | \Psi \rangle \\
&= \frac{1}{2\pi} \int dp \int dx' \int dx \langle \Psi | x' \rangle e^{ipx'/\hbar} \left( -\frac{\hbar}{i} \frac{d}{dp} \right) e^{-ipx/\hbar} \langle x | \Psi \rangle \\
&= \frac{1}{2\pi} \int dp \int dx' \int dx \Psi^*(x') e^{ip(x'-x)/\hbar} x \Psi(x) \\
&= \int dx' \int dx \Psi^*(x') \delta(x' - x) x \Psi(x) \\
&= \int dx \Psi^*(x) x \Psi(x)
\end{aligned}$$

Damit ist also der Ortsoperator  $\hat{x}$  in der Impulsdarstellung bestätigt, insbesondere auch das Minus-Zeichen in (6.23).

Als ein Beispiel für eine dynamische Variable in verschiedenen Darstellungen wollen wir den Hamiltonoperator für den eindimensionalen harmonischen Oszillator betrachten. Dieser hat die Form (siehe Abschnitt 2.5)

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2} m \omega^2 \hat{x}^2$$

In der Ortsdarstellung stellt man die Operatoren  $\hat{p}$  und  $\hat{x}$  nach (6.22) dar und wendet den resultierenden Hamiltonoperator auf eine Funktion  $\Psi(x)$  an

$$\hat{H}\Psi(x) = \left( -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2} m \omega^2 x^2 \right) \Psi(x) = f(x) \quad (6.24)$$

und erhält als Ergebnis dieser Operation eine Funktion  $f(x)$ . Entsprechend werden in der Impulsdarstellung die Operatoren  $\hat{p}$  und  $\hat{x}$  nach (6.23) eingesetzt und der resultierende Operator auf eine Funktion  $\tilde{\Psi}(p)$  der Impulsvariable  $p$  angewandt

$$\hat{H}\tilde{\Psi}(p) = \left( \frac{p^2}{2m} - \frac{1}{2} m \omega^2 \hbar^2 \frac{d^2}{dp^2} \right) \tilde{\Psi}(p) = \tilde{f}(p) \quad (6.25)$$

Unabhängig von der gewählten Darstellung können wir (6.24) und (6.25) schreiben als

$$\hat{H}|\Psi\rangle = |f\rangle$$

Der Hamiltonoperator  $\hat{H}$ , oder auch allgemein jeder Operator  $\hat{F}$ , der einer dynamischen Variablen zugeordnet ist, ist im mathematischen Sinne eine Abbildung vom Vektorraum der Ket-Zustände in den Vektorraum der Ket Zustände. Jedem Element  $|\Psi\rangle$  dieses Vektorraumes wird ein Bildelement  $|f\rangle$  zugeordnet.

Der Erwartungswert des Operators  $\hat{H}$  für den Zustand  $|\Psi\rangle$  ist dann

$$\langle \Psi | \hat{H} | \Psi \rangle = \langle \Psi | f \rangle$$

die Zahl, die sich als Skalarprodukt des Vektors  $\langle \Psi |$  mit diesem Bildvektor  $|f \rangle$  ergibt. Im folgenden sollen die Eigenschaften der Operatoren diskutiert werden, die den dynamischen Variablen zuzuordnen sind. Zunächst gilt, dass es sich bei den Operatoren, die in der Quantenmechanik dynamischen Variablen oder Observablen zugeordnet sind, um **Lineare Operatoren** handelt. Ein Operator  $\hat{A}$  ist genau dann linear, wenn man für jeden Zustand  $|\Psi \rangle$ , der sich als Linearkombination von Zuständen  $|\psi_i \rangle$  darstellt, schreiben kann:

$$\begin{aligned}\hat{A}|\Psi \rangle &= \hat{A}[\alpha_1|\psi_1 \rangle + \alpha_2|\psi_2 \rangle] \\ &= \alpha_1\hat{A}|\psi_1 \rangle + \alpha_2\hat{A}|\psi_2 \rangle\end{aligned}\quad (6.26)$$

wobei es sich bei den Koeffizienten  $\alpha_i$  um komplexe Zahlen handelt. Diese Eigenschaft ist in der Quantenmechanik, in der wir Zustände häufig als Überlagerung von Basiszuständen darstellen, von großer Bedeutung. Von der Gültigkeit dieser Forderung nach linearen Operatoren kann man sich leicht überzeugen: Die Basisoperatoren  $\hat{p}$  und  $\hat{x}$  sind lineare Operatoren (in der Orts- wie in der Impuls- oder jeder anderen Darstellung) und ebenso jede Summe oder jedes Produkt dieser Operatoren, wie sie bei dynamischen Variablen auftreten können.

Zur Charakterisierung einer zweiten Eigenschaft dieser Operatoren müssen wir zunächst den Begriff **Adjungierten Operators** einführen.

**Definition:** Zu einem Operator  $\hat{A}$  definiert man einen adjungierten Operator  $\hat{A}^\dagger$  so dass für alle Zustände  $\Phi$  und  $\Psi$  des betrachteten Hilbertraumes gilt:

$$\langle \Phi | \hat{A} | \Psi \rangle = \langle \hat{A}^\dagger \Phi | \Psi \rangle \quad (6.27)$$

Etwas ausführlicher kann man diese Definition auch wie folgt darstellen. Seien  $|f \rangle$  und  $|g \rangle$  die Bildvektoren für

$$\begin{aligned}\hat{A}|\Psi \rangle &= |f \rangle \\ \hat{A}^\dagger|\Phi \rangle &= |g \rangle\end{aligned}$$

so gilt:

$$\langle \Phi | \hat{A} | \Psi \rangle = \langle \Phi | f \rangle = \langle g | \Psi \rangle = \langle \hat{A}^\dagger \Phi | \Psi \rangle$$

Für Operatoren und die dazu adjungierten Operatoren gelten folgende Eigenschaften:

•

$$\langle \Phi | \hat{A} | \Psi \rangle = \langle \Psi | \hat{A}^\dagger | \Phi \rangle^* \quad (6.28)$$

Zum Beweis führen wir aus:

$$\begin{aligned}\langle \Phi | \hat{A} | \Psi \rangle &= \langle \hat{A}^\dagger \Phi | \Psi \rangle = \langle g | \Psi \rangle \\ &\stackrel{\text{siehe (6.4)}}{=} \langle \Psi | g \rangle^* = \langle \Psi | \hat{A}^\dagger | \Phi \rangle^*\end{aligned}$$

•

$$(\hat{A}^\dagger)^\dagger = \hat{A} \quad (6.29)$$

Den Beweis dieser Identität erhält man leicht durch zweifaches Anwenden der Beziehung (6.28):

$$\begin{aligned} \langle \Phi | \hat{A} | \Psi \rangle &= \langle \Psi | \underbrace{\hat{A}^\dagger}_{=\hat{B}} | \Phi \rangle^* \\ &= (\langle \Phi | \hat{B}^\dagger | \Psi \rangle)^* \\ &= \langle \Phi | (\hat{A}^\dagger)^\dagger | \Psi \rangle \end{aligned}$$

• Außerdem gilt auch

$$(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger. \quad (6.30)$$

Zum Beweis benutzen wir zunächst einmal (6.28)

$$\langle \Phi | (\hat{A}\hat{B}) | \Psi \rangle = \langle \Psi | (\hat{A}\hat{B})^\dagger | \Phi \rangle^* \quad (6.31)$$

Die gleiche Zahl können wir aber nach der Definition des adjungierten Operators auch umschreiben in

$$\begin{aligned} \langle \Phi | (\hat{A}\hat{B}) | \Psi \rangle &= \langle \hat{A}^\dagger \Phi | \hat{B} | \Psi \rangle \\ &= \langle \hat{B}^\dagger \hat{A}^\dagger \Phi | \Psi \rangle \\ &\stackrel{\text{siehe (6.28)}}{=} \langle \Psi | \hat{B}^\dagger \hat{A}^\dagger | \Phi \rangle \end{aligned}$$

Da diese Beziehungen für beliebige  $\Psi$  und  $\Phi$  gelten folgt aus dem Vergleich der letzten Zeile mit (6.31) die Behauptung (6.30).

Nach Einführung des Begriffes adjungierter Operator läßt sich direkt definieren, was unter einem **Hermiteischen Operator** zu verstehen ist:

**Definition:** Ein Operator  $\hat{A}$  heißt hermitesch oder auch selbstadjungiert wenn gilt:

$$\hat{A}^\dagger = \hat{A} \quad (6.32)$$

Wir können uns direkt davon überzeugen, dass ein Operator genau dann hermitesch ist, wenn die Erwartungswerte dieses Operators, berechnet für einen beliebigen Zustand  $\Psi$  reell sind. Sei  $\hat{A}$  hermitesch, so betrachten wir

$$\langle \Psi | \hat{A} | \Psi \rangle \stackrel{(6.28)}{=} \langle \Psi | \hat{A}^\dagger | \Psi \rangle^* \stackrel{A^\dagger=A}{=} \langle \Psi | \hat{A} | \Psi \rangle^* . \quad (6.33)$$

Die Erwartungswerte sind also genau dann reell, wenn  $\hat{A}$  selbstadjungiert bzw. hermitesch ist. Die Erwartungswerte entsprechen den Meßwerten für die physikalischen Observablen.

Da nur reellwertige Größen gemessen werden, ist es klar, dass wir für die Operatoren der Quantenmechanik, die physikalischen Observablen entsprechen, fordern, dass sie hermitesch sind.

Zum Schluß dieses Abschnittes soll der Begriff hermitescher Operator an zwei Beispielen verdeutlicht werden.

1. Der Impulsoperator ist in der Ortsdarstellung, und damit natürlich auch in jeder anderen Darstellung hermitesch. Seien  $\Phi$  und  $\Psi$  beliebige Zustände, die in der Ortsdarstellung durch die entsprechenden Funktion  $\Phi(x)$  und  $\Psi(x)$  dargestellt sind. Das Matrixelement für den Impulsoperator  $\hat{p}$  berechnet sich dann:

$$\begin{aligned} \langle \Phi | \hat{p} | \Psi \rangle &= \int_{-\infty}^{\infty} dx \Phi^*(x) \frac{\hbar}{i} \frac{d}{dx} \Psi(x) \\ &= \frac{\hbar}{i} \Phi^*(x) \Psi(x) \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} dx \left( \frac{\hbar}{i} \frac{d}{dx} \Phi^*(x) \right) \Psi(x) \\ &= \int_{-\infty}^{\infty} dx \left( \frac{\hbar}{i} \frac{d}{dx} \Phi(x) \right)^* \Psi(x) \\ &= \langle \hat{p} \Phi | \Psi \rangle \end{aligned}$$

Bei dem Übergang zur zweiten Zeile wurde partiell integriert und der erste Term in der zweiten Zeile ergibt Null, da die Wellenfunktion für  $x \rightarrow \pm\infty$  verschwinden. Vergleicht man diese Beziehung mit der Definition des adjungierten Operators (6.27) so stellt man fest, dass  $\hat{p} = \hat{p}^\dagger$ , d.h.  $\hat{p}$  hermitesch ist.

2. Natürlich ist auch der Ortsoperator  $\hat{x}$  hermitesch. Das Produkt aus diesen beiden Operatoren ist jedoch nicht hermitesch. Dazu betrachten wir

$$(\hat{p}\hat{x})^\dagger \stackrel{(6.30)}{=} \hat{x}^\dagger \hat{p}^\dagger \stackrel{(6.32)}{=} \hat{x}\hat{p}$$

Wäre das Produkt  $\hat{p}\hat{x}$  selbstadjungiert, so müßte also gelten  $\hat{x}\hat{p} = (\hat{p}\hat{x})^\dagger = \hat{p}\hat{x}$ . Wir können uns aber davon überzeugen, dass dies nicht gilt, indem wir in der Ortsdarstellung  $\hat{p}\hat{x}$  auf einen Zustand  $\Psi$  anwenden

$$\hat{p}\hat{x}\Psi(x) = \frac{\hbar}{i} \frac{d}{dx} (x\Psi(x)) = \frac{\hbar}{i} \Psi(x) + x \frac{\hbar}{i} \frac{d}{dx} \Psi(x) = \frac{\hbar}{i} \Psi + \hat{x}\hat{p}\Psi. \quad (6.34)$$

## 6.2 Eigenzustände und Matrixdarstellung

In der Quantenmechanik sind die Eigenvektoren und Eigenwerte von linearen und hermiteschen Operatoren, durch die ja die physikalischen Observablen dargestellt werden, von besonderer Bedeutung. Deshalb soll an dieser Stelle die Definition dieser Begriffe Eigenvektoren oder auch Eigenzustände und Eigenwerte aus der Linearen Algebra wiederholt werden:

**Definition:** Sei  $\hat{A}$  ein lineare Abbildung, beziehungsweise ein linearer Operator, definiert für die Zustände eines Hilbertraumes. Gilt für einen Zustand dieses Hilbertraumes

$$\hat{A}|\Psi\rangle = a|\Psi\rangle \quad (6.35)$$

mit einer komplexen Zahl  $a$ , so heißt diese Zahl  $a$  ein **Eigenwert** des Operators  $\hat{A}$  und  $|\Psi\rangle$  ist ein **Eigenzustand** oder **Eigenvektor** des Operators zu diesem Eigenwert  $a$ . Im folgenden wird ein Eigenvektor auch häufig durch den zugehörigen Eigenwert gekennzeichnet, wie z.B.  $|a\rangle$  einen Eigenzustand zum Eigenwert  $a$  bezeichnen soll.

Besitzt nun ein Operator  $\hat{A}$  zwei Eigenvektoren,  $|\Psi_1\rangle$  und  $|\Psi_2\rangle$  zum gleichen Eigenwert  $a$ , so ist auch jede Linearkombination dieser Zustände,  $\alpha|\Psi_1\rangle + \beta|\Psi_2\rangle$ , ein Eigenvektor zum selben Eigenwert  $a$ . Da die  $|\Psi_i\rangle$  Eigenvektoren sind gilt

$$\hat{A}|\Psi_i\rangle = a|\Psi_i\rangle. \quad (6.36)$$

Damit kann man nun leicht verifizieren

$$\begin{aligned} \hat{A}[\alpha|\Psi_1\rangle + \beta|\Psi_2\rangle] &\stackrel{\substack{= \\ \hat{A} \text{ ist linear}}}{=} \alpha\hat{A}|\Psi_1\rangle + \beta\hat{A}|\Psi_2\rangle \\ &\stackrel{\substack{= \\ (6.36)}}{=} \alpha a|\Psi_1\rangle + \beta a|\Psi_2\rangle \\ &= a[\alpha|\Psi_1\rangle + \beta|\Psi_2\rangle] \end{aligned}$$

dass die Definitionsgleichung (6.35) auch für diese Linearkombination gilt. Man kann sich leicht davon überzeugen, dass die Eigenvektoren zu einem Eigenwert  $a$  einen Unterraum des betrachteten Hilbertraumes bilden.

Für die Eigenwerte und Eigenvektoren eines hermiteschen Operators gilt folgender Satz:

**Satz:** Sei  $\hat{A}$  ein linearer und hermitescher Operator. Die Eigenwerte eines solchen Operators sind reell. Besitzt  $\hat{A}$  zwei unterschiedliche Eigenwerte  $a$  und  $b$ , so sind die zugehörigen Eigenvektoren  $|a\rangle$  und  $|b\rangle$  zueinander orthogonal.

dass die Eigenwerte von hermiteschen Operatoren reell sind, wird direkt durch die Beziehung

$$\frac{\langle a|\hat{A}|a\rangle}{\langle a|a\rangle} = \frac{\langle a|a\rangle}{\langle a|a\rangle} = a$$

deutlich. Der Zähler des Quotienten auf der rechten Seite der Gleichung ist reell, da hermitesche Operatoren reelle Erwartungswerte besitzen (siehe (6.33)), ebenso der Nenner (Norm des Zustandes  $|a\rangle$ ) und damit auch der Quotient, d.h. der Eigenwert  $a$ . Zum Nachweis der Orthogonalität der Zustände  $|a\rangle$  betrachten wir

$$\langle b|\hat{A}|a\rangle = a \langle b|a\rangle \quad (6.37)$$

Da  $\hat{A}$  hermitesch ist ( $\hat{A} = \hat{A}^\dagger$ ) und wegen (6.28) kann man aber auch schreiben:

$$\begin{aligned} \langle b|\hat{A}|a\rangle &= \langle a|\hat{A}|b\rangle^* \\ &= b \langle a|b\rangle^* \\ &= b \langle b|a\rangle. \end{aligned} \quad (6.38)$$

Subtrahiert man Gleichung (6.38) von (6.37), so ergibt sich

$$0 = (a - b) \langle b|a\rangle$$

Da die Eigenwerte  $a$  und  $b$  verschieden sind, ist  $(a - b) \neq 0$  und es muß gelten

$$\langle b|a\rangle = 0, \quad (6.39)$$

womit der Beweis, dass  $|a\rangle$  und  $|b\rangle$  zueinander orthogonal sind, erbracht ist.

Dieses Ergebnis ist eine wichtige Grundlage für die folgende Aussage über eine mögliche Basis des uns interessierenden Hilbertraumes:

**Satz:** Aus den Eigenzuständen zu einem linearen hermiteschen Operator kann eine Orthonormalbasis des Hilbertraumes der Zustände gebildet werden, auf den dieser Operator wirkt.

Der Beweis dieses Satzes wird an dieser Stelle nicht vollständig durchgeführt. Dazu sei auf Lehrbücher zur Linearen Algebra verwiesen. Zur Beweisführung und zu Konsequenzen dieser Aussage sollen folgende Bemerkungen gemacht werden

- a) Eigenvektoren eines linearen hermiteschen Operators  $\hat{A}|e_i\rangle = a_i|e_i\rangle$  mit verschiedenen Eigenwerten  $a_i$  sind ja, wie wir gerade gezeigt haben, zueinander orthogonal. Diese Eigenzustände müssen also nur noch entsprechend normiert werden, damit sie die Voraussetzungen als Elemente einer Orthonormalbasis

$$\langle e_i|e_j\rangle = \delta_{ij}$$

erfüllen.

- b) Gibt es mehrere, linear unabhängige Eigenzustände  $|e_{ik}\rangle$  zum selben Eigenwert  $a_i$

$$\hat{A}|e_{ik}\rangle = a_i|e_{ik}\rangle \quad \text{für } k = 1, \dots, N$$

so sagt man, dass der Eigenwert  $a_i$   $N$ -fach entartet ist. In diesem Fall kann man durch geeignete Orthonormierungsverfahren (z.B. Schmidtsche Orthogonalisierung) eine Orthonormalbasis dieses  $N$ -dimensionalen Unterraumes der Eigenvektoren zum Eigenwert  $a_i$  bilden.

- c) Sei nun  $|e_i\rangle$  ein vollständiger Satz von zueinander orthonormierten Eigenzuständen (wir verzichten hier und im folgenden auf den zweiten Index bei entarteten Eigenwerten), dann sagt unser Satz aus, dass jeder Zustand  $|\Psi\rangle$  des betrachteten Hilbertraumes als Linearkombination der Basiszustände  $|e_i\rangle$  dargestellt werden kann

$$|\Psi\rangle = \sum_i c_i |e_i\rangle .$$

Die Entwicklungskoeffizienten  $c_i$  sind dabei eindeutig als Ergebnis des Skalarprodukts (siehe (6.14))

$$c_i = \langle e_i | \Psi \rangle$$

gegeben und die Vollständigkeitsrelation (6.15)

$$\sum_i |e_i\rangle \langle e_i| = 1$$

gilt für diese Orthonormalbasis der Eigenzustände.

- d) In der vollständigen Orthonormalbasis seiner Eigenzustände  $|e_i\rangle$  kann man einen linearen hermiteschen Operator in der sogenannten **Spektraldarstellung** darstellen als

$$\hat{A} = \sum_i a_i |e_i\rangle \langle e_i| , \quad (6.40)$$

mit  $a_i$  den entsprechenden Eigenwerten. Zum Beweis dieser Aussage fügen wir den Einheitsoperator der Vollständigkeitsrelation (6.15) geeignet ein

$$\begin{aligned} \hat{A}|\Psi\rangle &= \hat{A} \underbrace{1}_{1} |\Psi\rangle = \hat{A} \sum_i |e_i\rangle \langle e_i| \Psi\rangle \\ &= \sum_i a_i |e_i\rangle \langle e_i| \Psi\rangle \end{aligned}$$

Da diese Beziehung für jedes  $|\Psi\rangle$  gilt, folgt die Behauptung (6.40).

### Beispiel: Eigenzustände des Hamiltonoperators

Die stationäre Schrödinger Gleichung (5.40), die wir zum Beispiel im Fall des Harmonischen Oszillators (siehe Abschnitt 2.5) schreiben können

$$\hat{H}|\Phi_n\rangle = E_n|\Phi_n\rangle \quad \text{mit } E_n = \hbar\omega\left(n + \frac{1}{2}\right) \text{ für } n = 0, 1, 2, \dots ,$$

ist ein typisches Beispiel für ein Eigenwertproblem mit diskreten Energieeigenwerten  $E_n$  und den Eigenzuständen zum Hamiltonoperator des Harmonischen Oszillators  $|\Phi_n\rangle$ . Zu jedem Energieeigenwert gibt es im Fall des eindimensionalen Harmonischen Oszillators genau einen linear unabhängigen Eigenzustand: die Energieeigenwerte sind nicht entartet. In der Ortsdarstellung sind diese Eigenzustände gegeben durch die Wellenfunktionen (5.82)

$$\Phi_n(x) = \langle x | \Phi_n \rangle = \frac{1}{\sqrt{2^n n!} \sqrt{\pi} b} H_n(x/b) e^{-\frac{x^2}{2b^2}} .$$

Mit den allgemeinen Überlegungen zu den Eigenschaften von Eigenfunktionssystemen können wir am Beispiel der Eigenfunktionen des Harmonischen Oszillators schließen:



- Die Eigenfunktionen des linearen Hermiteschen Hamiltonoperators  $\hat{H}$  bilden ein Orthonormalsystem:

$$\begin{aligned} \langle \Phi_m | \Phi_n \rangle &= \frac{1}{\sqrt{2^{m+n} n! m! \pi b}} \int_{-\infty}^{\infty} dx H_n(x/b) H_m(x/b) e^{-\frac{x^2}{b^2}} \\ &= \delta_{nm} \end{aligned}$$

- Die Eigenfunktionen  $|\Phi_n\rangle$  bilden eine vollständige Basis des Hilbertraumes von Ein-  
teilchenwellenfunktionen für Teilchen, die ( in einer Raumrichtung oder Dimension )  
gebunden sind. Dies entspricht dem Hilbertraum der quadratintegrierbaren Funktionen  
in einer Variablen  $x$ . Das bedeutet, dass jede beliebige quadratintegrierbare Funktion  
in einer Variable  $\Psi(x)$  in dieser Basis dargestellt oder entwickelt werden kann:

$$\begin{aligned} |\Psi\rangle &= \sum_n c_n |\Phi_n\rangle = \sum_n \langle \Phi_n | \Psi \rangle |\Phi_n\rangle \text{ oder} \\ \Psi(x) &= \sum_n c_n \Phi_n(x) \end{aligned}$$

### 6.2.1 Matrixdarstellung

Eine lineare Abbildung, und damit auch die Operatoren für physikalische Observable in der Quantenmechanik können durch  $N \times N$  Matrizen dargestellt werden, wobei  $N$  die Dimension des Vektorraumes oder Hilbertraumes bezeichnet. Mit der "Bracket" Darstellung der Quantenmechanik kann man dies leicht durch die folgenden Überlegungen verifizieren:

Ein Operator  $\hat{A}$  ist eindeutig dadurch definiert, dass man zu jedem Ursprungsvektor  $|\Psi\rangle$  den zugehörigen Bildvektor

$$|\Phi\rangle = \hat{A} |\Psi\rangle \quad (6.41)$$

angibt. Bezeichnet  $|e_i\rangle$  eine Orthonormalbasis des relevanten Hilbertraumes, so können sowohl  $|\Phi\rangle$  wie auch  $|\Psi\rangle$  in dieser Basis dargestellt werden

$$\begin{aligned} |\Phi\rangle &= \sum_{i=1}^N |e_i\rangle \langle e_i | \Phi \rangle \\ |\Psi\rangle &= \sum_{j=1}^N |e_j\rangle \langle e_j | \Psi \rangle \end{aligned} \quad (6.42)$$

und die Zustände sind durch die Entwicklungskoeffizienten  $\langle e_i | \Phi \rangle$ , bzw.  $\langle e_j | \Psi \rangle$  eindeutig definiert. Zur Berechnung der Koeffizienten  $\langle e_i | \Phi \rangle$  bilden wir das Skalarprodukt "der Gleichung" (6.41) mit  $\langle e_i |$

$$\begin{aligned} \langle e_i | \Phi \rangle &= \langle e_i | \hat{A} \underbrace{|\Psi\rangle}_{1 = \sum_j |e_j\rangle \langle e_j|} \\ &= \sum_{j=1}^N \langle e_i | \hat{A} |e_j\rangle \langle e_j | \Psi \rangle . \end{aligned} \quad (6.43)$$

Bezeichnet man die Zahlen  $\langle e_i | \hat{A} |e_j\rangle$  mit  $A_{ij}$  und schreibt man die Entwicklungskoeffizienten in Form eines Spaltenvektors der Dimension  $N$ , so kann man Gl. (6.43) auch als

Produkt einer Matrix mit einem Spaltenvektor schreiben

$$\begin{pmatrix} \langle e_1 | \Phi \rangle \\ \vdots \\ \langle e_N | \Phi \rangle \end{pmatrix} = \begin{pmatrix} A_{11} & \dots & A_{1N} \\ \vdots & \ddots & \vdots \\ A_{N1} & \dots & A_{NN} \end{pmatrix} \begin{pmatrix} \langle e_1 | \Psi \rangle \\ \vdots \\ \langle e_N | \Psi \rangle \end{pmatrix} \quad (6.44)$$

Die Zahlen  $A_{ij} = \langle e_i | \hat{A} | e_j \rangle$  bezeichnen also Matrixelemente, wobei der erste Index  $i$  dem Zeilenindex und der zweite Index  $j$  dem Spaltenindex entspricht. Die Definitionsgleichung für den Operator  $\hat{A}$  (6.41) kann also für beliebige Vektoren  $|\Phi\rangle$  angegeben werden, wenn die  $N^2$  Matrixelemente bekannt sind. In dieser Matrixdarstellung der Quantenmechanik sind die Operatoren durch Matrizen und die Zustände durch Spaltenvektoren (gebildet aus den Entwicklungskoeffizienten  $\langle e_i | \Phi \rangle$ ) dargestellt.

Wählt man nun als Basis des Hilbertraumes an Stelle der beliebigen Basis  $|e_i\rangle$  die Basis der Eigenfunktionen zum Operator  $\hat{A}$

$$\hat{A}|\alpha_i\rangle = a_i|\alpha_i\rangle$$

so ergibt sich für die Matrixelemente in dieser Basis:

$$\langle \alpha_j | \hat{A} \alpha_i \rangle = a_i \langle \alpha_j | \alpha_i \rangle = a_i \delta_{ji}$$

Die Matrix zum Operator  $\hat{A}$  ist also in der Basis der Eigenzustände diagonal, wobei die Diagonalelemente durch die Eigenwerte des Operators  $\hat{A}$ , nämlich  $a_i$ , gegeben sind. Will man also die Eigenwerte und Eigenvektoren eines Operators  $\hat{A}$  finden, so kann man in der Matrixdarstellung diese Aufgabe durch folgendes Verfahren lösen:

**I):** Bestimme in einer bekannten Basis  $|e_i\rangle$  die  $N \times N$  Matrix des Operators  $\langle e_i | \hat{A} | e_j \rangle$

**II):** Finde die Eigenwerte dieser Matrix. Dies kann man z.B. dadurch erreichen, dass man die Nullstellen des charakteristischen Polynoms

$$P(x) = \text{Det}(A_{ij} - x\delta_{ij})$$

bestimmt.

**III):** Bezeichnet  $a_1$  eine solche Nullstelle, bzw. Eigenwert der Matrix  $A_{ij}$ , so kann man die Entwicklungskoeffizienten des zugehörigen Eigenvektors  $\langle e_i | \alpha_1 \rangle$  durch Lösung des Gleichungssystems

$$\begin{pmatrix} A_{11} & \dots & A_{1N} \\ \vdots & \ddots & \vdots \\ A_{N1} & \dots & A_{NN} \end{pmatrix} \begin{pmatrix} \langle e_1 | \alpha_1 \rangle \\ \vdots \\ \langle e_N | \alpha_1 \rangle \end{pmatrix} = a_1 \begin{pmatrix} \langle e_1 | \alpha_1 \rangle \\ \vdots \\ \langle e_N | \alpha_1 \rangle \end{pmatrix}$$

bestimmen. Entsprechend verfährt man für die anderen Eigenwerte  $a_i$ .

Kennt man die Matrix des Operators  $\hat{A}$ :  $A_{ij} = \langle e_i | \hat{A} | e_j \rangle$ , so kann man auch direkt die Matrixelemente angeben, die den zu  $\hat{A}$  adjungierten Operator  $\hat{A}^\dagger$  definieren

$$A_{ij}^\dagger = \langle e_i | \hat{A}^\dagger | e_j \rangle$$

angeben. Dazu verwenden wir die Eigenschaften des adjungierten Operators (6.28)

$$\begin{aligned} A_{ij}^\dagger &= \langle e_i | \hat{A}^\dagger | e_j \rangle \\ &= \langle \hat{A}^\dagger e_j | e_i \rangle^* \\ &= \langle e_j | \hat{A} | e_i \rangle^* \\ &= A_{ji}^* \end{aligned}$$

Man erhält also die Matrixelemente des adjungierten Operators  $\hat{A}$  aus der Matrix des Operators  $\hat{A}$  durch Transponieren: vertausche Zeilen- und Spaltenindex, und komplex konjugieren. In Kurzform

$$A^\dagger = (A^T)^* \quad (6.45)$$

Im Fall eines selbstadjungierten oder hermiteschen Operators  $\hat{A}^\dagger = \hat{A}$  gilt also

$$A_{ij}^\dagger = A_{ij} = A_{ji}^*. \quad (6.46)$$

Insbesondere gehören also Matrizen, mit reellwertigen Matrixelementen, die symmetrisch sind zu einem hermiteschen Operator.

### 6.3 Messung von Observablen

In diesem Abschnitt soll die mathematische Darstellung der Quantenmechanik der vorangehenden Paragraphen wieder in Bezug gesetzt werden zur der Beschreibung physikalischer Systeme. Als Beispiel soll wieder Bezug genommen werden zum Experiment des Doppelspalt, das wir in der Einleitung, im Kapitel 7 diskutiert haben: Bei dem Experiment, bei dem nicht festgestellt werden konnte, welchen Spalt das Probeteilchen durchlief, war die Wellenfunktion in der Ebene der Wand des Doppelspalt gegeben als kohärente Überlagerung der 2 Komponenten der Wellenfunktion: Teilchen im Spalt 1 und Teilchen im Spalt 2. Wurde eine Messung vorgenommen, oder war diese vom Experimentaufbau her möglich, so wurde diese Kohärenz zerstört und wir hatten entweder die Wellenfunktion des Teilchens im Spalt 1, d.h. die Ortseigenfunktion  $\delta(x - x_{\text{Spalt1}})$  oder die entsprechende Wellenfunktion am anderen Spalt.

Entsprechendes geschah auch in der Ebene, in der die Teilchendetektoren aufgestellt sind. Zwischen dem Doppelspalt und dieser Detektorebene propagieren die Teilchen so, wie es durch die Wellenfunktion beschrieben wird. In der Detektorebene beobachtet man einzelne Teilchen, die jeweils an einer bestimmten Position ankommen. Die Wellenfunktion *kollabiert* zu einer Deltafunktion, d.h. zu einer Eigenfunktion des Ortsoperators, also des Operators, dessen Observable gemessen wird. Wir können normalerweise nicht vorher-sagen, an welcher Position das Teilchen ankommen wird, also zu welcher der möglichen Ortseigenfunktionen die Wellenfunktion der Quantenmechanik kollabieren wird. Das Betragsquadrat der Wellenfunktion  $|\Psi(x_1)|^2$  gibt lediglich an, mit welcher Wahrscheinlichkeit die Ortseigenfunktion an der Stelle  $x_1$  angenommen wird. Würde die Detektorwand gar nicht, oder 10 Meter entfernt aufgestellt, so würde die Wellenfunktion an der ursprünglichen Position der Detektorwand nicht kollabieren. Die Bewegung der Teilchen würde an dieser Position durch die gesamte Wellenfunktion beschrieben.

Durch das Aufstellen des Teilchendetektors beeinflussen wir die Wellenfunktion des Teilchens:

- Durch Aufstellen eines Teilchendetektors an der Position  $x_1$  *zwingen* wir die Wellenfunktion des Teilchens entweder in die Ortseigenfunktion  $\delta(x - x_1)$  zu kollabieren, das Teilchen wird in diesem Fall als vollständiges Teilchen im Detektor registriert, oder vollständig am Detektor vorbeizulaufen.
- Wird eine ganze Detektorwand aufgestellt, wird also das Teilchen sicher in irgendeinem der Detektoren registriert, so muss die Wellenfunktion zu einer der möglichen Ortseigenfunktionen kollabieren.
- Das Betragsquadrat der Wellenfunktion  $|\Psi(x_1)|^2$  gibt an mit welcher Wahrscheinlichkeit das Teilchen an einem bestimmten Ort  $x_1$  aufgefangen wird.

Dieser Überlegung gilt natürlich nicht nur für die Messung der Position eines Teilchens. Gleiches gilt auch bei der Messung anderer Observabler. Betrachtet man z.B. den Fall eines Harmonischen Oszillators (siehe Paragraph 8.5). Die Wellenfunktion eines Harmonischer Oszillator, der sich mit der Zeit ändert, wird als Überlagerung von Lösungen der stationären Schrödinger Gleichung

$$|\Psi\rangle = \sum_n |\Phi_n\rangle \langle \Phi_n | \Psi \rangle = \sum_n c_n |\Phi_n\rangle$$

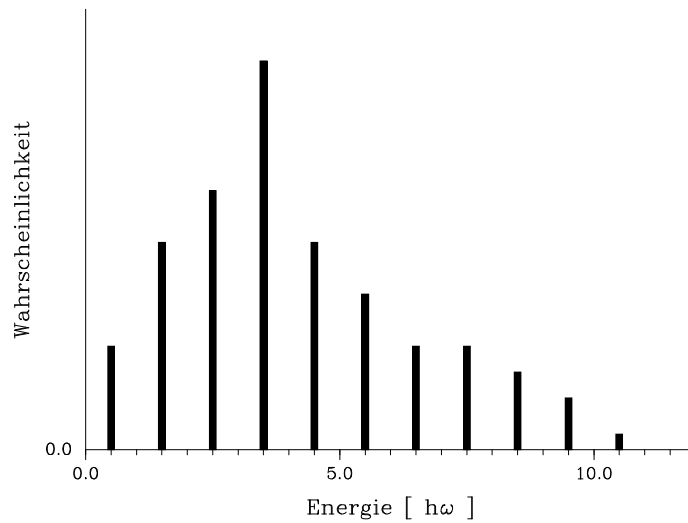


Abbildung 6.1: Ergebnis von Energiemessungen am Beispiel des Harmonischen Oszillators, siehe Gl. 6.47

Als Lösung der stationären Schrödinger Gleichung zum Harmonischen Oszillator sind die  $|\Phi_n\rangle$  Eigenzustände des Hamiltonoperators  $H$  mit den Energieeigenwerten  $E_n = \hbar\omega(n + 1/2)$ . Nimmt man nun an diesem System eine Energiemessung vor, so zwingt man den Zustand  $|\Psi\rangle$  in einen Energieeigenzustand zu kollabieren. Die Wahrscheinlichkeit eine bestimmte Energie  $E_n$  zu messen ist dann gegeben durch  $|c_n|^2$  beziehungsweise  $|\langle\Phi_n|\Psi\rangle|^2$ . Wiederholt man diese Energiemessung viele Male, so kann man insgesamt eine Wahrscheinlichkeitsverteilung erwarten von der Form

$$\rho(E) = \sum_n \delta(E - E_n) |\langle\Phi_n|\Psi\rangle|^2, \quad (6.47)$$

so wie sie z.B. in Abb. 6.1 dargestellt ist.

Als Physiker würde man diese Energieverteilung natürlich durch den Mittelwert, d.h. den Erwartungswert des Hamiltonoperators

$$\begin{aligned} \bar{E} &= \langle\Psi|H|\Psi\rangle \\ &= \sum_{i,n} \langle\Psi|\Phi_i\rangle \underbrace{\langle\Phi_i|H|\Phi_n\rangle}_{=E_n\delta_{in}} \langle\Phi_n|\Psi\rangle \\ &= \sum_n E_n |\langle\Phi_n|\Psi\rangle|^2 = \sum_n E_n c_n^2 \end{aligned}$$

charakterisieren und vielleicht auch noch das mittlere Schwankungsquadrat der Messung (siehe 5.9)

$$\begin{aligned} (\Delta E)^2 &= \langle\Psi|H^2|\Psi\rangle - \langle\Psi|H|\Psi\rangle^2 \\ &= \sum_n c_n^2 E_n^2 - \left( \sum_n c_n^2 E_n \right)^2 \end{aligned}$$

angeben.

Natürlich können wir diese Überlegungen verallgemeinern für die Messung einer beliebigen Observablen, der in der Quantenmechanik der hermitesche Operator  $A$  zugeordnet ist. Liefert nun in einer Messreihe dieser Observablen jede Einzelmessung den gleichen Wert  $a_n$ , einen Eigenwert des Operators  $A$ , so muß sich der Zustand des Systems in dem entsprechenden Eigenzustand befinden. Mathematisch formuliert gilt also:

$$(\Delta A)^2 = 0 \iff |\Psi\rangle = |\Phi_n\rangle, \quad (6.48)$$

wobei  $\Phi_n$  hier den Eigenzustand zum Operator  $A$  bezeichnen soll. Man spricht auch davon, dass in diesem Fall die Messung einen **scharfen** Messwert liefert, da die Verteilung der Abb. 6.1 entsprechend auf eine Delta Funktion reduziert wäre.

Zum Beweis der in 6.48 ausgedrückten Äquivalenz, zeigen wir zunächst die Richtung “ $\Leftarrow$ ”. Ist also  $|\Psi\rangle = |\Phi_n\rangle$ , so gilt

$$\begin{aligned} \langle \Psi | A^2 | \Psi \rangle &= \sum_i \underbrace{\langle \Phi_n | A | \Phi_i \rangle}_{=a_n \delta_{in}} \underbrace{\langle \Phi_i | A | \Phi_n \rangle}_{=a_n \delta_{in}} \\ &= a_n^2 \end{aligned} \quad (6.49)$$

Andererseits gilt  $\langle \Phi_n | A | \Phi_n \rangle = a_n$  und damit natürlich

$$(\Delta A)^2 = \langle \Psi | A^2 | \Psi \rangle - \langle \Psi | A | \Psi \rangle^2 = 0$$

Zum Beweis der Gegenrichtung “ $\Rightarrow$ ” überlegen wir uns, dass

$$(\Delta A)^2 = \sum_i c_i^2 a_i^2 - \left( \sum_i c_i^2 a_i \right)^2 = 0$$

(die  $c_i$  bezeichnen hier wieder die Entwicklungskoeffizienten  $\langle \Phi_i | \Psi \rangle$  nur dann erfüllt sein kann, wenn  $c_i = \delta_{in}$  ist, was aber bedeutet, dass  $\Psi = \Phi_n$  ist.

Für die weiteren Überlegungen wird der **Kommutator** von zwei Operatoren eine große Rolle. Wir definieren dazu:

*Gegeben seien 2 Operatoren  $\hat{A}$  und  $\hat{B}$ . Als Kommutator dieser beiden Operatoren bezeichnet man den Operator, der durch*

$$[\hat{A}, \hat{B}] := \hat{A}\hat{B} - \hat{B}\hat{A}, \quad (6.50)$$

*definiert ist.*

Diese Kommutatorrelation besitzt verschiedene Eigenschaften, von denen wir einige an dieser Stelle auflisten wollen. Für Operatoren  $\hat{A}$ ,  $\hat{B}$  und  $\hat{C}$  gelten insbesondere die folgenden Beziehungen

$$\begin{aligned} [\hat{A} + \hat{C}, \hat{B}] &= [\hat{A}, \hat{B}] + [\hat{C}, \hat{B}] \\ [\hat{B}, \hat{A}] &= -[\hat{A}, \hat{B}] \\ [\hat{A}, \hat{A}] &= 0. \end{aligned} \quad (6.51)$$

Der Beweis dieser Relationen ergibt sich sofort durch Einsetzen der Definition für den Kommutator (6.50).

Sind  $\hat{A}$  und  $\hat{B}$  2 hermitesche Operatoren, so ist der Kommutator dieser beiden Operatoren **antihermitesch**. Dies bedeutet, dass

$$\begin{aligned} [\hat{A}, \hat{B}] &= \hat{C} \quad \text{und} \\ ([\hat{A}, \hat{B}])^\dagger &= \hat{C}^\dagger = -\hat{C}. \end{aligned} \quad (6.52)$$

Zum Beweis dieser Eigenschaft berechnen wir unter Benutzung der in Abschnitt 6.1 bewiesenen Rechenregeln für adjungierte Operatoren

$$\begin{aligned} ([\hat{A}, \hat{B}])^\dagger &= (\hat{A}\hat{B})^\dagger - (\hat{B}\hat{A})^\dagger \\ &= \hat{B}^\dagger\hat{A}^\dagger - \hat{A}^\dagger\hat{B}^\dagger \\ &= \hat{B}\hat{A} - \hat{A}\hat{B} \\ &= -[\hat{A}, \hat{B}]. \end{aligned}$$

Beim Übergang von der zweiten zur dritten Zeile wurde ausgenutzt, dass  $\hat{A}$  und  $\hat{B}$  hermitesch sind. Multiplizieren wir also den Kommutator mit einer rein imaginären Zahl, also z.B.  $i$ , so ist dieser mit  $i$  multiplizierte Kommutator hermitesch, da ja bei der Adjungierung von Operatoren alle Faktoren komplex konjugiert werden. Anders formuliert können wir auch sagen, dass bei zwei hermiteschen Operatoren  $\hat{A}$  und  $\hat{B}$

$$[\hat{A}, \hat{B}] = i\hat{C}, \quad \hat{C} \text{ hermitesch ist.} \quad (6.53)$$

Mit dieser Definition und den zugehörigen Rechenregeln sind wir in der Lage eine wichtige Aussage über die gleichzeitige Messung von dynamischen Variablen eines Systems zu machen. Es gilt nämlich der folgende

**Satz:** Zwei dynamische Variable  $A$  und  $B$  eines Systems sind genau dann gleichzeitig scharf messbar, wenn die zugehörigen Operatoren  $\hat{A}$  und  $\hat{B}$  vertauschbar sind ( $\hat{A}\hat{B} = \hat{B}\hat{A}$ ), oder anders ausgedrückt, wenn der Kommutator dieser Operatoren  $[\hat{A}, \hat{B}]$  gleich Null ist.

Nach unseren Vorüberlegungen ist klar, dass gleichzeitig scharf messbar gleichbedeutend damit ist, dass es eine Basis des Hilbertraumes für das betrachtete System gibt, bei denen die Basiszustände gleichzeitig Eigenzustände des Operators  $\hat{A}$  und des Operators  $\hat{B}$  ist. Wir müssen also zeigen, dass eine solche Basis von gemeinsamen Eigenzuständen existiert genau dann, wenn die Operatoren vertauschen. Im ersten Schritt des Beweises nehmen wir an, dass eine Basis von gemeinsamen Eigenzuständen besteht in der Form

$$\begin{aligned} \hat{A}|a, b\rangle &= a|a, b\rangle \\ \hat{B}|a, b\rangle &= b|a, b\rangle, \end{aligned} \quad (6.54)$$

wobei  $a$  und  $b$  die zugehörigen Eigenwerte bezeichnen, also komplexe (bzw. da  $\hat{A}$  und  $\hat{B}$  hermitesche Operatoren sind, reelle) Zahlen. Wir werden nun zeigen, dass hieraus folgt

$$[\hat{A}, \hat{B}] = 0. \quad (6.55)$$

Wegen (6.54) und da  $\hat{A}$  und  $\hat{B}$  lineare Operatoren sind, gilt

$$\hat{A}\hat{B}|a, b\rangle = \hat{A}b|a, b\rangle = b\hat{A}|a, b\rangle = ab|a, b\rangle. \quad (6.56)$$

Ganz analog gilt aber auch

$$\hat{B}\hat{A}|a, b\rangle = \hat{B}a|a, b\rangle = a\hat{B}|a, b\rangle = ab|a, b\rangle. \quad (6.57)$$

Der Vergleich von (6.56) mit (6.57) zeigt also, dass für alle Basiszustände gilt:  $\hat{A}\hat{B}$  angewandt auf einen Basiszustand liefert das selbe Ergebnis wie  $\hat{B}\hat{A}$  angewandt auf diesen Zustand. Damit liefern diese beiden Operatoren für alle Zustände das selbe Ergebnis und es gilt  $\hat{A}\hat{B} = \hat{B}\hat{A}$ , was aber auch gleichbedeutend mit (6.55) ist.

Der Beweis in der Gegenrichtung (wir nehmen (6.55) an und zeigen dass daraus eine Basis von gemeinsamen Eigenzuständen folgt) ist etwas aufwändiger. Dazu betrachten wir eine Orthonormalbasis des Hilbertraumes, bestehend aus Eigenzuständen von  $\hat{A}$ , also

$$\hat{A}|a\rangle = a|a\rangle. \quad (6.58)$$

Kommutieren die Operatoren  $\hat{A}$  und  $\hat{B}$ , so gilt

$$\hat{A}(\hat{B}|a\rangle) = \hat{B}\hat{A}|a\rangle = a(\hat{B}|a\rangle),$$

was bedeutet, dass auch  $(\hat{B}|a\rangle)$  ein Eigenzustand des Operators  $\hat{A}$  ist mit dem selben Eigenwert  $a$ .

Für alle Eigenwerte  $a$ , die nicht entartet sind muss dann dieser Eigenzustand  $(\hat{B}|a\rangle)$  identisch sein mit  $|a\rangle$  bis auf eine komplexe Zahl  $b$ , es gilt also

$$\hat{B}|a\rangle = b|a\rangle,$$

was ja bedeutet, dass diese Basiszustände gleichzeitig Eigenzustände zu  $\hat{B}$  sind.

Ist der Eigenwert  $a$  des Operators  $n$ -fach entartet, so gibt es  $n$  orthogonale Eigenzustände zum Operator  $\hat{A}$  mit dem gleichen Eigenwert  $a$

$$\hat{A}|a, i\rangle = a|a, i\rangle \quad \text{für } i = 1 \dots n. \quad (6.59)$$

Diese  $n$  Zustände spannen den Unterraum von Eigenzuständen des Operators  $\hat{A}$  mit Eigenwert  $a$  auf. Da, wie oben gezeigt, auch  $\hat{B}|a\rangle$  zu diesem Unterraum gehört, gilt

$$\hat{B}|a, i\rangle = \sum_{j=1}^n |a, j\rangle \underbrace{\langle a, j|\hat{B}|a, i\rangle}_{=X_{ij}}. \quad (6.60)$$

Wir haben auf der rechten Seite dieser Gleichung praktisch eine 1 eingeschoben, wobei sich in diesem Fall die Summation auf Basiszustände des Unterraumes zum Eigenwert  $a$  beschränkt. Die  $n \times n$  Matrix  $X_{ij}$  ist die Darstellung des Operators  $\hat{B}$  in diesem Unterraum. Bestimmen wir die Eigenzustände dieser Matrix  $X_{ij}$ , so haben wir Eigenzustände zum Operator  $\hat{B}$ , die in diesem Unterraum liegen. Diese so bestimmten Eigenzustände sind also gleichzeitig Eigenzustände zu  $\hat{A}$  und  $\hat{B}$ . Damit haben wir also auch für die entarteten Eigenwerte den nötigen Beweis erbracht und den obigen Satz bewiesen.



Was passiert aber, wenn die Operatoren der Observablen  $A$  und  $B$  nicht kommutieren? Diese Frage bringt uns zur **Allgemeinen Heisenbergschen Unschärferelation**, die wir hier formulieren und beweisen wollen. Diese Unschärferelation besagt, dass für zwei hermitesche Operatoren, deren Kommutator durch (6.53) gegeben ist, für die mittleren Schwankungsquadrate gilt:

$$\langle \Psi | (\Delta \hat{A})^2 | \Psi \rangle \langle \Psi | (\Delta \hat{B})^2 | \Psi \rangle \geq \frac{1}{4} |\langle \Psi | \hat{C} | \Psi \rangle|^2. \quad (6.61)$$

Ist der Operator  $\hat{C}$  also ungleich der Null, so lassen sich  $\hat{A}$  und  $\hat{B}$  im Allgemeinen nicht gleichzeitig scharf messen und es gilt für das Produkt der mittleren Schwankungsquadrate der beiden Messungen die Beziehung (6.61).

Zum Beweis dieser Heisenbergschen Unschärferelation erinnern wir uns an die Schwarz'sche Ungleichung (6.7), die ja für beliebige Elemente des Hilbertraumes gilt

$$\langle f | f \rangle \langle g | g \rangle \geq |\langle f | g \rangle|^2, \quad (6.62)$$

und betrachten die Elemente

$$\begin{aligned} |f\rangle &:= (\hat{A} - \langle \hat{A} \rangle) |\Psi\rangle \\ |g\rangle &:= (\hat{B} - \langle \hat{B} \rangle) |\Psi\rangle, \end{aligned} \quad (6.63)$$

wobei  $\langle \hat{A} \rangle$  für den Erwartungswert von  $\hat{A}$  berechnet für den Zustand  $\Psi$  steht (entsprechendes gilt für  $\langle \hat{B} \rangle$ ). Damit berechnen wir z.B.

$$\begin{aligned} \langle f | f \rangle &= \langle (\hat{A} - \langle \hat{A} \rangle) \Psi | (\hat{A} - \langle \hat{A} \rangle) \Psi \rangle \\ &= \langle \Psi | (\hat{A} - \langle \hat{A} \rangle)^2 | \Psi \rangle \\ &= \langle \Psi | (\Delta \hat{A})^2 | \Psi \rangle. \end{aligned} \quad (6.64)$$

Bei dem Übergang zur zweiten Zeile wurde ausgenutzt, dass  $(\hat{A} - \langle \hat{A} \rangle)$  hermitesch ist, während bei dem Übergang zur letzten Zeile dieses Systems von Gleichungen die Definition des mittleren Schwankungsquadrats eingesetzt wurde. Entsprechend kann man auch zeigen, dass

$$\begin{aligned} \langle g | g \rangle &= \langle \Psi | (\Delta \hat{B})^2 | \Psi \rangle \quad \text{und} \\ \langle f | g \rangle &= \langle \Psi | (\hat{A} - \langle \hat{A} \rangle) (\hat{B} - \langle \hat{B} \rangle) | \Psi \rangle. \end{aligned} \quad (6.65)$$

Damit erhält die Schwarzsche Ungleichung aus (6.62) die Gestalt

$$\langle \Psi | (\Delta \hat{A})^2 | \Psi \rangle \langle \Psi | (\Delta \hat{B})^2 | \Psi \rangle \geq |\langle \Psi | (\hat{A} - \langle \hat{A} \rangle) (\hat{B} - \langle \hat{B} \rangle) | \Psi \rangle|^2. \quad (6.66)$$

Zum weiteren Beweis betrachten wir den Operator auf der rechten Seite dieser Ungleichung und schreiben ihn so um, dass wir ihn, jeweils multipliziert mit dem Faktor  $1/2$  einmal abschreiben und einmal ausmultiplizieren

$$\begin{aligned} (\hat{A} - \langle \hat{A} \rangle) (\hat{B} - \langle \hat{B} \rangle) &= \frac{1}{2} \{ (\hat{A} - \langle \hat{A} \rangle) (\hat{B} - \langle \hat{B} \rangle) + \hat{A} \hat{B} - \langle \hat{A} \rangle \hat{B} - \langle \hat{B} \rangle \hat{A} + \langle \hat{A} \rangle \langle \hat{B} \rangle \} \\ &= \frac{1}{2} \underbrace{ \{ (\hat{A} - \langle \hat{A} \rangle) (\hat{B} - \langle \hat{B} \rangle) \} }_{:= \hat{F}} + \frac{1}{2} [\hat{A}, \hat{B}] \\ &= \hat{F} + \frac{1}{2} i \hat{C}. \end{aligned} \quad (6.67)$$

Man kann sich leicht davon überzeugen, dass der Operator  $\hat{F}$  hermitesch ist. Das gleiche gilt, wie wir bereits oben gezeigt, haben für  $\hat{C}$ , der mit dem Kommutator über die Definition (6.61) verknüpft ist. Da  $\hat{F}$  und  $\hat{C}$  hermitesch sind, sind die entsprechenden Erwartungswerte reell und diese Darstellung des Operators in (6.67) liefert den Erwartungswert

$$\langle \Psi | \hat{F} + \frac{1}{2}i\hat{C} | \Psi \rangle = \langle \hat{F} \rangle + \frac{1}{2}i\langle \hat{C} \rangle$$

so dass die rechte Seite der Ungleichung (6.66) die Form annimmt

$$\begin{aligned} |\langle \hat{F} \rangle + \frac{1}{2}i\langle \hat{C} \rangle|^2 &= |\langle \hat{F} \rangle|^2 + \frac{1}{4}|\langle \hat{C} \rangle|^2 \\ &\geq \frac{1}{4}|\langle \hat{C} \rangle|^2. \end{aligned} \quad (6.68)$$

Die Ungleichung der letzten Zeile ergibt sich dadurch, dass der Summand  $|\langle \hat{F} \rangle|^2$  in der vorletzten Zeile positiv ist, durch sein Weglassen der Ausdruck als nur kleiner werden kann. Setzt man dies Ungleichung in (6.66) ein, so erhält man den Beweis für die Heisenbergsche Unschärferelation in (6.61).

Das wohl bekannteste Beispiel für die Anwendung der Heisenbergschen Unschärferelation ist natürlich die Unschärfe zwischen Orts- und Impulsbestimmung, die wir ja schon mehrfach angesprochen haben. Der Kommutator zwischen dem Impulsoperator  $\hat{p}$  und dem zugehörigen Ortsoperator  $\hat{x}$  berechnet sich z.B. in der Ortsdarstellung durch die Anwendung dieses Kommutators auf eine Wellenfunktion im Ortsraum

$$\begin{aligned} [\hat{p}, \hat{x}] \Psi(x) &= \hat{p}\hat{x}\Psi(x) - \hat{x}\hat{p}\Psi(x) \\ &= \frac{\hbar}{i} \frac{d}{dx} x \Psi(x) - \frac{\hbar}{i} x \frac{d}{dx} \Psi(x) \\ &= \frac{\hbar}{i} \Psi(x) + \frac{\hbar}{i} x \frac{d}{dx} \Psi(x) - \frac{\hbar}{i} x \frac{d}{dx} \Psi(x) \\ &= \frac{\hbar}{i} \Psi(x), \end{aligned}$$

oder in der Operatorschreibweise

$$[\hat{p}, \hat{x}] = \frac{\hbar}{i} = i(-\hbar) \quad (6.69)$$

In der Nomenklatur der Gleichung (6.53) ist also in diesem Fall der Operator  $\hat{C}$  gegeben durch die Zahl  $-\hbar$ . Damit ergibt sich also für das Produkt der Schwankungsquadrate in der Heisenbergschen Unschärferelation (6.61)

$$(\Delta p)^2 (\Delta x)^2 \geq \frac{\hbar^2}{4}. \quad (6.70)$$

Häufig wird allein diese Beziehung als Unschärferelation bezeichnet. Hier soll aber festgehalten bleiben, dass diese Beziehung nur ein Spezialfall der allgemeinen Relation (6.61) darstellt.

Zum Abschluss dieses Kapitels wollen wir uns noch mit dynamischen Variablen befassen, die in der klassischen Mechanik während der Entwicklung des Systems erhalten bleiben.

Ein Beispiel, das insbesondere für das folgende Kapitel eine große Bedeutung besitzt, ist der Drehimpuls eines Teilchens  $\vec{L} = \vec{r} \times \vec{p}$ , der bei einer Bewegung in einem Zentralfeld, erhalten bleibt. Die Komponenten dieses Vektors  $L_x, L_y$  und  $L_z$  sind Konstanten der Bewegung.

Sei also  $\hat{C}$  ein Operator, der in der Quantenmechanik eine solche Observable repräsentiert, die in der Klassischen Mechanik einer Konstanten der Bewegung entspricht. Wir werden im folgenden zeigen, dass in einem solchen Fall der Kommutator von  $\hat{C}$  mit dem Hamiltonoperator  $\hat{H}$  identisch 0 ist, also:

$$C = \text{Konst. der Bewegung} \quad \implies \quad [\hat{C}, \hat{H}] = 0, \quad (6.71)$$

sodass der Hamiltonoperator und  $\hat{C}$  ein gemeinsames Eigenfunktionssystem besitzen.

Wegen des Ehrenfestschen Theorems sollten bei einer Konstanten der Bewegung für alle Zustände  $|\Psi\rangle$  gelten, dass der Erwartungswert sich mit der Zeit nicht ändert, also

$$\begin{aligned} 0 &= \frac{d}{dt} \langle \Psi | \hat{C} | \Psi \rangle \\ &= \left\langle \frac{d}{dt} \Psi | \hat{C} | \Psi \right\rangle + \langle \Psi | \hat{C} \left| \frac{d}{dt} \Psi \right\rangle \end{aligned} \quad (6.72)$$

Wegen der Schrödingergleichung

$$\hat{H} |\Psi\rangle = i\hbar \frac{d}{dt} |\Psi\rangle$$

Gilt für den Bra-Zustand

$$\frac{d}{dt} \langle \Psi | = \frac{1}{i\hbar} \langle \hat{H} \Psi |. \quad (6.73)$$

Entsprechend gilt für den Ket Zustand

$$\langle \hat{H} \Psi | = \langle i\hbar \frac{d}{dt} \Psi | = -i\hbar \left\langle \frac{d}{dt} \Psi | \right.$$

bzw.

$$\left\langle \frac{d}{dt} \Psi | = \frac{-1}{i\hbar} \langle \hat{H} \Psi |. \quad (6.74)$$

Setzt man die Ergebnisse (6.73) und (6.74) in (6.72) ein, ergibt sich

$$\begin{aligned} 0 &= \frac{1}{i\hbar} \left\{ -\langle \hat{H} \Psi | \hat{C} | \Psi \rangle + \langle \Psi | \hat{C} \hat{H} | \Psi \rangle \right\} \\ &= \frac{1}{i\hbar} \left\{ -\langle \Psi | \hat{H} \hat{C} | \Psi \rangle + \langle \Psi | \hat{C} \hat{H} | \Psi \rangle \right\} \\ &= \frac{1}{i\hbar} \langle \Psi | [\hat{C}, \hat{H}] | \Psi \rangle. \end{aligned} \quad (6.75)$$

Dabei ergibt sich der Übergang zur zweiten Zeile aus der Tatsache, dass  $\hat{H}^\dagger = \hat{H}$ , da ja der Hamiltonoperator hermitesch ist. Da diese Gleichung für den Erwartungswert, berechnet mit allen möglichen Lösungen gilt, folgt

$$[\hat{C}, \hat{H}] = 0$$

was ja zu beweisen war.



# Kapitel 7

## Zentrale Kraftfelder

In diesem Kapitel werden die Werkzeuge der Quantenmechanik auf die Beschreibung der Bewegung von Massenpunkten in zentralen Kraftfeldern angewandt. Dazu werden die Drehimpulsoperatoren in der Ortsdarstellung explizit bestimmt (Abschnitt 1), die stationäre Schrödinger Gleichung auf eine Differenzialgleichung in einer Variablen reduziert (Abschnitt 2) und damit das Problem des Wasserstoffatoms bearbeitet

### 7.1 Drehimpuls in der Ortsdarstellung

Der Operator für den Drehimpuls eines Teilchens ergibt sich in natürlicher Weise aus der Definition des Drehimpulses in der klassischen Mechanik, indem wir einfach die Orts- und Impulsvariablen durch die entsprechenden Operatoren ersetzen. Für die Ortsdarstellung ergibt sich also

$$\hat{\vec{L}} = \hat{\vec{r}} \times \hat{\vec{p}} = \frac{\hbar}{i} \vec{r} \times \vec{\nabla}, \quad (7.1)$$

wobei  $\vec{\nabla}$  für den Operator des Gradienten steht. Stellt man den Ortsvektor  $\vec{r}$  und den Gradienten in der kartesischen Basis dar, so ergibt das Vektorprodukt die kartesische Darstellung der Drehimpulsoperatoren und liefert z.B. für den Operator

$$\hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x$$

Aber bereits in der klassischen Mechanik haben wir gelernt, dass die Kugelkoordinaten eine geeignetere Basis für die Behandlung von Zentralkraftproblemen darstellt (siehe z.B. Abschnitt 2.3 aus dem Skript zur Physik I). An dieser Stelle sollen einige Ergebnisse für diese Kugelkoordinaten rekapituliert werden.

Die Definition der Kugelkoordinaten  $r$ ,  $\vartheta$  und  $\varphi$  für einen Ortsvektor ist in Abb. 7.1 skizziert. Der Zusammenhang zwischen den kartesischen Koordinaten  $x$ ,  $y$  und  $z$  und den entsprechenden Kugelkoordinaten eines Vektors ist durch

$$\begin{aligned} x &= r \sin \vartheta \cos \varphi \\ y &= r \sin \vartheta \sin \varphi \\ z &= r \cos \vartheta, \end{aligned} \quad (7.2)$$

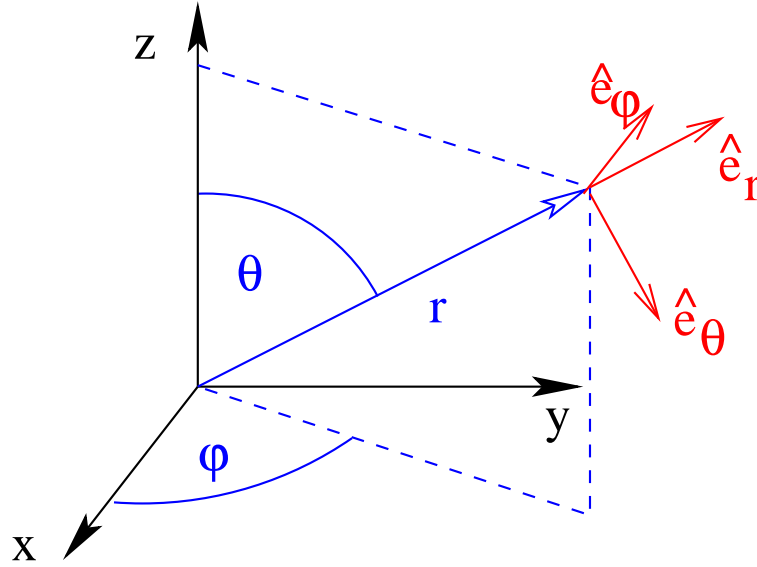


Abbildung 7.1: Vektor in Kugelkoordinaten

definiert. Daraus ergeben sich die Basisvektoren der Kugelkoordinaten

$$\hat{e}_r = \begin{pmatrix} \sin \vartheta \cos \varphi \\ \sin \vartheta \sin \varphi \\ \cos \vartheta \end{pmatrix}, \quad \hat{e}_\vartheta = \begin{pmatrix} \cos \vartheta \cos \varphi \\ \cos \vartheta \sin \varphi \\ -\sin \vartheta \end{pmatrix}, \quad \hat{e}_\varphi = \begin{pmatrix} -\sin \varphi \\ \cos \varphi \\ 0 \end{pmatrix}. \quad (7.3)$$

Man kann sich leicht davon überzeugen, dass diese Einheitsvektoren, die auch in der Abb. 7.1 dargestellt sind, ein Orthonormalsystem bilden. In diesen Einheitsvektoren stellt sich der Ortsvektor dar durch

$$\vec{r} = r \hat{e}_r, \quad (7.4)$$

und der Gradient Operator bekommt die Gestalt

$$\vec{\nabla} = \hat{e}_r \frac{d}{dr} + \hat{e}_\vartheta \frac{1}{r} \frac{d}{d\vartheta} + \hat{e}_\varphi \frac{1}{r \sin \vartheta} \frac{d}{d\varphi}. \quad (7.5)$$

Mit (7.1) erhält man also für den Operator des Bahndrehimpulses in Kugelkoordinaten

$$\begin{aligned} \hat{L} &= \frac{\hbar}{i} r \hat{e}_r \times \left\{ \hat{e}_r \frac{d}{dr} + \hat{e}_\vartheta \frac{1}{r} \frac{d}{d\vartheta} + \hat{e}_\varphi \frac{1}{r \sin \vartheta} \frac{d}{d\varphi} \right\} \\ &= \frac{\hbar}{i} \left\{ \underbrace{\hat{e}_r \times \hat{e}_r}_{=0} r \frac{d}{dr} + \underbrace{\hat{e}_r \times \hat{e}_\vartheta}_{=\hat{e}_\varphi} \frac{r}{r} \frac{d}{d\vartheta} + \underbrace{\hat{e}_r \times \hat{e}_\varphi}_{=-\hat{e}_\vartheta} \frac{r}{r \sin \vartheta} \frac{d}{d\varphi} \right\} \\ &= \frac{\hbar}{i} \left\{ \hat{e}_\varphi \frac{d}{d\vartheta} - \frac{1}{\sin \vartheta} \frac{d}{d\varphi} \right\}. \end{aligned} \quad (7.6)$$

Setzt man nun die Darstellung der Basisvektoren  $\hat{e}_\varphi$  und  $\hat{e}_\vartheta$  in der kartesischen Basis aus (7.3) ein, so ergibt sich

$$\hat{L} = \begin{pmatrix} \hat{L}_x \\ \hat{L}_y \\ \hat{L}_z \end{pmatrix} = \frac{\hbar}{i} \begin{pmatrix} -\sin \varphi \frac{d}{d\vartheta} & -\frac{\cos \vartheta}{\sin \vartheta} \cos \varphi \frac{d}{d\varphi} \\ \cos \varphi \frac{d}{d\vartheta} & -\frac{\cos \vartheta}{\sin \vartheta} \sin \varphi \frac{d}{d\varphi} \\ 0 & \frac{d}{d\varphi} \end{pmatrix}. \quad (7.7)$$

Aus dieser Darstellung können wir schon direkt ablesen, dass

$$\hat{L}_z = \frac{\hbar}{i} \frac{d}{d\varphi}. \quad (7.8)$$

Neben dem Operator  $\hat{L}_z$  interessiert uns besonders auch der Operator  $\hat{L}^2$ , den wir durch das Quadrat von (7.6) berechnen können

$$\hat{L}^2 = -\hbar^2 \left\{ \hat{e}_\varphi \frac{d}{d\vartheta} \hat{e}_\varphi \frac{d}{d\vartheta} - \hat{e}_\varphi \frac{d}{d\vartheta} \hat{e}_\vartheta \frac{1}{\sin \vartheta} \frac{d}{d\varphi} - \hat{e}_\varphi \frac{1}{\sin \vartheta} \frac{d}{d\varphi} \hat{e}_\vartheta \frac{d}{d\vartheta} + \hat{e}_\vartheta \frac{1}{\sin \vartheta} \frac{d}{d\varphi} \hat{e}_\vartheta \frac{1}{\sin \vartheta} \frac{d}{d\varphi} \right\}. \quad (7.9)$$

Für die weitere Berechnung müssen wir beachten, dass die Ableitungen auf alle Funktionen wirken, die jeweils rechts von der Ableitung stehen. Insbesondere muss also z.B. auch berücksichtigt werden, dass

$$\begin{aligned} \frac{d}{d\vartheta} \hat{e}_\vartheta &= \frac{d}{d\vartheta} \begin{pmatrix} \cos \vartheta \cos \varphi \\ \cos \vartheta \sin \varphi \\ -\sin \vartheta \end{pmatrix} = \begin{pmatrix} -\sin \vartheta \cos \varphi \\ -\sin \vartheta \sin \varphi \\ -\cos \vartheta \end{pmatrix} = -\hat{e}_r \\ \frac{d}{d\vartheta} \hat{e}_\varphi &= 0 \\ \frac{d}{d\varphi} \hat{e}_\varphi &= \begin{pmatrix} -\cos \varphi \\ -\sin \varphi \\ 0 \end{pmatrix} = -\{\cos \vartheta \hat{e}_\vartheta + \sin \vartheta \hat{e}_r\} \\ \frac{d}{d\varphi} \hat{e}_\vartheta &= \begin{pmatrix} -\cos \vartheta \sin \varphi \\ \cos \vartheta \cos \varphi \\ 0 \end{pmatrix} = \cos \vartheta \hat{e}_\varphi. \end{aligned} \quad (7.10)$$

Unter Ausnutzung dieser Beziehungen liefert (7.9) nach einigen Umrechnungen

$$\begin{aligned} \hat{L}^2 &= -\hbar^2 \left\{ \frac{d^2}{d\vartheta^2} + \frac{\cos \vartheta}{\sin \vartheta} \frac{d}{d\vartheta} + \frac{1}{\sin^2 \vartheta} \frac{d^2}{d\varphi^2} \right\} \\ &= -\hbar^2 \left\{ \frac{1}{\sin \vartheta} \frac{d}{d\vartheta} \sin \vartheta \frac{d}{d\vartheta} + \frac{1}{\sin^2 \vartheta} \frac{d^2}{d\varphi^2} \right\}. \end{aligned} \quad (7.11)$$

Man sieht also an den Darstellungen der Operatoren  $\hat{L}^2$  in (7.11) und  $\hat{L}_z$  in (7.8), dass diese ausschließlich auf die Winkelvariablen  $\vartheta$  und  $\varphi$  der Kugelkoordinaten wirken. Die gemeinsamen Eigenfunktionen zu diesen Operatoren, die es ja geben muss, da wir in den Übungen zeigen, dass der Kommutator

$$[\hat{L}^2, \hat{L}_z] = 0,$$

sind also als Funktionen dieser beiden Variablen auf der Oberfläche einer Kugel um den Koordinatenursprung definiert. Wir bezeichnen diese Eigenfunktionen als **Kugelflächenfunktionen** und stellen sie als Ortsdarstellung der Eigenzustände mit den Quantenzahlen  $l$  und  $m$  dar in der Form

$$\langle \vartheta, \varphi | lm \rangle = Y_{lm}(\vartheta, \varphi). \quad (7.12)$$

Im folgenden werden wir uns davon überzeugen, dass diese Eigenfunktionen den folgenden Eigenwertgleichungen genügen

$$\begin{aligned} \hat{L}^2 Y_{lm}(\vartheta, \varphi) &= \hbar^2 l(l+1) Y_{lm}(\vartheta, \varphi) \quad \text{für } l = 0, 1, 2, \dots \\ \hat{L}_z Y_{lm}(\vartheta, \varphi) &= \hbar m Y_{lm}(\vartheta, \varphi) \quad \text{für } m = -l, -l+1, -l+2, \dots, l-1, l. \end{aligned} \quad (7.13)$$

Ausserdem sind diese Kugelflächenfunktionen auch Eigenfunktionen zum Paritätsoperator. Der Paritätsoperator ist zunächst ein sehr abstraktes Konstrukt. Für eine beliebige Funktion  $f(\vec{r})$  bedeutet die Anwendung des Paritätsoperators  $\mathcal{P}$  auf diese Funktion, dass das Argument  $\vec{r}$  in den Ortsvektor  $-\vec{r}$  übergeführt wird

$$\mathcal{P}f(\vec{r}) = f(-\vec{r}), \quad (7.14)$$

durch Anwendung des Paritätsoperators wird also die Funktion am Koordinatenursprung "gespiegelt". Dynamische Variable oder auch die zugehörigen Operatoren der Quantenmechanik sind invariant unter der Paritätstransformation, wenn diese Variablen bei Anwendung des Paritätsoperators ihre Form nicht ändern. Eine Funktion  $f(\vec{r})$  ist also Eigenfunktion zum Paritätsoperator, wenn gilt

$$\mathcal{P}f(\vec{r}) = f(-\vec{r}) = pf(\vec{r}).$$

Bei wiederholter Anwendung des Paritätsoperators gilt

$$\mathcal{P}\mathcal{P}f(\vec{r}) = \mathcal{P}f(-\vec{r}) = f(\vec{r}),$$

was bedeutet, dass die möglichen Eigenwerte des Paritätsoperators  $\mathcal{P}$  durch  $p = 1$  und  $p = -1$  gegeben sind. In Kugelkoordinaten ausgedrückt bewirkt der Paritätsoperator

$$\mathcal{P}f(r, \vartheta, \varphi) = f(r, \pi - \vartheta, \varphi + \pi).$$

Man kann sich leicht davon überzeugen, dass dieser Paritätsoperator mit den Operatoren  $\hat{L}^2$  und  $\hat{L}_z$  vertauscht und es somit ein gemeinsames Eigenfunktionssystem gibt. Es gilt

$$\mathcal{P}Y_{lm}(\vartheta, \varphi) = (-1)^l Y_{lm}(\vartheta, \varphi). \quad (7.15)$$

Man sagt: die Kugelflächenfunktionen mit geradem Wert für  $l$  haben **positive Parität**, also den Paritätseigenwert  $+1$ , während die mit ungeradem  $l$  **negative Parität** besitzen.

In einem ersten Schritt betrachten wir die Eigenwertgleichung für  $\hat{L}_z$ . Mit (7.8) gilt:

$$\hat{L}_z Y_{lm}(\vartheta, \varphi) = \frac{\hbar}{i} \frac{d}{d\varphi} Y_{lm}(\vartheta, \varphi) = \hbar m Y_{lm}(\vartheta, \varphi). \quad (7.16)$$

Die zweite Gleichung stellt eine einfache Differenzialgleichung in der Winkelvariablen  $\varphi$  dar mit der Lösung ( $c_{lm}$  ist eine Konstante, die wir später bestimmen werden)

$$Y_{lm}(\vartheta, \varphi) = c_{lm} P_{lm}(\vartheta) e^{im\varphi}, \quad (7.17)$$

wobei  $m$  eine beliebige reelle Zahl sein kann. Die resultierende Funktion muss aber eindeutig definiert sein. Wenn die Winkelvariable  $\varphi$  um den Winkel  $2\pi$  erhöht wird, befindet man sich wieder auf der gleichen Stelle der Kugeloberfläche und es muss gelten

$$\begin{aligned} Y_{lm}(\vartheta, \varphi + 2\pi) &= c_{lm} P_{lm}(\vartheta) e^{im(\varphi + 2\pi)} \\ &= Y_{lm}(\vartheta, \varphi) = c_{lm} P_{lm}(\vartheta) e^{im\varphi}. \end{aligned}$$

Dies ist genau dann der Fall, wenn

$$e^{im2\pi} = 1,$$



was impliziert, dass  $m$  eine ganze Zahl sein muss.

In einem nächsten Schritt betrachten wir die Eigenwertgleichung für  $\hat{L}^2$  im Fall, dass  $m = 0$  ist. In diesem Fall hat die Differentialgleichung zur Bestimmung der Funktionen  $P_{l0}(\vartheta)$  die Form

$$\hat{L}^2 P_{l0}(\vartheta) = -\hbar^2 \left\{ \frac{1}{\sin \vartheta} \frac{d}{d\vartheta} \sin \vartheta \frac{d}{d\vartheta} \right\} P_{l0}(\vartheta) = \hbar^2 l(l+1) P_{l0}(\vartheta). \quad (7.18)$$

Zur weiteren Behandlung wird die Substitution

$$\vartheta \rightarrow \xi = \cos \vartheta,$$

benutzt, wodurch die Differentialgleichung (7.18) transformiert wird in

$$-\frac{d}{d\xi} (1 - \xi^2) \frac{d}{d\xi} P_{l0}(\xi) = l(l+1) P_{l0}(\xi). \quad (7.19)$$

Macht man dann für die Funktion  $P_{l0}(\xi)$  den Ansatz einer Reihe

$$P_{l0}(\xi) = \sum_k a_k \xi^k, \quad (7.20)$$

und setzt diesen Ansatz in (7.19) ein, so ergibt sich für die Koeffizienten

$$(k+2)(k+1)a_{k+2} - k(k+1)a_k = -l(l+1)a_k,$$

beziehungsweise die Rekursionsformel

$$a_{k+2} = \frac{k(k+1) - l(l+1)}{(k+2)(k+1)} a_k. \quad (7.21)$$

Diese Rekursionsformel bricht ab für  $k = l$ , sodass die Lösungsfunktionen  $P_{l0}$  (7.20) Polynome vom Grade  $l$  in der Variablen  $\xi = \cos \vartheta$  sind. Dabei treten bei geraden Zahlen  $l$  nur gerade Potenzen ( $k = 0, 2, \dots, l$ ) auf während bei ungeradem Wert für  $l$  nur ungerade Potenzen ( $k = 1, 3, \dots, l$ ) auftreten. Dadurch ist gewährleistet, dass die resultierenden Funktionen Eigenfunktionen zum Paritätsoperator sind, wie in (7.15) angegeben. Diese Funktionen  $P_{l0}(\cos \vartheta)$  heißen Legendrepolynome. Als Beispiele führen wir an

$$\begin{aligned} P_{00}(\cos \vartheta) &= 1, \\ P_{10}(\cos \vartheta) &= \cos \vartheta, \\ P_{20}(\cos \vartheta) &= \frac{3}{2} \cos^2 \vartheta - \frac{1}{2}, \\ P_{30}(\cos \vartheta) &= \frac{5}{2} \cos^3 \vartheta - \frac{3}{2} \cos \vartheta. \end{aligned} \quad (7.22)$$

Damit definieren wir die entsprechenden Kugelflächenfunktionen nach (7.17)

$$\begin{aligned} Y_{00}(\vartheta, \varphi) &= \frac{1}{\sqrt{4\pi}}, \\ Y_{10}(\vartheta, \varphi) &= \sqrt{\frac{3}{4\pi}} \cos \vartheta, \\ Y_{20}(\vartheta, \varphi) &= \sqrt{\frac{5}{4\pi}} \left( \frac{3}{2} \cos^2 \vartheta - \frac{1}{2} \right), \\ Y_{30}(\vartheta, \varphi) &= \sqrt{\frac{7}{4\pi}} \left( \frac{5}{2} \cos^3 \vartheta - \frac{3}{2} \cos \vartheta \right). \end{aligned} \quad (7.23)$$

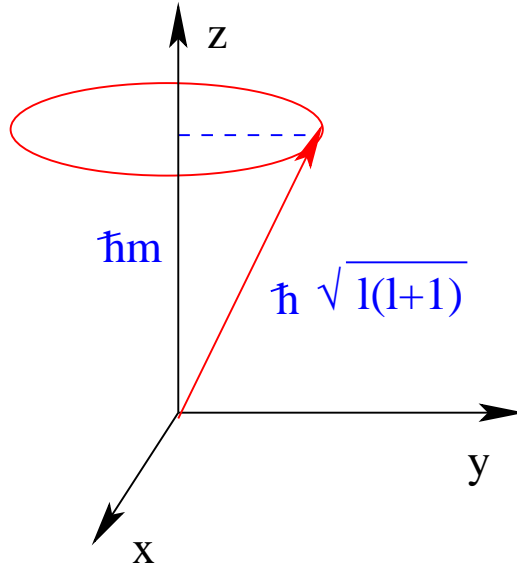


Abbildung 7.2: Schematische Darstellung des Drehimpulses charakterisiert durch die Quantenzahlen  $l$  und  $m$ . Siehe auch (7.26)

Die Koeffizienten  $c_{lm}$  sind dabei so definiert, dass die Kugelflächenfunktionen normiert sind gemäß

$$\begin{aligned}
 \langle l'm'|lm \rangle &= \underbrace{\int_0^{2\pi} d\varphi \int_0^\pi \sin\vartheta d\vartheta}_{=\int d\Omega} \langle l'm'|\vartheta, \varphi \rangle \langle \vartheta, \varphi|lm \rangle \\
 &= \int d\Omega Y_{l'm'}^*(\vartheta, \varphi) Y_{lm}(\vartheta, \varphi) \\
 &= \delta_{l'l} \delta_{m'm}.
 \end{aligned} \tag{7.24}$$

Dabei ist die Abhängigkeit der Kugelflächenfunktionen vom Winkel  $\vartheta$  für  $m \neq 0$  nicht durch die Legendre Polynome  $P_0$  sondern durch die sogenannten **assozierten Legendre Funktionen**  $P_{lm}$  definiert. Als Beispiele geben wir an

$$\begin{aligned}
 Y_{11}(\vartheta, \varphi) &= -\sqrt{\frac{3}{8\pi}} \sin\vartheta e^{i\varphi}, \\
 Y_{21}(\vartheta, \varphi) &= -\sqrt{\frac{15}{8\pi}} \sin\vartheta \cos\vartheta e^{i\varphi}, \\
 Y_{22}(\vartheta, \varphi) &= \sqrt{\frac{15}{32\pi}} \sin^2\vartheta e^{i\varphi}.
 \end{aligned} \tag{7.25}$$

Der Drehimpuls ist durch die Quantenzahlen  $l$  und  $m$  charakterisiert. Dabei steht  $l$  für den Betrag des Drehimpulsvektors und  $m$  für die  $z$ -Komponente  $L_z$ . Genauer gesagt gilt (siehe(7.13))

$$|\vec{L}| = \hbar\sqrt{l(l+1)} \quad \text{und} \quad L_z = \hbar m. \tag{7.26}$$

Nun kann natürlich der Betrag des Drehimpulsvektors nicht kleiner sein als der Betrag der  $z$ -Komponente, was aber bedeutet, dass

$$|m| \leq l,$$

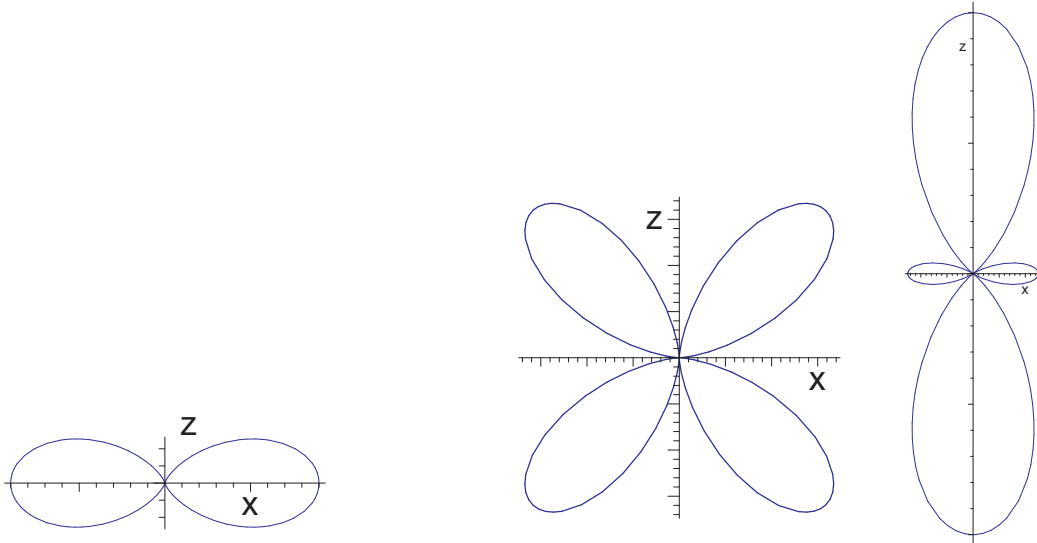


Abbildung 7.3: Darstellung der Dichte aus  $\rho_{lm}$  aus (7.27) für  $l = 2$  und von links nach rechts:  $m=0$ ,  $m=1$  und  $m=2$ .

wie es auch in den Auswahlregeln (7.13) zum Ausdruck kommt. Diese Charakteristika des Drehimpulsvektors sind auch in Abb. 7.2 dargestellt. In dieser Figur soll auch zum Ausdruck gebracht werden, dass über die Drehimpulskomponenten  $L_x$  und  $L_y$  keine Aussage gemacht werden kann, was bedeutet, dass der Endpunkt des Drehimpulsvektors auf dem Kreis liegt, der in dieser Darstellung angedeutet ist.

Ist also die  $z$ -Komponente maximal, wie z.B. im Fall  $l = m = 2$ , so ist der Drehimpulsvektor fast parallel zur  $z$ -Achse. Klassisch wird dies durch ein Teilchen realisiert, das sich auf einer Kreisbahn um die  $z$ -Achse in der  $x - y$  Ebene bewegt. Deshalb erwartet man also, dass auch in der Quantenmechanik die Aufenthaltswahrscheinlichkeit für ein Teilchen mit diesen Quantenzahlen in dieser Ebene maximal sein sollte. Diese Aufenthaltswahrscheinlichkeit

$$\rho_{lm}(\vartheta) = \alpha Y_{lm}^*(\vartheta, \varphi) Y_{lm}(\vartheta, \varphi) = \alpha c_{lm}^2 P_{lm}(\vartheta)^2, \tag{7.27}$$

ist unabhängig vom Azimutwinkel  $\varphi$  und in Abb. 7.3 dargestellt. Dabei steht der Abstand der dargestellten Kurve vom Koordinatenursprung für die Dichte in der entsprechenden Raumrichtung. Im Fall  $l = m = 2$  (linkes Teilbild von Abb. 7.3) ist in der Tat diese Dichte maximal für  $\vartheta = \pi/2$ . Man sieht auch, dass für  $l = 2, m = 1$  die Dichte maximal wird bei  $\vartheta \approx \pi/4$  und  $\vartheta \approx 3\pi/4$ . Auch das ist leicht verständlich, wenn man sich die Stellung des Drehimpulsvektors vergegenwärtigt und über alle möglichen Komponenten  $L_x$  und  $L_y$  mittelt. Im Fall  $m = 0$  liegt das Maximum auf der  $z$ -Achse.

Die Kugelflächenfunktionen sind nicht nur orthogonal, wie das ja bereits in (7.24) zum Ausdruck gebracht wurde, sie bilden auch ein vollständiges Funktionensystem für Funktionen, die auf einer Kugeloberfläche definiert sind. Ist  $\Upsilon(\vartheta, \varphi)$  eine solche Funktion, so würden wir sie in der Bracket Darstellung in der Form

$$\Upsilon(\vartheta, \varphi) = \langle \vartheta, \varphi | \Upsilon \rangle,$$

darstellen. Durch Einfügen einer Eins (siehe (6.15)) in der vollständigen Basis der Dre-

himpulseigenzustände ergibt sich

$$\begin{aligned}
 \Upsilon(\vartheta, \varphi) &= \langle \vartheta, \varphi | \Upsilon \rangle \\
 &= \sum_{l,m} \langle \vartheta, \varphi | lm \rangle \langle lm | \Upsilon \rangle \\
 &= \sum_{l,m} X_{lm} Y_{lm}(\vartheta, \varphi).
 \end{aligned} \tag{7.28}$$

Die Funktion  $\Upsilon$  wird nach den Kugelflächenfunktionen entwickelt. Die Entwicklungskoeffizienten  $X_{lm}$  berechnen sich, wie man ja aus (7.28) entnehmen kann durch das Skalarprodukt

$$X_{lm} = \langle lm | \Upsilon \rangle = \int_0^{2\pi} d\varphi \int_0^\pi \sin \vartheta d\vartheta Y_{lm}^*(\vartheta, \varphi) \Upsilon(\vartheta, \varphi). \tag{7.29}$$

Die vollständige Information über die Funktion  $\Upsilon(\vartheta, \varphi)$  ist in den Entwicklungskoeffizienten  $X_{lm}$  enthalten. Natürlich ist eine solche Entwicklung nur dann hilfreich, wenn die gesamte Funktion durch eine endliche Anzahl von Koeffizienten hinreichend charakterisiert ist und z.B. alle Koeffizienten mit  $l > l_0$  vernachlässigt werden können.

Als aktuelles Beispiel für eine solche Entwicklung nach Kugelflächenfunktionen sei die Darstellung der effektiven Temperatur der Hintergrundstrahlung, wie sie aus den verschiedenen Richtungen aus dem Weltall auf die Erde trifft, angeführt. Aus der Analyse der Entwicklungskoeffizienten dieser Temperaturverteilung werden zentrale Informationen für die Kosmologie extrahiert.

Dieser Abschnitt wird komplettiert durch das Protokoll einer Maple Sitzung. Das Worksheet dieser Sitzung kann von der Internet Adresse

<http://solid13.tphys.physik.uni-tuebingen.de/muether/physik3/physik3.html>

heruntergeladen werden

```
> with(orthopoly);
      [G, H, L, P, T, U]
```

Definition einer Prozedur zur Bestimmung der assoziierten Legendre Funktionen  $P_{lm}$

```
> Plm:=proc(theta,l,m);
> if type(l,integer)=true and type(m,integer)=true then
> x:=cos(theta);
> if m>0 then fak:=subs(y=x,diff(P(l,y),y$m));
> else fak:=subs(y=x,P(l,y));fi;
> (-1)^m*(sin(theta)^m)*fak;
> fi; end;
```

```

Plm := proc(theta, l, m)
local x, fak;
  if type(l, integer) = true and type(m, integer) = true then
    x := cos(theta);
    if 0 < m then fak := subs(y = x, diff(P(l, y), y $ m)) else fak := subs(y = x, P(l, y)) fi
    ;
    (-1)m × sin(theta)m × fak
  fi
end

```

Definition einer Prozedur für die Kugelflächenfunktionen (allerdings nicht normiert).

```

> Ylm:=proc(theta,phi,l,m);
> m1:=abs(m);
> if m1>l then RETURN(1,m); fi;
> exp(I*m*phi)*Plm(theta,l,m1); end;

```

```

Ylm := proc(theta, phi, l, m)
local m1;
  m1 := abs(m); if l < m1 then RETURN(l, m) fi; exp(I × m × phi) × Plm(theta, l, m1)
end

```

Beispiel für den Aufruf der Prozedur Ylm

```

> Ylm(theta,phi,2,1);
-3 e(I φ) sin(θ) cos(θ)

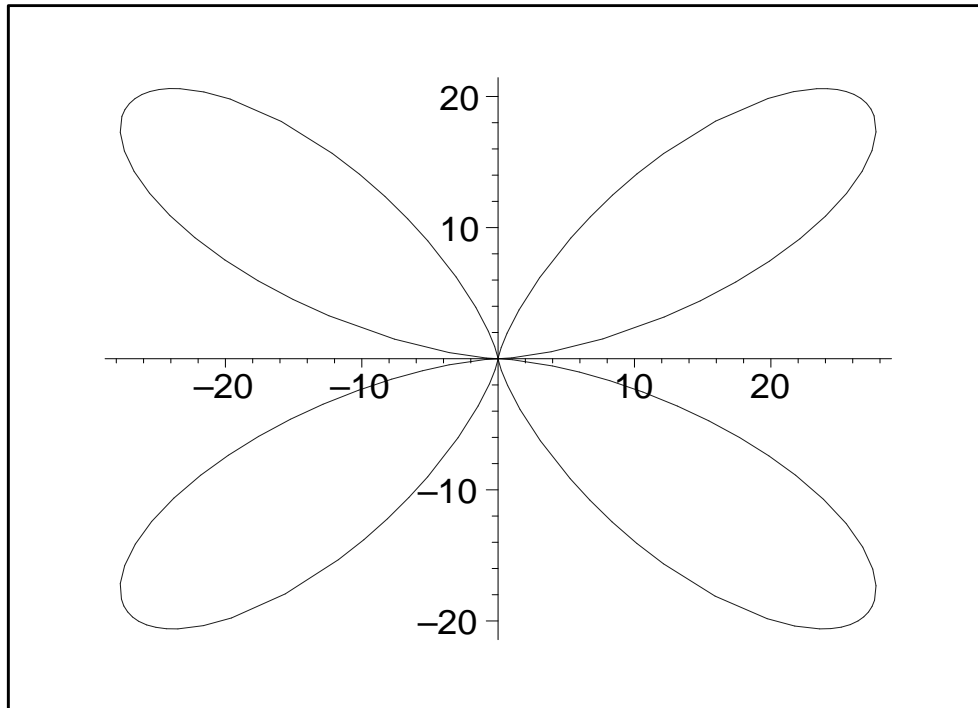
```

Berechnung und Darstellung der radialen Dichte nach (7.27). Die Prozedur xyexchange wird benötigt um die z-Achse in vertikaler Richtung darzustellen.

```

> rho:=proc(theta,m,n);
> Re(abs(evalf(Ylm(theta,0,m,n))))^2);
> end:
> with(plots):
> with(stats[statplots]):
> xyexchange(polarplot(rho(theta,3,2),theta=0..2*Pi,scaling='CONSTRAINED',color=blue));

```



```

> -(1/sin(theta)*diff(sin(theta)*diff(Ylm(theta,phi,2,1),theta),theta)+
> 1/sin(theta)^2*diff(Ylm(theta,phi,2,1),phi$2));

$$\frac{\cos(\theta) (-3e^{I\phi} \cos(\theta)^2 + 3e^{I\phi} \sin(\theta)^2) + 12 \sin(\theta)^2 e^{I\phi} \cos(\theta)}{\sin(\theta)} - 3 \frac{e^{I\phi} \cos(\theta)}{\sin(\theta)}$$

> simplify(%);

$$-18 e^{I\phi} \sin(\theta) \cos(\theta)$$

> Ylm(theta,phi,2,1);

$$-3 e^{I\phi} \sin(\theta) \cos(\theta)$$


```

## 7.2 Zentralfeldprobleme in der Quantenmechanik

Die Bewegung eines Teilchens in einem zentralen Kraftfeld wird durch ein Potenzial  $V(r)$  beschrieben, das nur vom Abstand  $r$  des Teilchens vom Kraftzentrum, welches mit dem Koordinatenursprung identisch sein soll, abhängt. In der klassischen Mechanik zeigt man, dass bei einer Bewegung in einem Zentralfeld der Vektor des Drehimpulses erhalten bleibt, also die Komponenten dieses Vektors  $L_x$ ,  $L_y$  und  $L_z$  Konstanten der Bewegung sind. Übersetzt in die Quantenmechanik bedeutet dies, dass der Hamiltonoperator, das Quadrat des Drehimpulses  $L^2$  und eine der Komponenten, also z.B.  $L_z$  ein gemeinsames Eigenfunktionsystem besitzen. Das heisst aber, dass die Lösungen der stationären Schrödingergleichung, also die Eigenfunktionen des Hamiltonoperators, so gewählt werden können, dass bei der Ortsdarstellung dieser Lösungen in Kugelkoordinaten der Winkelanteil der Wellenfunktion  $\Psi(r, \vartheta, \varphi)$  durch die Kugelflächenfunktionen gegeben ist, die ja nach den Ausführungen im vorhergehenden Abschnitt gerade die Eigenfunktionen zu  $\hat{L}^2$  und  $\hat{L}_z$  darstellen. Die Wellenfunktionen besitzen also die Gestalt

$$\langle r, \vartheta, \varphi | \Psi_{Elm} \rangle = \Psi_{Elm}(r, \vartheta, \varphi) = R_{Elm}(r) Y_{lm}(\vartheta, \varphi). \quad (7.30)$$

Die Indizes  $E, l, m$  an dieser Wellenfunktion beziehen sich auf die entsprechenden Eigenwerte für den Hamiltonoperator,  $E$ , den Betrag des Drehimpulses,  $l$ , und seiner Projektion auf die  $z$ -Achse. Den Radialanteil dieser Wellenfunktion,  $R_{Elm}(r)$ , erhalten wir durch die Lösung der stationären Schrödingergleichung für ein Teilchen der Masse  $\mu$  im Potenzial  $V(r)$  des zentralen Kraftfeldes

$$\begin{aligned} \hat{H} \Psi_{Elm}(r, \vartheta, \varphi) &= \left[ -\frac{\hbar^2}{2\mu} \Delta + V(r) \right] \Psi_{Elm}(r, \vartheta, \varphi) \\ &= E \Psi_{Elm}(r, \vartheta, \varphi). \end{aligned} \quad (7.31)$$

Dabei bezeichnet  $\Delta$  den **Laplace Operator**, die Summe der zweiten Ableitungen nach den kartesischen Koordinaten

$$\Delta = \frac{d^2}{dx^2} + \frac{d^2}{dy^2} + \frac{d^2}{dz^2}. \quad (7.32)$$

Für die Behandlung von Zentralfeldproblemen bietet es sich natürlich an, auch diesen Laplace Operator in Kugelkoordinaten darzustellen. Dies ergibt

$$\Delta = \frac{1}{r} \frac{d^2}{dr^2} r + \frac{1}{r^2} \underbrace{\left[ \frac{1}{\sin \vartheta} \frac{d}{d\vartheta} \sin \vartheta \frac{d}{d\vartheta} + \frac{1}{\sin^2 \vartheta} \frac{d^2}{d\varphi^2} \right]}_{= -\frac{\hat{L}^2}{\hbar^2}}. \quad (7.33)$$

Der Term in Abhängigkeit von den Winkelvariablen  $\vartheta$  und  $\varphi$  ist proportional zum Operator  $\hat{L}^2$  aus (7.11). Damit ergibt sich für den Hamiltonoperator in der Ortsdarstellung mit den Kugelkoordinaten die Gestalt

$$\hat{H} = -\frac{\hbar^2}{2\mu} \frac{1}{r} \frac{d^2}{dr^2} r + \frac{\hat{L}^2}{2\mu r^2} + V(r). \quad (7.34)$$

Wendet man diesen Hamiltonoperator auf die Wellenfunktion von (7.30) an, so kann man den Operator  $\hat{L}^2$  durch den zugehörigen Eigenwert ersetzen und erhält für die stationäre Schrödingergleichung (7.31)

$$\left[ -\frac{\hbar^2}{2\mu} \frac{1}{r} \frac{d^2}{dr^2} r + \frac{\hbar^2 l(l+1)}{2\mu r^2} + V(r) \right] R_{Elm}(r) Y_{lm}(\vartheta, \varphi) = E R_{Elm}(r) Y_{lm}(\vartheta, \varphi). \quad (7.35)$$

Man kann nun auf beiden Seiten der Gleichung den winkelabhängigen Faktor  $Y_{lm}(\vartheta, \varphi)$  eliminieren<sup>1</sup> und erhält eine Gleichung für den Radialteil  $R_{Elm}(r)$

$$\left[ -\frac{\hbar^2}{2\mu} \frac{1}{r} \frac{d^2}{dr^2} r + \frac{\hbar^2 l(l+1)}{2\mu r^2} + V(r) \right] R_{Elm}(r) = E R_{Elm}(r). \quad (7.36)$$

Diese Bestimmungsgleichung ist offensichtlich unabhängig von der Projektionsquantenzahl  $m$  und wir werden diese Quantenzahl im folgenden als Index an der Radialfunktion streichen. Damit sind also sowohl die Radialfunktion als auch die Energieeigenwerte unabhängig von der Projektionsquantenzahl  $m$ . Für jeden Energieeigenwert gibt es also Eigenzustände mit  $m = -l, \dots, +l$ , der Energieeigenwert ist also  $2l + 1$ -fach entartet.

Machen wir nun für die Radialfunktion den Ansatz

$$R_{El}(r) = \frac{u_{El}(r)}{r}, \quad (7.37)$$

setzen diesen Ansatz in (7.36) ein und multiplizieren die Gleichung mit  $r$ , so ergibt sich eine Schrödinger Gleichung für  $u_{El}(r)$  der Form

$$\left[ -\frac{\hbar^2}{2\mu} \frac{d^2}{dr^2} + \underbrace{\frac{\hbar^2 l(l+1)}{2\mu r^2}}_{=V_{eff}(r)} + V(r) \right] u_{El}(r) = E u_{El}(r). \quad (7.38)$$

Diese Differenzialgleichung zur Bestimmung der Energieeigenwerte hat also die gleiche Form wie die Schrödinger Gleichung für ein ein-dimensionales Problem mit einem effektiven Potenzial

$$V_{eff}(r) = V(r) + \frac{\hbar^2 l(l+1)}{2\mu r^2}, \quad (7.39)$$

das neben dem ursprünglichen Potenzial  $V(r)$  noch einen Zentrifugalanteil besitzt, der vom Betrag des Drehimpulses des Teilchens abhängt und für  $l > 0$  zum Koordinatenursprung hin divergiert. Beispiele für solche effektiven Potentiale sind in Abb. 7.4 skizziert.

Für weitere Aussagen zur Lösung der Differenzialgleichung (7.38) beschränken wir uns auf Potentiale, die für  $r \rightarrow 0$  schwächer divergieren als das Zentrifugalpotenzial, als schwächer als  $1/r^2$ . In diesem Fall kann man also bei der Differenzialgleichung im Grenzfall  $r \rightarrow 0$  den Term mit der konstanten Energie und den Term proportional  $V(r)$  gegenüber den weiteren Termen vernachlässigen und erhält also in diesem Grenzfall eine Differenzialgleichung der Form

$$\left[ -\frac{d^2}{dr^2} + \frac{l(l+1)}{r^2} \right] u_l(r) = 0. \quad (7.40)$$

<sup>1</sup>Etwas präziser gesagt multiplizieren wir beide Seiten von (7.35) mit  $Y_{lm}^*(\vartheta, \varphi)$ , integrieren über die Winkel und erhalten wegen der Orthonormalität der Kugelflächenfunktionen (7.36)



## 7.2. ZENTRALFELDE

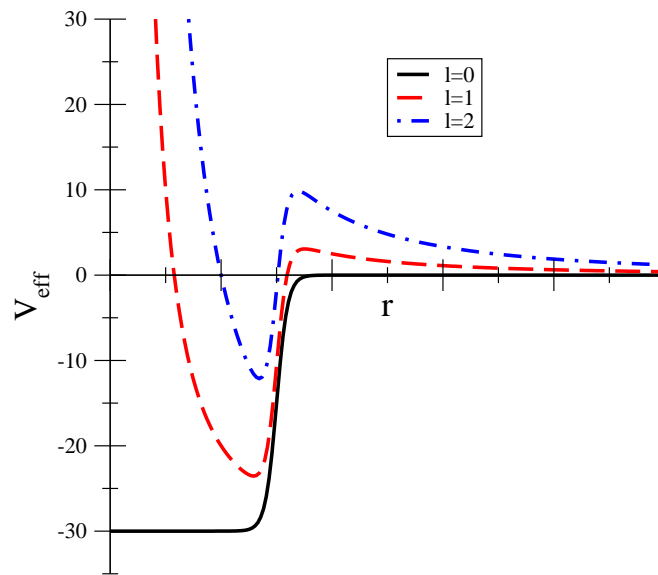


Abbildung 7.4: *Effektives Potenzial für die radiale Schrödingergleichung nach (7.39)*

Diese Differentialgleichung hat, wie man durch Einsetzen leicht feststellen kann die beiden unabhängigen Lösungen:

$$u_{l1}(r) = r^{l+1} \quad \text{und} \quad u_{l2}(r) = \frac{1}{r^l}.$$

Die Lösung  $u_{l2}(r)$  liefert eine Wahrscheinlichkeitsdichte  $\rho(r) \sim u^2/r^2$ , die für  $r \rightarrow 0$  divergiert und ist deshalb unphysikalisch. Wir schliessen daraus, dass die radiale Funktion  $u(r)$  für kleine Wert von  $r$  die asymptotische Form

$$\lim_{r \rightarrow 0} u_l(r) = a r^{l+1}, \quad (7.41)$$

annimmt.

### 7.3 Das Wasserstoffatom

Das Wasserstoffatom besteht aus einem Atomkern, der für den normalen Wasserstoff einfach durch ein Proton gegeben ist, mit der Masse  $m_p$ , und einem Elektron mit der Masse  $m_e$ . Vernachlässigen wir zunächst einmal die magnetischen Wechselwirkungen, die durch die magnetischen Momente der Protonen und Elektronen bewirkt werden, so werden Proton und Elektron allein durch die Coulomb Wechselwirkung zwischen der positiven Ladung des Protons und der negativen Ladung des Elektrons zusammengehalten. Die Hamiltonfunktion dieses Systems aus zwei Teilchen ergibt sich also als die Summe der kinetischen Energien des Protons und des Elektrons plus dem attraktiven Coulomb Potenzial zwischen diesen beiden Teilchen. Bezeichnet man mit  $\vec{p}_p, \vec{p}_e$  und  $\vec{r}_p, \vec{r}_e$  die Impuls- und Ortsvektoren dieser Teilchen, so ergibt sich die Hamiltonfunktion zu

$$H = \frac{\vec{p}_p^2}{2m_p} + \frac{\vec{p}_e^2}{2m_e} - \frac{\tilde{e}^2}{4\pi\epsilon_0 |\vec{r}_e - \vec{r}_p|}. \quad (7.42)$$

Dabei steht  $\tilde{e}$  für die Elementarladung, also den Betrag der Ladung eines Elektrons oder eines Protons. Im SI System ist diese Elementarladung durch

$$\tilde{e} = 1.60217 \cdot 10^{-19} \text{ C},$$

gegeben und die Influenzkonstante hat den Wert

$$\epsilon_0 = 8.8542 \cdot 10^{-12} \frac{\text{C}}{\text{Vm}}.$$

In der Atom-, Kern- und Teilchenphysik wird häufig das Gauß'sche Maßsystem benutzt. Man ersetzt dazu in (7.42)

$$\frac{\tilde{e}^2}{4\pi\epsilon_0} \rightarrow e^2 \quad (7.43)$$

und gibt Energien in Einheiten der Elementarladung  $e$  mal einer Beschleunigungsspannung in Volt an, also z.B. in Einheiten "eV" an. Da sich durch diese Ersetzung einige Ausdrücke einfacher formulieren lassen und um den Vergleich mit Lehrbüchern zu diesem Thema zu vereinfachen, soll auch in diesem Abschnitt das Gauß'sche System benutzt werden.

Außerdem haben wir bereits im ersten Semester bei der Diskussion von Vielteilchensystemen gesehen, dass die Beschreibung von 2 Massenpunkten, auf die nur eine Kraft zwischen diesen beiden Punkten wirkt, am besten in den Koordinaten des Schwerpunktes des Systems und der Relativkoordinate beschrieben wird. Wir transformieren also die Ortsvektoren des Atomkerns und des Elektrons auf die Koordinaten des Schwerpunktvektors

$$\vec{R} = \frac{m_e \vec{r}_e + m_p \vec{r}_p}{m_e + m_p},$$

und die Koordinaten des Relativvektors

$$\vec{r} = \vec{r}_e - \vec{r}_p.$$

In diesen Schwerpunkts- und Relativkoordinaten lautet die Hamiltonfunktion

$$H = \frac{\vec{P}_{cm}^2}{2M} + \frac{\vec{p}^2}{2\mu} - \frac{e^2}{r}. \quad (7.44)$$

Dabei steht  $\vec{P}_{cm}$  für den Impuls der Bewegung des Schwerpunktes,  $M = m_e + m_p$  für die Gesamtmasse,  $\vec{p}$  für den Impuls der Relativbewegung und  $\mu$  für die sogenannte reduzierte Masse mit

$$\begin{aligned}\frac{1}{\mu} &= \frac{1}{m_e} + \frac{1}{m_p}, \quad \text{bzw.} \\ \mu &= \frac{m_e m_p}{M}.\end{aligned}\quad (7.45)$$

Für die Behandlung des Problems im Rahmen der Quantenmechanik bestimmen wir ausgehend von dieser Hamiltonfunktion den entsprechenden Hamiltonoperator. In der Ortsdarstellung ersetzen wir

$$\vec{P}_{cm} \rightarrow \frac{\hbar}{i} \vec{\nabla}_R \quad \text{und} \quad \vec{p} \rightarrow \frac{\hbar}{i} \vec{\nabla}_r,$$

wobei  $\vec{\nabla}_R$  den Operator des Gradienten bezüglich der Schwerpunktskoordinaten und  $\vec{\nabla}_r$  den bezogen auf die Relativkoordinate bezeichnet. Damit ergibt sich der Hamiltonoperator

$$\hat{H} = \underbrace{-\frac{\Delta_R}{2M}}_{=\hat{H}_{cm}} - \frac{\Delta_r}{2\mu} - \frac{e^2}{r} =: \hat{H}_r, \quad (7.46)$$

mit den Laplace Operator  $\Delta_R$  ( $\Delta_r$ ) für die Schwerpunkts- (Relativ-) koordinaten.

Der gesamte Hamiltonoperator ist also eine Summe aus 2 Termen, wobei der erste Summand  $\hat{H}_{cm}$  nur auf die Koordinaten des Schwerpunktvektors und der zweite Summand,  $\hat{H}_r$  nur auf die des Relativvektors wirkt.

In einem solchen Fall, in dem der Hamiltonoperator eine Summe von 2 Operatoren ist, die auf unterschiedliche Koordinaten, sagen wir  $r_1$  und  $r_2$ , wirken,

$$\hat{H} = \hat{H}_1 + \hat{H}_2,$$

gibt es für die stationäre Schrödingergleichung

$$\hat{H}\Psi(r_1, r_2) = E\Psi(r_1, r_2), \quad (7.47)$$

Lösungen der Form

$$\begin{aligned}\Psi(r_1, r_2) &= \phi(r_1)\chi(r_2) \quad \text{mit} \\ \hat{H}_1\phi(r_1) &= E_1\phi(r_1), \\ \hat{H}_2\chi(r_2) &= E_2\chi(r_2) \quad \text{und} \\ E &= E_1 + E_2.\end{aligned}\quad (7.48)$$

Die Lösung faktorisiert also in ein Produkt von 2 Wellenfunktionen, die jeweils als Lösung einer Schrödingergleichung für die Koordinaten 1 und 2 gewonnen werden. Die Gesamtenergie ist dann die Summe der jeweiligen Eigenwerte  $E_1$  und  $E_2$ .

Zum Beweis dieser faktorisierten Lösung nehmen wir den Produktansatz an und setzen ihn in die stationäre Schrödingergleichung ein

$$(\hat{H}_1 + \hat{H}_2) \phi(r_1)\chi(r_2) = \chi(r_2)\hat{H}_1\phi(r_1) + \phi(r_1)\hat{H}_2\chi(r_2) = E\phi(r_1)\chi(r_2).$$

Die erste Gleichung ergibt sich dadurch, dass z.B. der Operator  $H_1$  nicht auf die Koordinaten  $r_2$  wirkt, wir also die Funktion  $\chi(r_2)$  auch links vom Operator  $H_1$  schreiben können. Dividiert man den zweiten Teil dieser Gleichung durch  $\phi(r_1)\chi(r_2)$  so ergibt sich

$$\frac{1}{\phi(r_1)}\hat{H}_1\phi(r_1) + \frac{1}{\chi(r_2)}\hat{H}_2\chi(r_2) = E.$$

Der erste der beiden Summanden auf der linken Seite dieser Gleichung hängt allenfalls von den Koordinaten  $r_1$  ab, während der zweite allenfalls von  $r_2$  abhängen kann. Die rechte Seite der Gleichung ist aber ein Konstante  $E$ , unabhängig von  $r_1$  und  $r_2$ . Deshalb müssen aber auch die beiden Summanden auf der linken Seite Konstanten sein:

$$\frac{1}{\phi(r_1)}\hat{H}_1\phi(r_1) = E_1 \quad , \quad \frac{1}{\chi(r_2)}\hat{H}_2\chi(r_2) = E_2 \quad \text{mit} \quad E_1 + E_2 = E.$$

Multipliziert man die erste Gleichung mit  $\phi(r_1)$  und die zweite mit  $\chi(r_2)$ , so erhält man gerade die beiden Schrödingergleichungen aus (7.48).

Für den Hamiltonoperator des Wasserstoffatoms aus (7.46) bedeutet dies, dass es Lösungen der zugehörigen stationären Schrödingergleichung gibt in der Form

$$\Psi(\vec{r}, \vec{R}) = \phi(\vec{r})\chi(\vec{R}).$$

Der Teil der Wellenfunktion, der die Abhängigkeit von der Schwerpunktskoordinate beschreibt ergibt sich als Lösung der Schrödinger Gleichung

$$-\frac{\Delta_R}{2M}\chi(\vec{R}) = E_{cm}\chi(\vec{R}). \quad (7.49)$$

Dies ist die Schrödinger Gleichung für ein freies Teilchen der Masse  $M$ . Die Lösungen sind durch ebene Wellen gegeben der Form

$$\chi(\vec{R}) = \frac{1}{(2\pi)^{3/2}}e^{i\vec{K}\vec{R}},$$

mit einem Wellenzahlvektor  $\vec{K}$  und einer zugehörigen kinetischen Energie dieser Bewegung des Gesamtatoms

$$E_{cm} = \frac{\hbar^2 \vec{K}^2}{2M}.$$

Sehr viel interessanter ist natürlich die innere Struktur des Atoms, die durch die Schrödinger Gleichung

$$\left\{ -\frac{\hbar^2}{2\mu}\Delta_r - \frac{e^2}{r} \right\} \phi(\vec{r}) = E\phi(\vec{r}), \quad (7.50)$$

beschrieben wird. Diese Gleichung hat die Struktur einer Schrödinger Gleichung für ein Teilchen der Masse  $\mu$ , das sich in einem zentralen Kraftfeld, beschrieben durch das Coulomb Potenzial, bewegt. Durch die Faktorisierung der Gesamtwellenfunktion in einen Anteil für Schwerpunkts- und Relativkoordinaten ist also das Zweiteilchenproblem des Wasserstoffatoms auf die Gleichung eines Einteilchenproblems reduziert worden.

Zu beachten ist dabei lediglich, dass die Koordinaten des Relativvektors die relevanten Freiheitsgrade beschreiben und dass die Masse für dieses Einteilchenproblem durch die

reduzierte Masse  $\mu$  beschrieben wird. Im Fall des “normalen” Wasserstoffatoms ist die Masse des Atomkerns etwa 2000 mal so groß wie die des Elektrons, da ja mit der Lichtgeschwindigkeit  $c$  für die Ruheenergie dieser Teilchen gilt

$$m_p c^2 = 938 \text{ MeV} \quad \text{und} \quad m_e c^2 = 0.511 \text{ MeV}.$$

In diesem Fall ist der Vektor des Schwerpunktes des Systems praktisch identisch mit dem Ortsvektor des Protons und die reduzierte Masse (siehe (7.45)) praktisch identisch mit der Masse des Elektrons

$$\mu c^2 = 0.5107 \text{ MeV}.$$

Mit der Schrödinger Gleichung (7.50) können wir aber auch exotischere Atome beschreiben, wie z.B. das myonische Atom. Dabei wird das Elektron ersetzt durch ein **Myon**. Ein Myon ist ein Elementarteilchen wie das Elektron. Es gehört ebenfalls zu den **Leptonen** und hat die gleiche Ladung. Lediglich seine Masse ist erheblich größer mit  $m_{my} c^2 = 105 \text{ MeV}$ , so dass die reduzierte Masse für das myonische Atom mit  $\mu c^2 = 94 \text{ MeV}$  deutlich kleiner ist als  $m_{my} c^2$ .

Im vorhergehenden Abschnitt haben wir gesehen, dass die Lösungen einer stationären Schrödinger Gleichung für ein zentrales Kraftfeld wie z.B. in (7.50) faktorisiert werden kann in der Form (siehe (7.30) und (7.37))

$$\phi_{nlm}(\vec{r}) = \frac{u_{nl}(r)}{r} Y_{lm}(\vartheta, \varphi), \quad (7.51)$$

mit den Kugelflächenfunktionen  $Y_{lm}$  und einer Radialfunktion  $u_{nl}$ , die im Fall des Wasserstoffatoms bestimmt wird aus der Lösung der radialen Gleichung

$$\left\{ -\frac{\hbar^2}{2\mu} \frac{d^2}{dr^2} + \frac{\hbar^2 l(l+1)}{2\mu r^2} - \frac{e^2}{r} \right\} u_{nl}(r) = E_{nl} u_{nl}(r), \quad (7.52)$$

wobei die Quantenzahl  $n$  eingeführt wurde, um die verschiedenen Lösungen dieser Gleichung bei gegebenem  $l$  zu unterscheiden.

Bevor wir uns der expliziten Lösung dieser Gleichung zuwenden, wollen wir die Energie für den tiefsten gebundenen Zustand abschätzen. Da der Zentrifugalterm in der Gleichung (7.52) für  $l > 0$  einen repulsiven Beitrag zur Energie liefert, ist es klar, dass wir den energetisch tiefsten Zustand für  $l = 0$  erwarten. In diesem Fall besteht der Hamiltonoperator nur aus einem Beitrag der radialen kinetischen Energie, der positiv sein wird, und dem attraktiven Energiebeitrag des Coulomb Potentials.

Je näher das Elektron am Atomkern ist, um so attraktiver ist der Beitrag des Coulomb Potentials. Im Abstand  $r_0$  beträgt ja der Wert dieses Coulomb Potentials

$$V_0 = -\frac{e^2}{r_0}.$$

Schränkt man aber den Aufenthaltsort des Elektrons auf einen Bereich um  $r_0$  ein so dass das mittlere Schwankungsquadrat  $\Delta^2(r)$  etwa gleich  $r_0^2/4$  ist, so hat die Heisenbergsche Unschärferelation

$$\Delta^2(p) \Delta^2(r) \geq \frac{\hbar^2}{4},$$

zur Folge, dass

$$\Delta^2(p) = \langle p^2 \rangle - \langle p \rangle^2 = \langle p^2 \rangle \geq \frac{\hbar^2}{r_0^2},$$

sein muss. Dabei haben wir benutzt, dass der Erwartungswert für den Impuls  $\langle p \rangle = 0$  sein wird. Je kleiner  $r_0$  um so größer wird also der Erwartungswert des Impulsquadrates,  $\langle p^2 \rangle$ , sein und entsprechend positiver die kinetische Energie

$$T_0 = \frac{p^2}{2\mu} = \frac{\hbar^2}{2\mu r_0^2}.$$

Zur Abschätzung der minimalen Energie fassen wir die Summe aus dieser kinetischen Energie und der potenziellen Energie zusammen

$$E(r_0) = \frac{\hbar^2}{2\mu r_0^2} - \frac{e^2}{r_0}, \quad (7.53)$$

und bestimmen den Wert von  $r_0$  an dem diese Funktion minimal ist

$$\frac{dE}{dr_0} = -\frac{\hbar^2}{\mu r_0^3} + \frac{e^2}{r_0^2} = 0$$

was zu einem Wert für

$$r_0 = \frac{\hbar^2}{\mu e^2} = 0.53 \times 10^{-10} \text{ m}, \quad (7.54)$$

führt. Der numerisch Wert von 0.53 Angstrom ergibt sich, wenn man die reduzierte Masse  $\mu$  für das Elektron, also für das gewöhnliche Wasserstoffatom einsetzt. Diesen Ausdruck für  $r_0$  bezeichnet man auch als den **Bohr'schen Radius** für das Wasserstoffatom.

Berechnet man die Energie für diesen Wert von  $r_0$  nach (7.53), so ergibt sich

$$\begin{aligned} E_0 &= \frac{\mu e^4}{2\hbar^2} - \frac{\mu e^4}{\hbar^2} \\ &= -\frac{\mu e^4}{2\hbar^2} \\ &= -13.5 \text{ eV} \end{aligned} \quad (7.55)$$

Der numerische Wert für die Energie von -13.5 eV ergibt sich wieder für den Fall des normalen Wasserstoffatoms. Bemerkenswert am Ausdruck in der ersten Zeile ist, dass der Wert für die potenzielle Energie mit dem Wert der kinetischen Energie verknüpft ist durch

$$T_0 = -\frac{1}{2}V_0.$$

Dies entspricht der Vorhersage des Virialsatzes für ein Potenzial der Form  $1/r$  in der klassischen Mechanik. (siehe Abschnitt 3.4 aus der Vorlesung Physik I).

Den Ausdruck (7.55) kann man einfach umschreiben in

$$E_0 = -\frac{1}{2} \left( \frac{e^2}{\hbar c} \right)^2 \mu c^2, \quad (7.56)$$

wobei der Quotient

$$\alpha = \frac{e^2}{\hbar c} = \frac{\tilde{e}^2}{4\pi\epsilon_0\hbar c} \approx \frac{1}{137}, \quad (7.57)$$

die sogenannte **Sommerfeldsche Feinstrukturkonstante**, ein dimensionslose Zahl bestehend aus Naturkonstanten, bezeichnet.

Betrachtet man jetzt ein exotisches Atom, wie z.B. ein myonisches Atom, mit einem schweren Lepton, also einem größerem Wert für die reduzierte Masse  $\mu$ , so wird aus diesen Überlegungen deutlich, dass die kinetische Energie bei gleichem Wert für  $p^2$  kleiner sein wird als im Fall des Elektrons. Das Minimum der Funktion  $E(r_0)$  wird also bei einem kleineren Radius auftreten (siehe (7.54)) was gleichzeitig zu mehr Bindungsenergie, also einem negativeren Wert für  $E(r_0)$  führt (siehe (7.56)).

Nach dieser Abschätzung wenden wir uns nun wieder der Lösung der radialen Gleichung (7.52) für gebundene Zustände, also negativen Werten für  $E_{nl}$  zu. Dazu führen wir eine Substitution der Koordinate  $r$  durch auf die dimensionslose Variable

$$\rho = \sqrt{\frac{8\mu(-E_{nl})}{\hbar^2}} r. \quad (7.58)$$

Damit bringen wir (7.52) auf die Form

$$\left\{ \frac{d^2}{d\rho^2} - \frac{l(l+1)}{\rho^2} + \frac{\lambda}{\rho} - \frac{1}{4} \right\} u_{nl}(\rho) = 0, \quad (7.59)$$

mit

$$\lambda := \frac{e^2}{\hbar} \sqrt{\frac{\mu}{2(-E_{nl})}}. \quad (7.60)$$

Zunächst betrachten wir die Asymptotik dieser Gleichung für große Werte von  $r$  bzw.  $\rho$ . Bei dieser asymptotischen Betrachtung können also die Terme proportional zu  $1/\rho$  und  $1/\rho^2$  vernachlässigt werden. Dadurch reduziert sich (7.59) auf

$$\left\{ \frac{d^2}{d\rho^2} - \frac{1}{4} \right\} u_{nl}(\rho) = 0, \quad (7.61)$$

mit den Lösungen

$$u_{nl,1} = e^{-\rho/2} \quad \text{und} \quad u_{nl,2} = e^{+\rho/2}.$$

Da die Lösung  $u_{nl,2}$  zu Wahrscheinlichkeitsdichten führt, die für große Werte von  $r$  divergieren hat die reguläre Lösung also die Asymptotik der Form  $u_{nl,1}$ . Zusammen mit dem Verhalten bei kleinen Werten von  $r$  bzw.  $\rho$ , das wir bereits im vorhergehenden Abschnitt gefunden haben (siehe (7.41)) ergibt sich also der Ansatz

$$u_{nl} = \rho^{l+1} e^{-\rho/2} H_{nl}(\rho). \quad (7.62)$$

Wir sehen auch in diesem Fall, dass die Lösung der Schrödinger Gleichung für negative Energien, also gebundene Elektronen, durch die Asymptotik für  $r \rightarrow 0$  und  $r \rightarrow \infty$  zwei Randbedingungen erfüllen muss. Diese Zahl der Randbedingungen ist so groß, dass es nicht für jede Energie eine Lösung geben kann. Wir erwarten also auch hier Lösungen der stationären Schrödinger Gleichung bei  $E < 0$  nur für diskrete Energien.

Für  $E > 0$  wird der Ausdruck (7.58) für  $\rho$  rein imaginär. In diesem Fall sind also beide Lösungen der asymptotischen Gleichung für  $r \rightarrow \infty$  regulär, so dass es nur eine Randbedingung für  $r \rightarrow 0$  zu beachten gilt. Deshalb wird es für jede positive Energie eine Lösung der Schrödinger Gleichung geben, die dann allerdings ein ungebundenes Elektron beschreibt.

Die unbekannte Funktion  $H_{nl}(\rho)$  im Ansatz (7.62) darf das asymptotische Verhalten von  $u_{nl}$  für  $\rho \rightarrow 0$  und  $\rho \rightarrow \infty$  nicht zerstören. Deshalb setzen wir für diese Funktion  $H_{nl}(\rho)$  ein Polynom endlichen Grades  $n'$  an:

$$H_{nl}(\rho) = \sum_{i=0}^{n'} a_i \rho^i. \quad (7.63)$$

Setzt man den Ansatz (7.62) in die Differenzialgleichung (7.59) ein, so ergibt sich für diese Funktion  $H_{nl}(\rho)$  die Gleichung

$$\rho \frac{d^2 H_{nl}}{d\rho^2} + (2l + 2 - \rho) \frac{dH_{nl}}{d\rho} + (\lambda - l - 1) H_{nl} = 0,$$

was mit dem Ansatz (7.63) für die Koeffizienten des Polynoms zu einer Rekursionsformel führt:

$$a_{i+1} = \frac{i + l + 1 - \lambda}{(i + 1)(i + 2l + 2)} a_i. \quad (7.64)$$

Damit dieses Rekursionsformel bei dem Wert  $i = n'$  abbricht muss gelten

$$\lambda = n' + l + 1, \quad (7.65)$$

Durch diese Rekursionsformel sind die Polynome bis auf eine Konstante eindeutig definiert. Wir können ja z.B. mit  $a_0 = 1$  starten und dann über die Rekursionsformel (7.65) alle Koeffizienten  $a_i$  bis  $i = n'$  bestimmen. Das Ergebnis ist ein Polynom vom Grade  $n'$ , das bis auf einen globalen Normierungsfaktor  $\gamma$  identisch ist mit den sogenannten **Laguerre Polynom**, so dass

$$H_{nl}(\rho) = \gamma L_n^{2l+1}(\rho), \quad (7.66)$$

wobei die Laguerreschen Polynome auch definiert sind durch

$$L_n^k(\rho) = \sum_{i=0}^n (-1)^i \frac{[(n+k)!]^2}{(p-i)!(k+i)!i!} \rho^i. \quad (7.67)$$

Die Abbruchbedingung (7.65) führt aber wegen (7.60) auch (mit  $n' = n$ ) zu einer Bestimmungsgleichung für die möglichen Energien

$$E_{nl} = -\frac{1}{2} \frac{e^4}{\hbar^2 c^2} \mu c^2 \frac{1}{(n+l+1)^2}. \quad (7.68)$$

Die niedrigste Energie ergibt sich also für  $n = l = 0$  und entspricht genau der Energie unserer Abschätzung in (7.55) bzw. (7.56).

Neben dieser Energie für den niedrigsten Zustand, den **Grundzustand** des Wasserstoffatoms ergeben sich aber auch unendlich viele Energieeigenwerte der Form

$$E_{nl} = -(13.5 \text{ eV}) \frac{1}{N^2}, \quad \text{mit } N = n + l + 1. \quad (7.69)$$



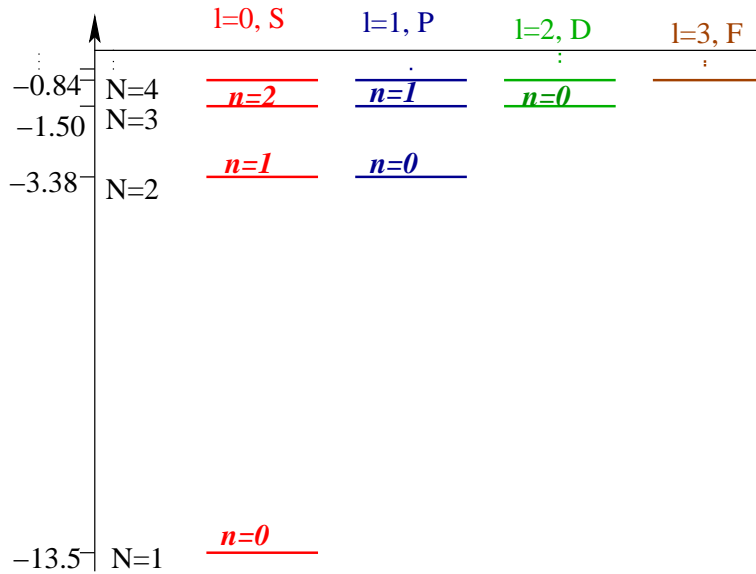


Abbildung 7.5: Darstellung der niedrigsten Energieeigenwerte des Wasserstoffatoms

Der numerische Wert -13.5 eV ist eine Näherung für die entsprechende Konstante in (7.68) im Fall des normalen Wasserstoffatoms und trägt den Namen **Rydberg Konstante**. Die Zahl  $N = n + l + 1$  bezeichnet man in der Regel als **Hauptquantenzahl** für die Energien des Wasserstoffatoms.

Diese Energien des Wasserstoffatoms, das sogenannte Termschema, ist in Abb. 7.5 skizziert. In der Nähe des Energienullpunktes wird die Niveaudichte für  $E < 0$  unendlich groß; bei positiven Energien, also ungebundenen Elektronen gibt es, wie bereits oben diskutiert, ein Kontinuum von Energien.

Wir wollen nun einige dieser Lösungen etwas genauer betrachten. Für den Grundzustand ist die Hauptquantenzahl  $N = n + l + 1 = 1$ , was bedeutet, dass  $n = l = 0$  sein muss. Damit ist auch die Quantenzahl  $m = 0$  festgelegt und die Wellenfunktion für den Grundzustand ist eindeutig definiert. Da die Radialquantenzahl  $n$  identisch null ist, ist das Polynom  $H_{nl}$  (vom Grade  $n=0$ ) eine Konstante und die Wellenfunktion ergibt sich nach (7.51) und (7.62) zu

$$\phi_{000}(r, \vartheta, \varphi) = \gamma \frac{\rho}{\rho} e^{-\rho/2} Y_{00}(\vartheta, \varphi). \quad (7.70)$$

Mit der Definition von  $\rho$  in (7.58) und der Energie  $E_{00}$  nach (7.68) ergibt sich

$$\rho = 2\mu \frac{e^2}{\hbar^2} r = 2 \frac{r}{r_0},$$

wobei  $r_0$  den Bohrschen Radius nach (7.54) bezeichnet. Setzt man diesen Ausdruck in (7.70) und berücksichtigt, dass  $Y_{00} = 1/\sqrt{4\pi}$ , so ergibt sich

$$\phi_{000}(r, \vartheta, \varphi) = \gamma \frac{1}{\sqrt{4\pi}} e^{-r/r_0}. \quad (7.71)$$

Die Normierungskonstante  $\gamma$  ergibt sich aus der Normierungsbedingung

$$1 = \int d^3r \phi_{000}^2(r, \vartheta, \varphi)$$

$$\begin{aligned}
&= \gamma^2 \int_0^\infty r^2 e^{-2r/r_0} dr \\
&= \gamma^2 \frac{1}{4} r_0^3.
\end{aligned} \tag{7.72}$$

Dies bedeutet also, dass die Normierungskonstante

$$\gamma = 2 \left( \frac{1}{r_0} \right)^{3/2}.$$

Damit ergibt sich also die Radialfunktion für  $\phi_{000}$

$$R_{00}(r) = \frac{u_{00}(r)}{r} = 2 \left( \frac{1}{r_0} \right)^{3/2} e^{-r/r_0}. \tag{7.73}$$

Mit dieser Wellenfunktion können auch andere Erwartungswerte berechnet werden. So ergibt sich z.B. als Maß für den Radius des Atoms im Grundzustand der Erwartungswert

$$\langle n = l = 0 | r | n = l = 0 \rangle = \int_0^\infty r^2 dr r R_{00}^2(r) = \frac{3}{2} r_0, \tag{7.74}$$

was erstaunlich gut mit der Abschätzung (7.54) übereinstimmt.

Für den Grundzustand gibt es nur eine einzige Wellenfunktion, man sagt der Grundzustand ist nicht entartet oder einfach entartet<sup>2</sup>. Für die Zustände mit der Hauptquantenzahl  $N = 2$  gibt es wegen (7.69) die Möglichkeiten

$$\begin{aligned}
n = 1 \quad l = 0 \quad m = 0 \quad \text{oder} \\
n = 0 \quad l = 1 \quad m = -1, 0, 1.
\end{aligned} \tag{7.75}$$

Wir haben also 4 Kombinationen von Quantenzahlen vorliegen, die alle zur gleichen Hauptquantenzahl  $N = 2$  führen und damit die gleiche Energie besitzen. Die zugehörige Energie ist also 4-fach (mit Berücksichtigung des Elektronenspins 8-fach) entartet.

Die radialen Wellenfunktionen können wieder wie für  $N = 1$  angedeutet bestimmt werden. Man beachte dabei, dass die Definition von  $\rho$  nach (7.58) von der zugehörigen Energie abhängt. Sie ergeben sich zu

$$\begin{aligned}
R_{10}(r) &= 2 \left( \frac{1}{2r_0} \right)^{3/2} \left( 1 - \frac{r}{2r_0} \right) e^{-r/2r_0} \\
R_{01}(r) &= \frac{1}{\sqrt{3}} \left( \frac{1}{2r_0} \right)^{3/2} \frac{r}{r_0} e^{-r/2r_0}.
\end{aligned} \tag{7.76}$$

Die Zustände für  $l = 0$  werden in der Atomspektroskopie und in der Folge davon auch in anderen Bereichen der Physik als s-Zustände bezeichnet, die mit  $l = 1, 2, \dots$  als  $p, d, f, g, \dots$  Zustände (siehe auch die Bezeichnungen in Abb. 7.5).

Die s-Zustände führen zu kugelsymmetrischen Dichteverteilungen für den Aufenthaltsort des Elektrons relativ zum Atomkern. Während im Fall  $n = 0$  diese Dichteverteilung nur

---

<sup>2</sup>Wie wir noch sehen werden, ist diese Aussage nicht ganz richtig: Wegen des Spins des Elektrons und der damit verbundenen Entartung von 2 Zuständen, die sich durch die Spinorientierung unterscheiden, ist der Entartungsgrad dieses Modells ohne Berücksichtigung des Elektronenspins mit einem Faktor 2 zu multiplizieren

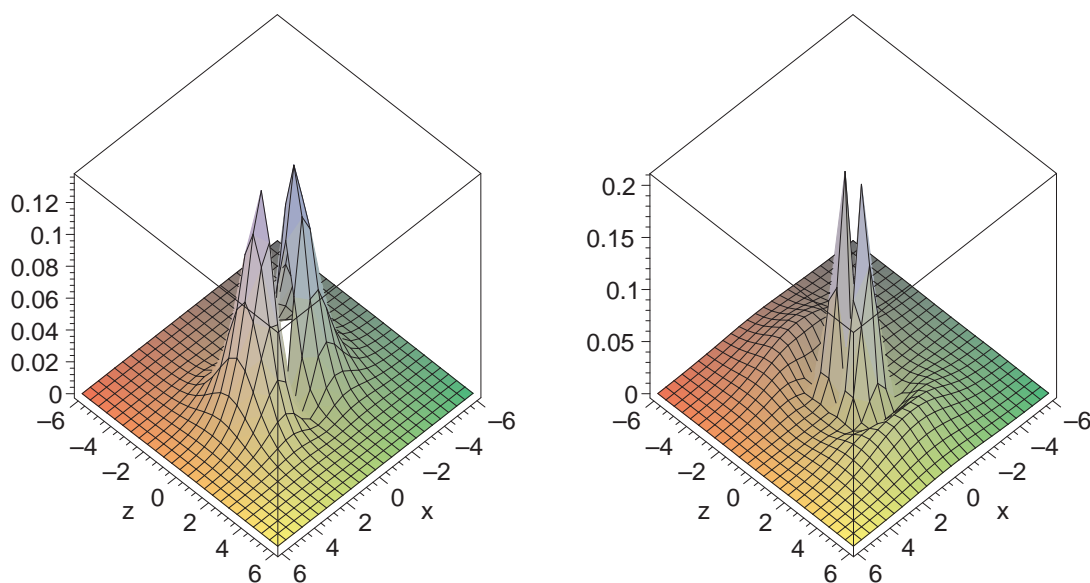


Abbildung 7.6: Darstellung der Dichteverteilung für  $p$ -Zustände des Wasserstoffatoms in der  $xz$ -Ebene. Im linken Teilbild ist die Verteilung für  $n = 0$ ,  $l = 1$ ,  $m = 1$  dargestellt, während das rechte Teilbild die für  $n = 1$ ,  $l = 1$ ,  $m = 0$  zeigt. Beachte, dass die Koordinate in Einheiten von  $\rho$  angegeben sind.

ein Maximum am Koordinatenursprung aufweist, gibt es im Fall  $n = 1$  eine Nullstelle in der radialen Wellenfunktion  $R_{10}$  (siehe (7.76)) was zu zwei lokalen Maxima bei  $r = 0$  und bei einem weiteren Wert von  $r$  führt.

Die  $p$ -Zustände zeigen neben dem entsprechenden Laguerre Polynom eine weiteren Faktor  $r$  in der Radialfunktion. Dieser Faktor ist eine Folge des Zentrifugalpotenzials und führt dazu, dass die Wahrscheinlichkeitsdichte für die Elektronen in Zuständen mit  $l = 1$  (und größer) bei  $r = 0$  identisch null ist.

Bei den  $p$ -Zuständen ergibt sich durch die Kugelflächenfunktionen auch eine Richtungsabhängigkeit in der Dichteverteilung. Als Beispiel ist im linken Teilbild von Abb. 7.6 die Dichteverteilung für  $n = 0$ ,  $l = 1$ ,  $m = 1$ . In diesem Fall ist der Drehimpulsvektor also parallel zur  $z$ -Achse ausgerichtet. Das hat zur Folge, dass die Dichteverteilung bei  $z = 0$  und einem endlichen Wert von  $|x|$  maximal wird.

Anders ist die Situation im rechten Teilbild von Abb. 7.6. Hier handelt es sich um die Dichteverteilung für  $n = 1$ ,  $l = 1$ ,  $m = 0$ . Die Dichteverteilung wird hier maximal auf der  $z$ -Achse. Da wir hier  $n = 1$  also die Hauptquantenzahl  $N = 3$  vorliegen haben, hat die Radialfunktion eine weitere Nullstelle und es gibt zwei Maxima für positive Werte von  $z$  und zwei für negative Werte von  $z$ .

Als weiteres Beispiel sind in der Abb. 7.7 Dichteverteilungen für verschiedene  $f$ -Zustände ( $l = 3$ ) mit  $n = 0$  also einer Hauptquantenzahl  $N = 4$  dargestellt.

Als weiteres Ergebnis geben wir ohne expliziten Beweis an, dass der Entartungsgrad der

Energieniveaus mit der Hauptquantenzahl  $N$  gegeben ist durch

$$\sum_{l=0}^{N-1} (2l+1) = N^2. \quad (7.77)$$

Ausserdem kann man ausrechnen, dass der Erwartungswert für den radialen Abstand des Elektrons vom Atomkern in den verschiedenen Zuständen gegeben ist durch

$$\langle nlm|r|nlm\rangle = \int_0^\infty r^2 dr r R_{nl}^2(r) = \frac{r_0}{2} \left( 3(n+l+1)^2 - l(l+1) \right). \quad (7.78)$$

(vergleiche auch (7.74)). Der Erwartungswert wächst also mit der Hauptquantenzahl quadratisch an. Einerseits liegt das an den Polynomen in den Radialfunktionen, deren Ordnung bei festgehaltenem  $l$  mit  $n$  anwächst. Andererseits liegt dies aber auch etwa im Fall ( $n=0, N=l+1$ ) an der Zentrifugalbarriere, bzw. dem daraus resultierenden Faktor  $r^l$  in der radialen Wellenfunktion.

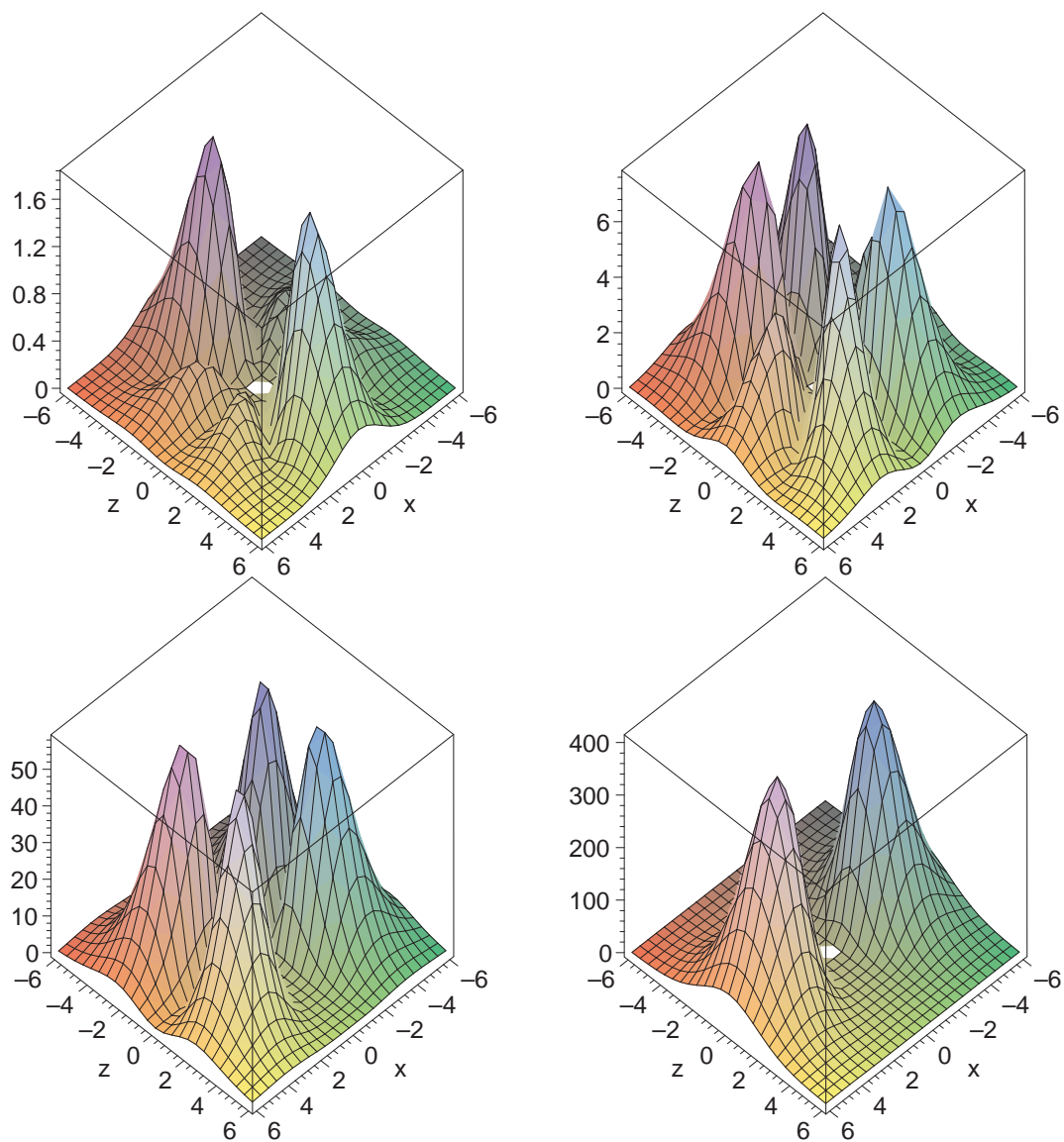


Abbildung 7.7: Darstellung der Dichteverteilung für  $f$ -Zustände des Wasserstoffatoms in der  $xz$ -Ebene. In den Teilbildern sind für  $n = 0$  von oben links nach unten rechts die Ergebnisse für  $m = 0, 1, 2$  und  $3$  dargestellt.