

Markus Vogt

---

# **Neue Wege in der Schulphysik?**

**Ein Beitrag zur Kritik aktueller  
Entwicklungen in der Physikdidaktik**

---



**Cuvillier Verlag Göttingen**  
Internationaler wissenschaftlicher Fachverlag

Markus Vogt

# Neue Wege in der Schulphysik?

Ein Beitrag zur Kritik aktueller Entwicklungen in der Physikdidaktik



**Cuvillier Verlag Göttingen**  
Internationaler wissenschaftlicher Fachverlag

**Bibliografische Information der Deutschen Nationalbibliothek**

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

1. Aufl. - Göttingen: Cuvillier, 2010

978-3-86955-242-2

© CUVILLIER VERLAG, Göttingen 2010

Nonnenstieg 8, 37075 Göttingen

Telefon: 0551-54724-0

Telefax: 0551-54724-21

[www.cuvillier.de](http://www.cuvillier.de)

Alle Rechte vorbehalten. Ohne ausdrückliche Genehmigung des Verlages ist es nicht gestattet, das Buch oder Teile daraus auf fotomechanischem Weg (Fotokopie, Mikrokopie) zu vervielfältigen.

1. Auflage, 2010

Gedruckt auf säurefreiem Papier

978-3-86955-242-2

# Inhaltsverzeichnis

<b>Vorwort</b>	<b>7</b>
<b>Einleitung</b>	<b>9</b>
<b>1 Bemerkungen zum Karlsruher Physikkurs</b>	<b>13</b>
1.1 Vorbemerkungen . . . . .	13
1.2 Physikalische Vorbetrachtungen . . . . .	15
1.2.1 Energie und der Energieerhaltungssatz . . . . .	15
1.2.2 Die physikalische Arbeit . . . . .	17
1.2.3 Die statistische Definition der Entropie . . . . .	18
1.2.4 Ein wenig Thermodynamik . . . . .	21
1.2.4.1 Zustandsgrößen . . . . .	22
1.2.4.2 Entropie als Zustandsgröße . . . . .	23
1.2.4.3 Irreversible Prozesse und der Satz von Clausius . . . . .	26
1.2.4.4 Das Nernstsche Wärmetheorem . . . . .	29
1.2.5 Die Gibbssche Fundamentalbeziehung . . . . .	32
1.2.5.1 Die kanonische Gesamtheit . . . . .	32
1.2.5.2 Die verallgemeinerte großkanonische Gesamtheit . . . . .	35
1.2.6 Anleihen aus der Strömungsmechanik . . . . .	39
1.3 Zwölf Argumente gegen den Karlsruher Physikkurs . . . . .	41
1.3.1 Physikalische Aspekte . . . . .	41
1.3.1.1 Die Bedeutung der Thermodynamik . . . . .	41
1.3.1.2 Energieträger oder Energieformen? . . . . .	43
1.3.1.3 Fehlvorstellungen zur Entropie . . . . .	48
1.3.1.4 Quantenmechanik . . . . .	51
1.3.2 Wissenschaftstheoretische Probleme . . . . .	52
1.3.2.1 Was sind die physikalischen Grundgrößen? . . . . .	52
1.3.2.2 Woher kommt der universelle Antrieb? . . . . .	55
1.3.2.3 Rückkehr zur Kontinuumsphysik? . . . . .	56
1.3.2.4 Fehlende Begriffsdefinitionen . . . . .	58
1.3.3 Didaktische Überlegungen . . . . .	59
1.3.3.1 Präkonzepte . . . . .	59
1.3.3.2 Analogien . . . . .	61
1.3.3.3 Verzicht auf wichtige Begriffe . . . . .	62
1.3.3.4 Was ist mit der etablierten Fachsprache? . . . . .	63

<b>2</b>	<b>Bemerkungen zum Zeigerformalismus</b>	<b>66</b>
2.1	Quantenmechanik und Schulphysik	66
2.2	Pfadintegrale in der Quantenmechanik	68
2.2.1	Der Schrödinger-Propagator	68
2.2.2	Pfadintegral-Darstellung des Schrödinger-Propagators	70
2.2.3	Feynman-Integrale und Standard-Quantenmechanik	79
2.2.4	Die Methode der stationären Phase	83
2.2.5	Pfadintegrale im Phasenraum	84
2.3	Der Satz von Cameron	86
2.3.1	Integraldarstellung des Zeitentwicklungsoperators	87
2.3.1.1	Die Trotter-Produktformel	87
2.3.1.2	Anwendung auf den Zeitentwicklungsoperator	92
2.3.2	Mathematische Unzulänglichkeiten der Feynman-Integrale	95
2.3.2.1	Unendlichdimensionale Maße	95
2.3.2.2	Maße auf Räumen stetiger Funktionen	96
2.3.2.3	Komplexe Maße	97
2.3.2.4	Weitere Schwierigkeiten	99
2.4	Die Feynman-Kac-Formel	101
2.4.1	Wiener-Prozeß, Wiener-Maß und Wiener-Integrale	101
2.4.1.1	Grundbegriffe aus der Stochastik	102
2.4.1.2	Integration über Brownsche Pfade	104
2.4.2	Beweis der Feynman-Kac-Formel	107
2.4.3	Analytische Fortsetzung	110
2.4.3.1	Operatorwertige Funktionalintegrale	110
2.4.3.2	Anwendung auf Feynman-Integrale	111
2.4.3.3	Euklidische Feldtheorie	114
2.5	Fresnel-Integrale	116
2.5.1	Definition und Eigenschaften	116
2.5.2	Die Feynman-Itô-Formel	119
2.5.3	Unendlichdimensionale oszillatorische Integrale	126
2.6	Weißes Rauschen	129
2.6.1	Gelfandsche Raumtripel	129
2.6.2	Analysis des weißen Rauschens	131
2.6.3	Verallgemeinerte Funktionenräume	137
2.6.3.1	Distributionen und Fock-Raum	137
2.6.3.2	Weitere Räume verallgemeinerter Funktionen	140
2.6.3.3	Hida- und Kondratiev-Distributionen	143
2.6.3.4	Das Charakterisierungstheorem	144
2.6.4	Feynman-Integrale und weißes Rauschen	149
2.7	Weitere Varianten	152
2.7.1	Promäße und Prodistributionen	153
2.7.2	Imaginäre Resolventen	155
2.7.3	Non-Standard-Analysis	156
2.7.4	Pfadintegrale im Phasenraum	157
2.8	Erkenntnistheoretische Einordnung der Pfadintegrale	160
2.9	Quantenmechanik und rotierende Zeiger	161

2.9.1	Zeigerformalismus und klassische Wellenoptik . . . . .	162
2.9.2	Zeigerformalismus und Quantenmechanik . . . . .	164
2.10	Zehn Argumente gegen den Zeigerformalismus . . . . .	169
2.10.1	Physikalische Aspekte . . . . .	169
2.10.1.1	Nichtrelativistische Quantenmechanik . . . . .	169
2.10.1.2	Klassische Näherung und Quantenmechanik . . . . .	170
2.10.1.3	Anschaulichkeit und Unanschaulichkeit . . . . .	171
2.10.1.4	Quantenmechanik und Quantenelektrodynamik . . . . .	173
2.10.2	Erkenntnistheoretische Aspekte . . . . .	174
2.10.2.1	Quantenmechanik als Naturbeschreibung . . . . .	174
2.10.2.2	Defizite der Feynmanschen Pfadintegrale . . . . .	175
2.10.3	Didaktische Aspekte . . . . .	176
2.10.3.1	Pfadintegrale für die Schule . . . . .	176
2.10.3.2	Klassische Physik durch die Hintertür . . . . .	177
2.10.3.3	Aufwand und Nutzen des Zeigerformalismus . . . . .	179
2.10.3.4	Blick über den Horizont . . . . .	181
<b>3</b>	<b>Entropie in Klasse 10</b>	<b>183</b>
3.1	Vorbemerkungen . . . . .	184
3.2	Notwendige Vorkenntnisse . . . . .	185
3.3	Der zweite Hauptsatz der Thermodynamik . . . . .	186
3.3.1	Entropie als Maß des inneren Durcheinanders . . . . .	186
3.3.2	Entropiedifferenzen . . . . .	189
3.3.3	Der zweite Hauptsatz . . . . .	190
3.4	Nutzbare und verlorene Arbeit . . . . .	192
3.4.1	Entropieübertragung und Entropieerzeugung . . . . .	192
3.4.2	Verlorene Arbeit . . . . .	194
3.4.3	Der Carnotsche Wirkungsgrad . . . . .	198
3.4.4	Beispiele . . . . .	201
3.5	Der dritte Hauptsatz der Thermodynamik . . . . .	203
3.6	Entropie in Technik und Natur . . . . .	204
3.6.1	Wärmepumpen . . . . .	204
3.6.2	Entropiebilanzen der Natur . . . . .	206
<b>4</b>	<b>Quantenmechanik in der Oberstufe</b>	<b>208</b>
4.1	Vorbemerkungen . . . . .	208
4.2	Axiomatische Basis der Quantenmechanik . . . . .	212
4.2.1	Die Axiome der Quantenmechanik . . . . .	212
4.2.2	Bemerkungen zu den Axiomen . . . . .	213
4.2.2.1	Zustände und Zustandsvektoren . . . . .	213
4.2.2.2	Observablen und Operatoren . . . . .	215
4.2.2.3	Meßwerte und Spektren . . . . .	216
4.2.2.4	Quantenmechanische Wahrscheinlichkeiten . . . . .	217
4.2.2.5	Das Projektionspostulat . . . . .	221
4.2.2.6	Zeitentwicklungsoperator und Schrödingergleichung . . . . .	222
4.3	Welle-Teilchen-Dualismus . . . . .	231

4.3.1	Das Komplementaritätsprinzip . . . . .	232
4.3.2	Welcher-Weg-Information und Sichtbarkeit von Interferenzmustern	236
4.3.2.1	Interferenzexperimente mit einzelnen Quantenobjekten .	236
4.3.2.2	Photonen im Mach-Zehnder-Interferometer . . . . .	238
4.3.2.3	Welcher-Weg-Information und Interferenzmuster . . . . .	241
4.3.2.4	Exakte Formulierung des Welle-Teilchen-Dualismus . . .	246
4.3.3	Englerts Dualitätsrelation . . . . .	249
4.3.3.1	Sichtbarkeit von Interferenzmustern . . . . .	249
4.3.3.2	Unterscheidbarkeit der Wege . . . . .	252
4.3.3.3	Beweis der Dualitätsrelation . . . . .	254
4.3.3.4	Verallgemeinerungen des Formalismus . . . . .	257
4.4	Notwendige Vorkenntnisse . . . . .	260
4.5	Klassische Physik und Quantenmechanik . . . . .	261
4.5.1	Rückblick: Klassische Physik . . . . .	261
4.5.2	Quantenmechanik als Neuformulierung . . . . .	263
4.5.2.1	Teilchen, die sich wie Wellen zu verhalten scheinen . . .	263
4.5.2.2	Wellen, die sich wie Teilchen zu verhalten scheinen . . .	266
4.5.3	Der Welle-Teilchen-Dualismus . . . . .	269
4.5.3.1	Experimente mit Zwei-Wege-Interferometern . . . . .	270
4.5.3.2	Das Prinzips vom Welle-Teilchen-Dualismus . . . . .	274
4.5.3.3	Anwendung: Wechselwirkungsfreie Messungen . . . . .	275
4.5.3.4	Vorüberlegungen zum Zustandsbegriff . . . . .	277
4.5.4	Elementare Axiomatik der Quantenmechanik . . . . .	280
4.5.4.1	Die Grundregeln der Quantenmechanik . . . . .	281
4.5.4.2	Interpretation und Bemerkungen . . . . .	281
4.5.4.3	Beispiele . . . . .	286
4.5.5	Die Heisenbergsche Unschärferelation . . . . .	292
4.5.5.1	Streuungen von Meßwerten . . . . .	293
4.5.5.2	Unschärferelation für Ort und Impuls . . . . .	294
4.5.5.3	Das Komplementaritätsprinzip . . . . .	296
4.5.6	Einfache quantenmechanische Systeme . . . . .	298
4.5.6.1	Beschreibung quantenmechanischer Systeme . . . . .	299
4.5.6.2	Der lineare Potentialtopf . . . . .	300
4.5.6.3	Das Wasserstoffatom . . . . .	304
4.5.6.4	Quantenmechanik schwererer Atome . . . . .	314
4.5.7	Superpositionsprinzip und verschränkte Zustände . . . . .	318
4.5.7.1	Lokalitätseigenschaften der Physik . . . . .	318
4.5.7.2	Das Superpositionsprinzip der Quantenmechanik . . . .	319
4.5.7.3	Verschränkte Zustände . . . . .	322
4.5.7.4	Verborgene Variable und Bellsche Ungleichungen . . . .	325

<b>Anhang</b>	<b>338</b>
<b>A Der thermodynamische Limes</b>	<b>339</b>
A.1 Thermodynamische Systeme . . . . .	339
A.2 Existenz des thermodynamischen Limes . . . . .	342
A.3 Die mikrokanonische Gesamtheit . . . . .	349
A.4 Die kanonische Gesamtheit . . . . .	352
A.5 Die großkanonische Gesamtheit . . . . .	355
<b>B Der Fock-Raum</b>	<b>358</b>
B.1 Definition und mathematische Eigenschaften . . . . .	358
B.2 Basissysteme für Fock-Räume . . . . .	360
<b>C Lösungen der Aufgaben</b>	<b>362</b>
<b>Literaturverzeichnis</b>	<b>381</b>





# Vorwort

Das vorliegende Buch entstand als Reaktion auf Entwicklungen in der gymnasialen Physikdidaktik, deren Anfänge bereits Jahrzehnte zurückliegen, deren Umsetzung jedoch erst seit einigen Jahren verstärkt betrieben wird. Dabei wird speziell in zwei Bereichen verbreitet massives Umdenken propagiert, einerseits im Zusammenhang mit Begriffen wie Energie und Entropie sowie der Bedeutung der Thermodynamik für die Schulphysik, andererseits in Bezug auf den Umgang mit der Quantenmechanik in der Oberstufe des Gymnasiums. Zwei Begriffe stehen hier im Mittelpunkt: Es geht um den Karlsruher Physikkurs und den Zeigerformalismus. Die Protagonisten dieser Konzepte fordern teilweise nicht weniger, als wohletablierte Zugänge zu physikalischen Inhalten aufzugeben oder zumindest stark zu modifizieren. Wir plädieren hier dafür, dieses Umdenken selbst zu überdenken und kommen zum Schluß, daß die herkömmlichen Zugänge keinesfalls vorschnell über Bord geworfen werden sollten.

Wir gehen davon aus, daß die Leserin oder der Leser normalerweise keine Anleitung zur Lektüre von Fachbüchern nötig hat; dennoch sind ein paar Hinweise womöglich hilfreich, da die Inhalte dieses Buches erheblich unterschiedliche Ausrichtungen aufweisen, insbesondere was den fachlichen Schwierigkeitsgrad angeht. Entsprechend kann es mit ganz unterschiedlichen Zielsetzungen gelesen werden. Die Idealvorstellung ist natürlich die Lektüre des gesamten Buches von der ersten bis zur letzten Seite; die erforderlichen Vorkenntnisse sind jedoch ebenso unterschiedlich wie die Inhalte der Abschnitte, sodaß auch einzelne Teile als eigene Einheiten betrachtet werden können.

Das gilt insbesondere für die Abschnitte 1.3 und 2.10, die jeweils für sich allein stehen können und eine Auflistung der gesammelten Argumente im Rahmen unserer kritischen Auseinandersetzung mit den beiden oben erwähnten vermeintlichen Aushängeschildern der gegenwärtigen Physikdidaktik darstellen. Sie können unabhängig vom Rest des Buches gelesen werden und erfordern lediglich Kenntnisse über Schulphysik, aber selbstverständlich ist das physikalische Darumherum der restlichen Abschnitte von Kapitel 1 und 2 nicht zuletzt auch dazu geeignet, die Argumente weiter zu untermauern. Die Abschnitte 1.2 auf der einen und 2.2 bis 2.8 auf der anderen Seite beschreiben den fachlichen Hintergrund des Karlsruher Physikkurses beziehungsweise des Zeigerformalismus in der Quantenmechanik. Gleichzeitig stellen die Abschnitte 2.1 bis 2.8 eine zwar geraffte, aber dafür streng formale und für sich völlig autonome Einführung in die Theorie der Pfadintegrale dar, soweit sie für die mathematische Physik von Interesse sind. Sie setzen Grundkenntnisse in mathematischer Physik voraus, die erheblich über das hinausgehen, was als Hintergrundwissen zum Unterrichten des Fachs Physik erforderlich ist; das zeigt aber umso deutlicher, daß zumindest angezweifelt werden darf, ob die entsprechenden Sachverhalte als Grundlage für die Entwicklung von Konzepten der Schulphysik tatsächlich brauchbar sind.

Da konstruktive Kritik reinem Gemotze stets vorzuziehen ist, beschränken wir uns natürlich nicht einfach auf letzteres, sondern zeigen in den Kapiteln 3 und 4 ausführlich, wie man es aus unserer Sicht besser machen kann. Diese beiden Kapitel enthalten mehrfach erprobte Unterrichtsgänge zur Einführung der Entropie in der Mittelstufe beziehungsweise der Quantenmechanik in der Oberstufe, wobei beiderseits konsequent all das vermieden wird, was in den Kapiteln zuvor moniert wurde. Insbesondere stehen unsere Unterrichtsgänge vollständig auf der Basis der etablierten Fachsprache, verzichten vollkommen auf die eigenmächtige Einführung unüblicher Begriffe und ermöglichen so bei Bedarf einen reibungslosen Übergang zur Hochschulphysik. Die Unterrichtsentwürfe können direkt als Handreichung für das Unterrichten der entsprechenden Lehrplaneinheiten in der neunten oder zehnten beziehungsweise zwölften Klasse des Gymnasiums verwendet werden. Aufgaben am Ende jedes Abschnittes der Unterrichtseinheiten unterstützen das zusätzlich.

Ein ausführlicher Anhang behandelt einerseits mit den mathematischen Grundlagen thermodynamischer Grenzübergänge und des Begriffs des Fock-Raumes zwei Gebiete, welche die eher formalen Abschnitte des Buches ergänzen und liefern andererseits die vollständigen und ausführlichen Lösungen zu allen Aufgaben, die in der Beschreibung der beiden Unterrichtsgänge aufgeführt sind.

Mit dem vorliegenden Buch verfolgen wir zwei Zielsetzungen, eine sehr optimistische und eine eher realistische. Die optimistische Zielsetzung besteht darin, eine Entwicklung, die in der gymnasialen Physikdidaktik seit Jahren mit zunehmender Intensität zu beobachten ist und nach unserer Ansicht in die falsche Richtung geht, zu bremsen und umzudirigieren zugunsten eines Festhaltens an bewährten Konzepten, was die Thermodynamik, und zugunsten der Abkehr von Fehlinterpretationen, was die Quantenmechanik angeht. Vielleicht kann auch dazu beigetragen werden, daß die Hochschulphysik, besonders auch, was ihre formalen und mathematischen Aspekte betrifft, natürlich nicht im Unterricht selbst, aber umso mehr beim Nachdenken über und Planen und Weiterentwickeln von Unterrichtskonzepten, verstärkt Berücksichtigung findet. Die realistische Zielsetzung ist das Anstoßen einer entsprechenden Diskussion, die gegenwärtig viel zu wenig stattfindet, sicherlich auch aufgrund des traditionell zurückhaltenden Gedankenaustauschs zwischen Fachphysikern und Fachdidaktikern. Im Idealfall erweisen sich beide Zielsetzungen als identisch.

E-Mail-Adresse des Autors: Vogt.Markus@t-online.de

# Einleitung

Die Schulphysik versucht seit eh und je im Hinblick auf ihr Verhältnis zur Physik als Wissenschaft, einen eigenen Weg zu beschreiten. Das mag in mancherlei Hinsicht berechtigt und aus psychologischer Sicht vielleicht auch verständlich sein und ist so lange unbedenklich, wie sich dieses Phänomen auf Details wie die Auswahl von Variablennamen, einführende Beispiele oder Anwendungsmöglichkeiten beschränkt. Seit einigen Jahren gibt es jedoch starke Bewegungen, diesen Sonderweg ganz andere Ausmaße annehmen zu lassen, die überdies mit merkwürdigen Abschottungstendenzen verbunden sind und kurz gesagt darauf hinauslaufen, alles ganz anders zu machen. Man kann sich schwerlich an der Vermutung hindern, es solle der Eindruck der Existenz zweier getrennter Fächer erweckt werden, der Schul- und der Hochschulphysik, und das keineswegs nur im Hinblick auf das fachliche und formale Niveau. Vielmehr findet man in ersterer inzwischen verbreitet Zugänge, Begriffe und Erklärungen, die letzterer im günstigeren Fall völlig fremd sind, im ungünstigeren ihr sogar widersprechen.

Zwei didaktische Modelle stehen im Zentrum dieser Entwicklung. Sie sind völlig unabhängig voneinander, unterscheiden sich insbesondere in Bezug auf die von ihnen jeweils beanspruchten Zuständigkeitsbereiche und haben genau besehen nur die Gemeinsamkeit, in der Physik des Gymnasiums inzwischen offensichtlich verbreitet für unverzichtbar gehalten zu werden. Das eine Modell tritt üblicherweise unter der Bezeichnung *Karlsruher Physikkurs* auf. Gegenstand dieses Konzepts ist eigentlich nur die elementare phänomenologische Gleichgewichtsthermodynamik, es erhebt jedoch gleichzeitig den Anspruch, die Grundlage der gesamten Physik darzustellen. Entsprechend soll die gesamte Schulphysik vom Anfangsunterricht bis zum Abitur nach diesem Muster aufgezogen werden. Das andere Modell versucht, unter Verwendung des sogenannten *Zeigerformalismus* einen für die Schulphysik gangbaren Zugang zur Quantenmechanik zu finden. Der Allgemeinheitsanspruch des Karlsruher Physikkurses ist diesem Konzept zwar fremd, es hat mit jenem aber zumindest die Eigenschaft gemein, auf breiter Ebene Begriffe zu verwenden, die nicht nur von der Terminologie der traditionellen Schulphysik, sondern vor allen Dingen von derjenigen der Hochschulphysik zum Teil sehr stark abweichen. Begründet wird das unter anderem mit dem Ziel, Lernschwierigkeiten abzubauen – was ein wenig den Vergleich mit einem Sprachkurs aufdrängt, bei dem die echten Vokabeln als zu schwer eingestuft und folglich durch andere, selbsterfundene ersetzt werden. Das mag tendenziös klingen, ist aber geeignet, eine Vorstellung von der Tagweite solcher Konzepte zu vermitteln.

Gleichzeitig illustriert dieser Vergleich ein zentrales Prinzip der Physikdidaktik, das nach Auffassung des Autors allen anderen denkbaren solchen Prinzipien übergeordnet ist: Grundlage jeglicher Unterrichtsplanung muß die etablierte Hochschulphysik sein. Das

bedeutet – man behalte den Vergleich mit den Vokabeln im Sinn – einen völligen Verzicht der Schulphysik auf begriffliche Eigenzuchtungen und damit das Vermeiden einer vom Standard abweichenden Fachsprache genauso wie die Beschränkung auf Zugänge zu und Modelle von physikalischen Gesetzmäßigkeiten, die sich in entsprechend subtilerer Form in der Hochschulphysik wiederfinden oder zumindest mit deren Zugängen und Modellen in Einklang bringen lassen.

Das alles wird von den oben erwähnten Unterrichtskonzepten konsequent ignoriert. Was den Karlsruher Physikkurs angeht, liegt das erheblich an einer didaktischen Strategie, die mit dem Schlagwort „kumulatives Lernen“ umschrieben wird und Begriffsbildungen im Physikunterricht unter Verzicht auf klare Definitionen nach und nach durch Zusammentragen von Eigenschaften der betrachteten physikalischen Größe bewerkstelligen will. Dabei soll von den Vorstellungen ausgegangen werden, die von den Schülerinnen und Schülern schon mitgebracht werden. Abgesehen von der Frage, inwieweit diese mitgebrachten Vorstellungen tatsächlich mitgebracht werden und sich überhaupt einheitlich formulieren lassen, muß einerseits damit gerechnet werden, daß sich diese Vorstellungen auch dann verfestigen, wenn sie völlig falsch sind, da an keiner Stelle der notwendige Bruch zwischen Alltagsmythen und tatsächlichen physikalischen Gesetzmäßigkeiten provoziert wird, und andererseits fällt dabei ein wesentliches oder vielleicht das wichtigste Bildungsziel des Physikunterrichts unter den Tisch, nämlich das Kennenlernen dieses Faches als einer mathematisch orientierten Naturwissenschaft, die das Berechnen quantitativer Resultate erlaubt. Beim Zeigerformalismus sind die Ursachen zwar nicht derart didaktisch tiefschürfend, dafür aber nicht weniger folgenschwer. Hauptbeweggrund ist der Wunsch, quantitative Berechnungen anzustellen, wo, wie man gleichzeitig feststellt, komplexe Zahlen eigentlich unentbehrlich sind. Da man letztere in der Schule jedoch nicht zur Verfügung hat, verwendet man einen Hilfsformalismus, der auf den ersten Blick harmlos erscheint, sich auf den zweiten jedoch als enorm aufwendig und ohne Computereinsatz kaum beherrschbar erweist und zudem Resultate erbringt, die, wiewohl weitgehend fachlich korrekt, entweder für das Gymnasium überflüssig oder aber anderweitig viel einfacher beschaffbar sind. Insbesondere aber taucht dieser Formalismus in der hier zur Diskussion stehenden Form fachwissenschaftlich an keiner Stelle auf, so daß auch hier die Schülerinnen und Schüler auf einen Weg gebracht werden, der nicht, wie es eigentlich zu wünschen wäre, ein wenig in Richtung der etablierten Fachwissenschaft führt, sondern weit abseits davon.

Wir wir sehen werden, liegt darin jeweils eine wesentliche, wenngleich keineswegs die einzige Schwachstelle sowohl des Karlsruher Physikkurses als auch des Zeigerformalismus. Sie führt beiderseits dazu, daß Schülerinnen und Schüler, die auf diese Weise unterrichtet wurden, eine Sprache lernen, von der sie nichts in der wirklichen Physik wiederfinden; sie haben, wenn man sie nichts anderes kennenlernen lassen hat, kaum eine Chance, zur traditionellen und damit insbesondere für alle weiterführenden Wege unverzichtbaren Physik zurückzufinden. Neben dieser gemeinsamen Problematik weisen die beiden Konzepte sehr unterschiedliche, teilweise sogar diametral entgegengesetzte Ansatzpunkte für Kritik auf. Letzteres bildet einen wesentlichen Bestandteil des vorliegenden Buches und läßt sich vorab kurz wie folgt skizzieren:

Grundlage des Karlsruher Physikkurses ist die Gleichgewichtsthermodynamik und damit eine physikalische Disziplin, die nicht nur wohletabliert und -verstanden, sondern inzwischen auch weitgehend mathematisch streng formulierbar ist. Diese wird jedoch

einerseits fehlerhaft und andererseits in einer Breite angewendet, die weit über ihren tatsächlichen Zuständigkeitsbereich hinausgeht, mit entsprechenden Folgen für die Vorstellungen, die dabei bei Schülerinnen und Schülern ausgelöst werden. Grundlage des Zeigerformalismus ist die Feynmansche Pfadintegralformulierung der Quantenmechanik, also ein Konzept, das nicht nur sehr viel anspruchsvoller als die anderen elementaren Zugänge zur Quantenmechanik ist, sondern eine strenge mathematische Formulierung gerade nicht erlaubt. Pfadintegrale sind fiktive Objekte, mit denen Geübte zwar virtuos rechnen können, die aber genau besehen nicht existieren, eine Nachteil, der erst in der Quantenfeldtheorie und damit weit jenseits jeglicher Schulphysik vollständig reparierbar ist. Damit darf der Pfadintegral-Zugang zumindest im Hinblick auf die nichtrelativistische Quantenmechanik nicht als naturphilosophisch fundamentale Theorie angesehen werden. Das ist ein Schwachpunkt, der auch durch die Tatsache, daß der Zeigerformalismus im Gegensatz zum Karlsruher Physikkurs weitgehend fachlich korrekt arbeitet und seine engen Zuständigkeitsgrenzen kaum überschreitet, keineswegs aufgewogen werden kann.

Kurz gesagt arbeitet der Karlsruher Physikkurs fehlerhaft und dessen zulässige Grenzen weit überschreitend mit einem mathematisch streng formulierbaren Formalismus, während der Zeigerformalismus ein Konzept, das streng mathematisch gar nicht existiert, immerhin überwiegend richtig und mit der gebotenen Zurückhaltung in Bezug auf die vorgebliche Gültigkeitsspanne anwendet. Entsprechend unterschiedlich detailliert ist die fachliche Vorbetrachtung, die wir unserer Kritik voranstellen. Beim Karlsruher Physikkurs können wir uns auf eine elementare theoretische Betrachtung der thermodynamischen, statistischen und kontinuumsmechanischen Grundlagen beschränken; der Vollständigkeit halber befindet sich eine kurze Einführung in die Thermodynamik aus der Sicht der mathematischen Physik im Anhang. Das genügt, um die anschließend aufgeführten Argumente zu untermauern. Vor der Kritik des Zeigerformalismus dagegen ist ein ausführlicher Ausflug in die mathematische Physik angemessen. Nur so wird wirklich deutlich, was vermutlich unbemerkt von den meisten seiner Protagonisten und allen seiner Anhänger hinter dem Zeigerformalismus steckt, warum das mathematisch nicht funktionieren kann und wie extrem aufwendig sämtliche Versuche zur Konsolidierung dieses Problems sind, die noch dazu ausnahmslos nur für sehr spezielle Systeme und niemals in voller Allgemeinheit funktionieren. Die ausführlich kommentierten Listen unserer Gegenargumente sind nach unserer Auffassung auch für sich allein schon überzeugend; mit dem hier gelieferten mathematischen Hintergrund betrachten wir sie jedoch als noch wesentlich fundierter.

Diese Gegenargumente liefern jeweils bereits detaillierte Hinweise, wie man es anders machen sollte. Was läge somit näher als diese sogleich umzusetzen? Das vorliegende Buch soll auch von Nutzen für die alltägliche Unterrichtsarbeit sein, daher beschreiben wir zusätzlich ausführlich unsere Vorstellung von einem Unterricht der hier zur Diskussion stehenden physikalischen Inhalte, wie er den Anforderungen, die wir vom Karlsruher Physikkurs und vom Zeigerformalismus als nicht erfüllt betrachten, standhalten kann. Die vorgestellten Unterrichtseinheiten zur Thermodynamik und zur Quantenmechanik können zwar abgesehen davon, daß sie sich seit Jahren im Unterrichtsalltag bewähren, keine empirischen Bewertungen, Evaluationen oder dergleichen vorweisen, sie haben jedoch zumindest die Eigenschaft, fachlich absolut hieb- und stichfest zu sein. Damit soll nebenbei erneut verdeutlicht werden, woran sich Physikunterricht nach unserer Auffas-

sung stets orientieren sollte, nämlich an den Maßstäben und Begriffen der etablierten Wissenschaft. Daß das möglich ist, ohne mit Niveau und Schwierigkeitsgrad meilenweit über das Ziel hinauszuschießen, dürfte anhand der Unterrichtsgänge ebenfalls erkennbar sein.

Über all den Diskussionen über Zugänge und Begriffsbildungen im Schulunterricht sollte eines jedoch nicht zu kurz kommen: Das intellektuelle Vergnügen, welches die Physik bereitet. Es ist ganz unzweifelhaft ein solches, sich mit den Gesetzen der theoretischen Physik und den zugehörigen axiomatisch strengen Beweisen der mathematischen Physik zu beschäftigen, und auch die Auseinandersetzung mit anderen didaktischen Vorstellungen, gleichgültig wie abweichend von den eigenen sie sind, darf nicht zuletzt auch aus dieser Perspektive gesehen werden. Es wäre schön, wenn das vorliegende Buch dazu beiträgt, daß Leserinnen und Leser diese Einstellung teilen.

# Kapitel 1

## Bemerkungen zum Karlsruher Physikkurs

### 1.1 Vorbemerkungen

Gegenwärtig drängt sich unbeteiligten Beobachtern der Verdacht auf, daß der sogenannte Karlsruher Physikkurs im gymnasialen Physikunterricht mit Gewalt durchgesetzt werden soll. Selbst wenn man von Aspekten der didaktischen Freiheit, die in vernünftigem Ausmaß eigentlich nicht zur Diskussion stehen sollte, einmal völlig absieht, muß eine solche Entwicklung Widerspruch herausfordern.

Grundlage des Karlsruher Physikkurses ist eine von G. Falk vorgelegte Darstellung der theoretischen Physik, die stark thermodynamisch orientiert ist und auf der Dynamik der sogenannten extensiven Größen makroskopischer Systeme aufbaut [155], [156]. Das Konzept wurde darauf aufbauend an der Universität Karlsruhe von Falk und F. Herrmann gegründet [158]; die selbe Idee veröffentlichte etwa gleichzeitig A. A. diSessa am MIT [124]. Seither wurde es stetig weiterentwickelt<sup>1</sup>. Ausgangspunkt ist dabei entsprechend der Vorlage die Betrachtung mengenartiger und damit bilanzierbarer Größen, die sich in allen Bereichen der Physik finden, und eine darauf aufbauende streng analoge Vorgehensweise in den unterschiedlichen Teilgebieten. Im Zentrum der Betrachtung stehen dabei zeitliche Änderungen solcher mengenartiger Größen und deren Übergänge über Systemgrenzen hinweg, was als „Strömungen“ gedeutet wird. In diesem Sinn erfolgt dann jeweils die Einführung von Strömen und Stromstärken, welche die entsprechenden herkömmlichen Begriffe ersetzen sollen<sup>2</sup>.

Außerdem wird der Begriff des Energietransports in den Mittelpunkt gestellt, oder

---

<sup>1</sup>Hinweise auf den Karlsruher Physikkurs finden sich auch im Baden-Württemberger Lehrplan für allgemeinbildende Gymnasien, in der Fassung von 1994 nur an einer Stelle und ausdrücklich nur als mögliche Alternative, in derjenigen von 2004 schon etwas deutlicher, dafür aber undifferenziert und diffus.

<sup>2</sup>Gelegentlich vergleicht man den Karlsruher Physikkurs mit Blick auf seine von der traditionellen Terminologie abweichenden Sprechweise mit der von David Hestenes als neue Grundlage für Mathematik und Physik propagierten sogenannten geometrischen Algebra [236], [238]. Das ist nicht gerechtfertigt, denn Hestenes erklärt zwar gewissermaßen mit Hilfe von Clifford-Algebren die ganze Welt, er startet dabei jedoch mit Konzepten der relativistischen Quantenfeldtheorie und steht damit stets fest auf der Basis der etablierten Mikrophysik. Er verwendet lediglich eine besondere mathematische Sprache, versucht aber an keiner Stelle, das begriffliche Gerüst der Physik umzukrempeln.



genauer gesagt die Aussage, daß Energietransport nie allein, sondern stets zusammen mit einer zweiten physikalischen Größe erfolgt, die dann generell als „Energieträger“ bezeichnet wird. Wie oben angedeutet wird dann von Energieströmen und der Energiestromstärke geredet, wobei sich hinter letzterer nichts anderes als die Leistung verbirgt<sup>3</sup>. Um die Sache anschaulich zu halten, wird das Konzept zunächst ausführlich am Beispiel von Wasserstromkreisen eingeführt. Analogien hierzu werden nun in allen anderen Bereichen gesucht und gefunden. So wird beispielsweise in der Mechanik mit dem Impuls angefangen, dieser als Energieträger erkannt, und Kräfte werden dann als Impulsstromstärken eingeführt. In der Thermodynamik muß entsprechend die Entropie sehr frühzeitig eingeführt werden, um eine als Energieträger verwendbare Größe zur Verfügung zu haben [415]. Ähnlich geht man mit dem Drehimpuls oder in der Elektrizitätslehre<sup>4</sup> und in sonstigen Teilgebieten der Physik vor.

Selbst wenn wir fachliche Aspekte für den Moment außer Acht lassen, kommen wir nicht umhin, die Konsequenzen eines derart umgekrempten Physikunterrichts als fatal einzuschätzen. Das Problem liegt unter anderem in einer extremen Überstrapazierung der Verwendung von Analogien und dem Gebrauch von Begriffen, deren Abstraktionsgrad deutlich über das in der Schule vernünftige Maß hinausgeht. So ist etwa das Energieträger-Konzept im Rahmen der Thermodynamik innerhalb gewisser enger Grenzen tragfähig; die physikalischen Größen, die als Energieträger herangezogen werden, sind auch gelegentlich recht anschaulich (Licht, elektrischer Strom, Wasser, ...), sehr oft jedoch äußerst abstrakt (Impuls, Drehimpuls, Entropie, ...), so daß die damit gebildeten Analogien eben gerade nicht die gewünschte Erklärungskraft entfalten, sondern bestenfalls zu einem Scheinverständnis führen dürften. Inwiefern beispielsweise die Einführung von Kräften als Impulsstromstärken Lernschwierigkeiten beseitigen helfen soll, ist schwer nachvollziehbar. Ähnliches gilt für die Verwendung der Entropie in diesem Zusammenhang, was die Schülerinnen und Schüler dazu bringt, Entropie für dasselbe wie Wärme zu halten<sup>5</sup>. Und hierin zeichnet sich ein weiteres zentrales Problem ab: Die Schülerinnen und Schüler werden mit dem Karlsruher Physikkurs auf einen begrifflichen Weg gebracht, der sie weit von den etablierten Vorstellungen wegführt, eine Entwicklung, die schwer rückgängig zu machen sein dürfte.

Bevor wir unsere Einwände präzisieren, beschäftigen wir uns kurz mit den Grundlagen

---

<sup>3</sup>Diese Idee ist nicht neu. Ein entsprechendes Konzept wurde bereits 1898 von Gustav Mie veröffentlicht [366], allerdings selbstverständlich nicht mit der Absicht, daraus eine Didaktik der Schulphysik zu fabrizieren.

<sup>4</sup>Die Einführung einer Substanz namens „Elektronium“ als elektrischem Grundstoff, aus dem etwa auch die Atomhüllen aufgebaut und deren „Quanten“ dann die Elektronen sein sollen, ist ein besonders abwegiges Beispiel der Kreation neuer, der Fachwelt unbekannter Begriffe, auf das noch zurückzukommen sein wird.

<sup>5</sup>Es mag zwar unglaublich klingen, aber es existieren inzwischen tatsächlich ernstgemeinte Vorschläge, im Unterricht Entropie mit Wärme gleichzusetzen [345], also objektiv falsche Aussagen zu vermitteln. Diese Position wird bei Fortbildungsveranstaltungen offen vertreten. Im Übrigen stand die Gleichsetzung von Entropie und Wärme im Mittelpunkt der undifferenzierten und einseitigen Berichterstattung, die in der Tagespresse im Jahr 2004 zu lesen war [13], [248], [314]; dort wird der Karlsruher Physikkurs völlig kritiklos und fachlich unqualifiziert als alternativlose neue Physikdidaktik und Lösung aller Probleme des Physikunterrichts gefeiert, Umdenken und -lernen der Physiklehrerinnen und Physiklehrer gefordert und deren mangelnde Bereitschaft hierzu beklagt. Die diametral falsche Aussage zur Entropie, mit der in [13] eine Schülerin zitiert wird, mag auf ein Versehen der Autorin zurückzuführen sein, sie paßt gleichwohl hervorragend zur Lage der Dinge.

des Karlsruher Physikkurses aus der Sicht der theoretischen Physik. Es werden sich dabei bereits erste Hinweise auf die Problematik dieses Konzepts ergeben.

## 1.2 Physikalische Vorbetrachtungen

Die Strategie des Karlsruher Physikkurses besteht im wesentlichen darin, die Thermodynamik zu *der* fundamentalen Disziplin der Physik überhaupt zu erklären. Wenn wir das zunächst einmal unkommentiert stehen lassen, so erscheint es in jedem Fall sinnvoll, diejenigen physikalischen Begriffe und Gesetze, die für diese Strategie von Bedeutung sind, kurz zu rekapitulieren. Wir setzen dabei voraus, daß der Leser mit den Grundlagen der Mechanik, der Thermodynamik und der Statistik vertraut ist. Der Abschnitt kann jedoch auch ohne Nachteile für das Verständnis des restlichen Kapitels überschlagen werden.

### 1.2.1 Energie und der Energieerhaltungssatz

Der übliche umgangssprachliche Gebrauch des Begriffs der *Energie* täuscht schnell darüber hinweg, daß es sich aus physikalischer Sicht hierbei um ein sehr abstraktes Konzept zur Beschreibung des Zustands physikalischer Systeme handelt. Ausgangspunkt zur Definition dieses Begriffs ist das *Noethersche Theorem* [397], wonach jede Symmetrie der Lagrange-Dichte eines Systems auf einen Erhaltungssatz einer bestimmten Größe führt<sup>6</sup>. Im vorliegenden Zusammenhang ist die Zeittranslationssymmetrie von Interesse, also salopp gesagt die Tatsache, daß es aus der Sicht der Physik eigentlich egal sein sollte, wann ein Experiment durchgeführt wird. Man erkennt das am besten in relativistisch kovarianter Schreibweise.

Dazu betrachten wir ein System mit Lagrange-Dichte

$$\mathcal{L} = \mathcal{L} \left( \psi_1, \psi_2, \dots, \psi_N, \frac{\partial \psi_1}{\partial x^\mu}, \frac{\partial \psi_2}{\partial x^\mu}, \dots, \frac{\partial \psi_N}{\partial x^\mu} \right),$$

wobei die  $\psi_A$  irgendwelche Felder sein sollen; wir verwenden im folgenden die Einsteinsche Summenkonvention und schreiben kurz  $\partial A / \partial x^\mu \equiv A_{|\mu}$  sowie  $\mathcal{L} = \mathcal{L}(\psi_A, \psi_{A|\mu})$ . Die Lagrange-Dichte sei invariant unter Raum-Zeit-Translationen  $x^\mu \rightarrow x^\mu + a^\mu$ . Für eine infinitesimale Translation  $x^\mu \rightarrow x^\mu + \varepsilon^\mu$ , wobei  $\varepsilon^\mu$  klein sein soll, gilt wegen  $\delta x^\mu = \varepsilon^\mu$  für die Felder und die Lagrange-Dichte zunächst

$$\begin{aligned} \delta \psi_A &= \varepsilon^\mu \psi_{A|\mu}, \\ \delta \psi_{A|\mu} &= \varepsilon^\nu \psi_{A|\mu|\nu}, \\ \delta \mathcal{L} &= \varepsilon^\mu \mathcal{L}_{|\mu}; \end{aligned}$$

andererseits gilt für die Lagrange-Dichte aber auch

$$\begin{aligned} \delta \mathcal{L} &= \frac{\partial \mathcal{L}}{\partial \psi_A} \delta \psi_A + \frac{\partial \mathcal{L}}{\partial \psi_{A|\mu}} \delta \psi_{A|\mu} \\ &= \frac{\partial \mathcal{L}}{\partial \psi_A} \varepsilon^\mu \psi_{A|\mu} + \frac{\partial \mathcal{L}}{\partial \psi_{A|\mu}} \varepsilon^\nu \psi_{A|\mu|\nu}, \end{aligned}$$

---

<sup>6</sup>Für Details siehe zum Beispiel [135].

was mit Hilfe der Euler-Lagrange-Gleichungen

$$\frac{\partial \mathcal{L}}{\partial \psi_A} - \left( \frac{\partial \mathcal{L}}{\partial \psi_{A|\mu}} \right)_{|\mu} = 0$$

umgeformt werden kann zu

$$\begin{aligned} \delta \mathcal{L} &= \varepsilon^\nu \left[ \left( \frac{\partial \mathcal{L}}{\partial \psi_{A|\mu}} \right)_{|\mu} \psi_{A|\nu} + \frac{\partial \mathcal{L}}{\partial \psi_{A|\mu}} \psi_{A|\mu|\nu} \right] \\ &= \varepsilon^\nu \left( \frac{\partial \mathcal{L}}{\partial \psi_{A|\mu}} \psi_{A|\nu} \right)_{|\mu}. \end{aligned}$$

Zusammengenommen findet man also

$$\left( \mathcal{L} \delta^\mu{}_\nu - \frac{\partial \mathcal{L}}{\partial \psi_{A|\mu}} \psi_{A|\nu} \right)_{|\mu} \varepsilon^\nu = 0,$$

das heißt einen differentiellen Erhaltungssatz der Form

$$T^\mu{}_{\nu|\mu} = 0 \tag{1.1}$$

mit dem kanonischen Energie-Impuls-Tensor

$$T^\mu{}_\nu = \mathcal{L} \delta^\mu{}_\nu - \frac{\partial \mathcal{L}}{\partial \psi_{A|\mu}} \psi_{A|\nu}.$$

Integriert man über (1.1) und verwendet den Gaußschen Satz, so läßt sich daraus ein integraler Erhaltungssatz herleiten. Es gilt

$$\begin{aligned} \iiint T^\mu{}_{\nu|\mu} d^3x &= \iiint T_0^\mu{}_{|0} d^3x + \iiint T_i^\mu{}_{|i} d^3x \\ &= \frac{d}{dt} \iiint T_0^\mu d^3x + \oint T_i^\mu dS_i = 0 \end{aligned}$$

(griechische Indizes laufen von 0 bis 3, lateinische Indizes von 1 bis 3), und da die Felder  $\psi_A$  im Unendlichen als hinreichend schnell verschwindend angenommen werden dürfen, fällt der Oberflächenterm weg. Der daraus folgende integrale Erhaltungssatz lautet

$$\frac{d}{dt} P^\mu = 0. \tag{1.2}$$

Die hier auftretende Größe

$$P^\mu \equiv \iiint T_0^\mu d^3x$$

ist also eine Erhaltungsgröße und heißt *4-Impuls*; ihre drei räumlichen Komponenten nennt man Komponenten des relativistischen Impulses, die zeitliche Komponente

$$E \equiv \iiint T_0^0 d^3x$$

erhält die Bezeichnung *Energie*. (1.2) ist gerade der Energie-Impuls-Erhaltungssatz, der sich damit als eine Folge der Invarianz der Lagrange-Dichte unter Raum-Zeit-Translationen erweist. Damit haben wir die allgemeinste Definition der Energie:

*Energie ist diejenige Erhaltungsgröße, die aus der Zeittranslationssymmetrie eines physikalischen Systems folgt.*

Diese Definition hat allerdings den Nachteil, sehr formal zu sein. Im nächsten Abschnitt folgt eine operationale und damit anschaulichere Variante.

### 1.2.2 Die physikalische Arbeit

Auch der Begriff der *Arbeit* ist aus der Umgangssprache bekannt und erfährt in der Physik eine Umdeutung; allerdings ist der Abstraktionsgrad hier bei weitem nicht so hoch. Wir kehren zur Definition dieser Größe von der relativistischen zur gewöhnlichen Schreibweise zurück und betrachten eine Kraft, die durch ein Vektorfeld  $\vec{F}(\vec{r})$  beschrieben werden soll. Wird ein Körper, der mit diesem Feld wechselwirkt, in demselben entlang eines Weges  $\Gamma$  bewegt, so ist die dabei verrichtete beziehungsweise freiwerdende Arbeit definiert durch das Kurvenintegral

$$W = \int_{\Gamma} \vec{F} d\vec{r}.$$

Anschaulich läßt sich das interpretieren, indem man sagt, daß an einem System Arbeit verrichtet wird, wenn man es gegen eine auf dieses wirkende Kraft verschiebt; umgekehrt kann das System selbst Arbeit verrichten, wenn es sich in Richtung einer auf sich wirkenden Kraft bewegt. Ist  $\vec{F}$  auf ein möglicherweise geschwindigkeitsabhängiges Potential

$$U = U(\vec{r}, \dot{\vec{r}}) = V(\vec{r}) + D(\dot{\vec{r}})$$

zurückführbar gemäß

$$F_j = \frac{\partial U}{\partial r_j} + \frac{d}{dt} \frac{\partial U}{\partial \dot{r}_j} = \frac{\partial V}{\partial r_j} + \frac{d}{dt} \frac{\partial D}{\partial \dot{r}_j},$$

wobei von der vereinfachenden, aber häufig in guter Näherung zutreffenden Annahme ausgegangen wird, daß sich das Potential aus einem konservativen und einem geschwindigkeitsabhängigen Anteil zusammensetzt, dann folgt

$$W = \int_{\Gamma} \sum_j \left( \frac{\partial V}{\partial r_j} + \frac{d}{dt} \frac{\partial D}{\partial \dot{r}_j} \right) dr_j = \int dV + \int_{\Gamma} \sum_j \frac{d}{dt} \frac{\partial D}{\partial \dot{r}_j} dr_j = V + \frac{dB}{dt}.$$

Die Größe

$$B = \int_{\Gamma} \sum_j \frac{\partial D}{\partial \dot{r}_j} dr_j$$

hat die Dimension einer Wirkung und ist ein Maß dafür, wie stark das System energetisch an die Umgebung gekoppelt ist, beispielsweise durch dissipative Prozesse; wenn  $B$  zeitlich konstant ist, wird keine Energie an die Umgebung abgegeben oder aus dieser aufgenommen. In diesem Fall ist das System rein konservativ, und es gilt  $\Delta W = \Delta V$ , das heißt, es wird Arbeit verrichtet, um die potentielle Energie des Systems zu vergrößern, oder letztere nimmt ab, und das System verrichtet selbst Arbeit.

Ein wichtiger Sonderfall ist durch das geschwindigkeitsabhängige „Potential“

$$D = \sum \frac{1}{2} m_j \dot{r}_j^2$$

gegeben. In diesem Fall gilt

$$F_j = \frac{\partial V}{\partial r_j} + m_j \ddot{r}_j$$

und somit

$$W = V + \int_{\Gamma} \sum_j m_j \ddot{r}_j dr_j = V + \int_{t_0}^t \sum_j \ddot{r}_j m_j \dot{r}_j dt' = V + \sum_j \frac{1}{2} m_j (v_j^2 - v_{j0}^2).$$

Das ist die Standard-Situation der Mechanik; die an einem System verrichtete Arbeit wird beispielsweise dazu verwendet, die potentielle und die kinetische Energie desselben zu vergrößern<sup>7</sup>.

Arbeit wird also genau dann verrichtet, wenn Energie unter der Wirkung einer Kraft über eine Systemgrenze hinweg übertragen wird und dabei im allgemeinen in eine andere Form umgewandelt wird<sup>8</sup>. Wird an einem System Arbeit verrichtet, so nimmt dessen Energie zu, verrichtet das System selbst Arbeit, nimmt seine Energie ab. Das läßt sich auch folgendermaßen formulieren:

*Energie ist die Fähigkeit eines Systems, Arbeit zu verrichten.*

Dabei ist zu beachten, daß es sich hierbei zunächst nicht um eine thermodynamische Aussage handelt. Wird die Thermodynamik, und das heißt nichts anderes als die Komplexität makroskopischer Systeme, mitberücksichtigt, wird klar, daß es sich dabei keineswegs vollständig um technisch nutzbare Arbeit handeln muß<sup>9</sup>. Im allgemeinen findet ein erheblicher Teil der verrichteten Arbeit auf mikroskopischer Ebene und noch dazu in völlig ungeordneter Art und Weise und folglich absolut unkontrollierbar statt<sup>10</sup>. Führt man den Begriff der Wärme als über eine Systemgrenze hinweg transportierte thermische Energie ein, läßt sich das besonders einfach ausdrücken. Man hat dann zwei Arten von Arbeit zu unterscheiden: *Mechanische Arbeit* und *thermische Arbeit*, also Wärme. Erstere ist Energieübertragung *ohne* Entropieerhöhung, letztere ist Energieübertragung *mit* Entropieerhöhung.

### 1.2.3 Das Boltzmannsche Prinzip und die statistische Definition der Entropie

Im Gegensatz zur Energie ist die Entropie kein aus der Umgangssprache geläufiger Begriff. Eine einigermaßen anschauliche Vorstellung gewinnt man durch deren von Boltzmann entdeckte statistische Deutung [71]; diese ist im Bereich der Betrachtung von Systemen im Gleichgewicht zur auf Clausius zurückgehenden thermodynamischen Entropie äquivalent, aber darüberhinausgehend sowie insbesondere aus naturphilosophischer Sicht

<sup>7</sup>Natürlich können Zu- und Abnahme der beteiligten Energieformen in beliebiger Kombination auftreten.

<sup>8</sup>Man sagt daher auch, Arbeit sei eine *Prozeßgröße*. Weitere Beispiele für Prozeßgrößen sind Wärme und Entropie, wie wir noch sehen werden. Im Gegensatz dazu ist Energie eine *Systemgröße*, genauer gesagt sogar eine *Zustandsgröße*. Näheres dazu steht in Abschnitt 1.2.4.1.

<sup>9</sup>Im Gegensatz zum Wunschenken mancher Physikdidaktiker kümmert sich die Physik keineswegs um ihre technische Anwendbarkeit und ist insbesondere auch nicht primär für letztere da.

<sup>10</sup>Das ist genau das, was mit dem populären Begriff der Energieentwertung gemeint ist.

fundamentaler als letztere. Die Thermodynamik hat in ihrer Eigenschaft, auf makroskopische Größen beschränkt zu sein, rein phänomenologischen Charakter, im Gegensatz zur statistischen Mechanik, auf die sie weitgehend zurückführbar ist. Machen wir also zunächst einen kurzen Abstecher dorthin.

In der statistischen Mechanik werden im allgemeinen Systeme mit sehr vielen Freiheitsgraden betrachtet, typischerweise  $10^{23}$  und mehr. Betrachten wir ein klassisches System aus  $N$  identischen Einzelsystemen mit jeweils  $f$  Freiheitsgraden, so ist der Phasenraum dieser Einzelsysteme  $2f$ -dimensional, und jeder mögliche Zustand eines Einzelsystems entspricht darin einem Punkt. Zur statistischen Beschreibung unterteilt man diesen Phasenraum in kleine jeweils gleich große Bereiche oder *Elementarzellen*, mathematisch gesehen in unendlich viele. Beschränkt auf den tatsächlich besetzten Bereich ist ihre Anzahl eine endliche, jedoch gewöhnlich unvorstellbar große Zahl. Entsprechend ist eine noch unvorstellbar viel größere Anzahl verschiedener Verteilungen der Einzelsysteme auf diese Zellen denkbar. Jede dieser möglichen Verteilungen nennt man einen *Mikrozustand* des Gesamtsystems. Allerdings lassen sich die Mikrozustände in Gruppen einteilen, innerhalb derer jeweils jede Zelle  $j$  eine feste Teilchenzahl  $N_j$  aufweist und sich die zugehörigen Mikrozustände folglich nach außen nicht unterscheiden. Jede solche Gruppe nennt man einen *Makrozustand* des Gesamtsystems. Insbesondere bei sehr großen System sind häufig nur diese Makrozustände unmittelbar zugänglich, da sie sich in makroskopischen Eigenschaften wie beispielsweise Volumen, Druck oder Temperatur unterscheiden, und deshalb versucht man häufig, große Systeme unter Verwendung solcher makroskopischer, durch Mittelung über mikroskopische Eigenschaften entstehender Größen zu beschreiben. Dazu muß man über die Verteilung der Mikrozustände im Phasenraum Bescheid wissen. Hierin liegt gerade der Unterschied zwischen Thermodynamik und statistischer Mechanik: Während sich erstere mit makroskopischen Größen beschäftigt, ohne Bezug auf mikroskopische Sachverhalte zu nehmen, also rein phänomenologisch vorgeht, beschreibt letztere makroskopische Eigenschaften unter expliziter Zuhilfenahme der mikroskopischen, indem sie, ganz ihrem Namen verpflichtet, diese mit statistischen Mitteln betrachtet.

Die Makrozustände sind durch Angabe der *Besetzungszahlen*  $N_j$  vollständig bestimmt, und es gilt natürlich  $\sum_j N_j = N$ . Elementare Kombinatorik lehrt, daß die Anzahl der Mikrozustände, die zu einem bestimmten Makrozustand führen, durch

$$W = \frac{N!}{\prod_j N_j!} \quad (1.3)$$

gegeben ist. Die  $W$  werden dabei in Abweichung von der üblichen Sprechweise *thermodynamische Wahrscheinlichkeiten* der zugehörigen Makrozustände genannt oder auch statistische Gewichte der Makrozustände; auch sie sind im allgemeinen sehr große Zahlen. Sie sind keine Wahrscheinlichkeiten in der mathematischen Bedeutung dieses Wortes<sup>11</sup>, sie geben aber dennoch Auskunft darüber, denn die statistische Wahrscheinlichkeit des Eintretens eines bestimmten Makrozustandes ist um so größer, je größer die Anzahl der Mikrozustände ist, durch die dieser Makrozustand realisiert wird<sup>12</sup>.

<sup>11</sup>Das ist schon an deren allgemeiner Eigenschaft  $W \geq 1$  erkennbar.

<sup>12</sup>Die Berechnung dieser statistischen Wahrscheinlichkeit ist nicht allgemein möglich, da die  $W$  natür-

Die *Entropie* ist definiert durch

$$S = k_B \ln W, \quad (1.4)$$

mit der *Boltzmann-Konstanten*  $k_B = 1,38 \cdot 10^{-23} \text{ J/K}$ . Gleichung (1.4) nennt man auch das *Boltzmannsche Prinzip*<sup>13</sup>. Danach ist die Entropie eines Makrozustands proportional zum Logarithmus der Anzahl der Mikrozustände, die bei diesem Makrozustand möglich sind. Offensichtlich ist die Entropie um so größer, je größer die thermodynamische Wahrscheinlichkeit des betrachteten Makrozustandes ist. Das läßt sich auch folgendermaßen interpretieren: Je kleiner  $W$ , desto genauer muß die Besetzung der Mikrozustände, also deren Anordnung im Phasenraum, stimmen, damit der gewünschte Makrozustand tatsächlich realisiert wird. Die Entropie ist daher in gewissem Sinne ein Maß für die innere Unordnung des Makrozustandes. Das hat jedoch nichts mit der Symmetrie oder mit sonstigen räumlichen oder anderen Ordnungsstrukturen des betrachteten Systems zu tun; Unordnung ist hier als Anzahl der für das System mikroskopisch erreichbaren Zustände zu verstehen. Je größer diese ist, desto größer ist dessen Entropie.

Verwendet man die für sehr große Zahlen gültige Stirlingsche Formel

$$N! \approx \left(\frac{N}{e}\right)^N, \quad N_j! \approx \left(\frac{N_j}{e}\right)^{N_j}$$

und setzt  $\rho_j \equiv \frac{N_j}{N}$ , so erhält man aus (1.3)

$$\ln W \approx -N \sum_j \rho_j \ln \rho_j$$

und weiter mit (1.4) für  $N \rightarrow \infty$  die mittlere Entropie pro Einzelsystem

$$s = \frac{S}{N} = -k_B \sum_j \rho_j \ln \rho_j. \quad (1.5)$$

$s$  nimmt für den Fall  $\rho_j = \text{konst.}$  ein Maximum an und wird minimal, nämlich  $s = 0$ , wenn genau ein  $\rho_j = 1$  ist und alle anderen verschwinden. Beim ersten Fall handelt es sich

---

lich von der Größe der verwendeten Elementarzellen abhängen. Formal geht man dazu folgendermaßen vor: Ist  $G$  die Anzahl der Elementarzellen im für das betrachtete System physikalisch zugänglichen Phasenraum, dann ist die Anzahl  $M$  der Möglichkeiten,  $N$  identische Teilchen auf diese  $G$  Zellen zu verteilen, wobei in jeder Zelle beliebig viele Teilchen sein dürfen, gegeben durch

$$M = \frac{(G + N - 1)!}{(G - 1)! N!}.$$

Die Wahrscheinlichkeit des Eintretens eines Makrozustands mit  $W$  möglichen Mikrozuständen beträgt dann

$$p = \frac{W}{M} = \frac{(G - 1)! (N!)^2}{(G + N - 1)! \prod_j N_j},$$

ein Ausdruck, der allerdings wenig praktische Bedeutung hat.

<sup>13</sup>Boltzmann entdeckte die Proportionalität der Entropie zum natürlichen Logarithmus der Anzahl der einen Makrozustand realisierenden Mikrozustände 1877 [71]. Er schrieb diese Beziehung jedoch nicht in der Form (1.4) auf; das tat erst Planck 1900 bei der Aufstellung seiner Strahlungsformel, wobei er auch den Proportionalitätsfaktor  $k_B$  einführte [409].



um das thermodynamische Gleichgewicht, das heißt, die Mikrozustände sind im Phasenraum gleichverteilt, und folglich entsprechen dem vorliegenden Makrozustand maximal viele Mikrozustände<sup>14</sup>. Im zweiten wird der Makrozustand durch nur genau einen Mikrozustand realisiert. Das bedeutet, daß das Eintreten des ersten Falls besonders wahrscheinlich, das des zweiten dagegen besonders unwahrscheinlich ist. Es bedeutet außerdem, daß im ersten Fall minimale, im zweiten Fall maximale Information über den tatsächlichen Mikrozustand des Systems verfügbar ist. Oder anders gesagt: Je größer die Entropie, desto geringer der Informationsgehalt des Systems und umgekehrt. Das paßt zur statistischen Deutung der Entropie; je mehr Mikrozustände einen bestimmten Makrozustand eines Systems repräsentieren, desto weniger kann man wissen, um welchen dieser Mikrozustände es sich dabei tatsächlich handelt und umgekehrt. Die Entropie ist damit ein Maß für die Unkenntnis des tatsächlich vorliegenden Mikrozustandes.

(1.4) heißt ganz ihrer Interpretation entsprechend auch *statistische Entropie*, während (1.5) gelegentlich nicht ganz korrekt als informationstheoretische Entropie bezeichnet wird, womit angedeutet wird, daß in der Informatik und in verwandten Gebieten ein mathematisch äquivalenter Ausdruck von Bedeutung ist<sup>15</sup>. Wir werden im nächsten Abschnitt eine weitere Definition der Entropie betrachten.

### 1.2.4 Ein wenig Thermodynamik

Auf den Unterschied zwischen Thermodynamik und statistischer Mechanik wurde bereits hingewiesen. Die Thermodynamik beschäftigt sich mit der Beschreibung makroskopischer Eigenschaften, beispielsweise Druck, Temperatur, Magnetisierung und dergleichen, von makroskopischen Systemen. Dabei ist sie eine rein phänomenologische Theorie, das heißt, sie versucht, empirische Beziehungen zwischen den thermodynamischen Größen zu ordnen und mathematisch zu formulieren, ohne dabei Begründungen dafür zu liefern, die

<sup>14</sup>Die Aussage, daß die Entropie im thermodynamischen Gleichgewicht maximal ist, ist der Inhalt des berühmten *H-Theorems* von Boltzmann [70]. Der Name kommt dabei von der von Boltzmann definierten Größe

$$H(f) = \iint_{\Omega \times \mathbb{R}^3} f(\vec{r}, \vec{v}) \ln f(\vec{r}, \vec{v}) d^3v d^3r,$$

wobei  $\Omega$  das vom System eingenommene Raumgebiet und  $f(\vec{r}, \vec{v})$  die zeitabhängige Teilchendichte im Phasenraum ist. Das H-Theorem besagt, daß diese Größe stets nur abnehmen kann, was aufgrund der Relation  $S(f) = -H(f)$  zur Aussage äquivalent ist,  $S$  könne nur zunehmen.

<sup>15</sup>Es handelt sich dabei um die sogenannte *Informationsentropie* nach Shannon, mit deren Hilfe man die minimale mittlere Anzahl von Bits abschätzen kann, die benötigt werden, um einen bestimmten, aus einer vorgegebenen Menge von  $N$  Symbolen aufgebauten Ausdruck  $X$  zu codieren [472], [473], [474]. Die Informationsentropie ist definiert als

$$S(X) = - \sum_{i=0}^N p_i \log_2 p_i;$$

dabei sind die  $p_i$  die relativen Häufigkeiten der einzelnen Symbole. Die minimale mittlere Anzahl der benötigten Bits ist dann

$$n(X) = \min \{q \in \mathbb{N} \mid q \geq S(X)\}.$$

Jaynes konnte zeigen [276], daß bei geeigneter Interpretation ein Zusammenhang zur statistischen Mechanik hergestellt werden kann, wenn man die thermodynamische Entropie als diejenige Informationsmenge auffaßt, die notwendig wäre, um den exakten Mikrozustand eines Systems zu bestimmen.



etwa auf den mikroskopischen Aufbau der betrachteten Systeme zurückführbar wären. In den folgenden Abschnitten werden wir Entropie und Energie, die wir aus mikroskopischer Sicht schon kennen, aus thermodynamischer Perspektive diskutieren.

#### 1.2.4.1 Zustandsgrößen

Zustandsgrößen sind Größen, die einem physikalischen System *unabhängig von dessen Vorgeschichte* zugeordnet werden können und damit allein durch dessen augenblicklichen Zustand gegeben sind. Mathematisch wird das dadurch ausgedrückt, daß das Differential  $dA$  einer Zustandsgröße  $A$ , aufgefaßt als alternierende Differentialform, ein *vollständiges Differential* ist<sup>16</sup>, oder äquivalent dazu durch die Forderung

$$\oint dA = 0.$$

Wir beginnen gleich mit zwei Gegenbeispielen. Der *Arbeit*  $dW$ , die ein System verrichtet oder ihm zugeführt wird, kann keine Zustandsgröße  $W$  zugeordnet werden, denn sonst müßte  $\oint dW = 0$  gelten; diese Größe beschreibt beispielsweise den Flächeninhalt eines Kreisprozesses in einem  $p$ - $V$ -Diagramm, der offensichtlich nicht verschwindet.

Das selbe gilt für die Wärmemenge  $dQ$ . Unter der *Wärme*  $Q$  versteht man in der Thermodynamik über eine Systemgrenze hinweg transportierte thermische Energie. Auch hier gilt ersichtlich  $\oint dQ \neq 0$ , denn Wärme ist eine *Prozeßgröße*, bei der es wesentlich auf den Transportcharakter ankommt: Wärme ist immer in Bewegung. Daher kann man dem Zustand eines Systems keine von dessen Vorgeschichte unabhängige charakteristische Wärme zuschreiben.

Betrachten wir stattdessen die *innere Energie*  $U$ , so ergibt sich ein anderes Bild. Das ist gerade der Inhalt des *ersten Hauptsatzes der Thermodynamik*:

*Für jedes thermodynamische System gibt es eine Zustandsgröße  $U$ , die wir dessen Energie nennen. Sie wächst mit der zugeführten Wärmemenge  $dQ$  und nimmt mit der vom System geleisteten Arbeit  $dW$  ab.*

Das Differential

$$dU = dQ - dW$$

ist somit ein vollständiges Differential, und es gilt

$$\oint dU = 0.$$

Ein weiteres Beispiel für eine Zustandsgröße ist die *Enthalpie*, definiert durch

$$H = U + pV.$$

Für deren Differential  $dH = dU + p dV + V dp$  gilt wegen  $dU = dQ - p dV$

$$dH = dQ + V dp,$$

woraus erkennbar ist, daß die Enthalpie anschaulich die einem System bei konstant gehaltenem Druck von außen zugeführte Wärme darstellt.

<sup>16</sup>Hier ist auch der Ausdruck *exaktes Differential* üblich.

### 1.2.4.2 Entropie als Zustandsgröße

Der hier beschriebene Formalismus erlaubt nun einen sehr direkten formalen mathematischen Zugang zum Begriff der Entropie, der allerdings anschließend noch mit physikalischem Inhalt zu füllen sein wird.

Weiter oben wurde bemerkt, daß die Wärme  $Q = U + W$  keine Zustandsgröße ist;  $dQ = dU + dW$  ist folglich kein vollständiges Differential. Es läßt sich jedoch zu einem solchen machen, und zwar durch Multiplikation mit einem *integrierenden Faktor*. Wir verleihen diesem die Bezeichnung  $1/T$  und verraten damit gleichzeitig, daß dessen Inverses natürlich nichts anderes als die *absolute Temperatur*  $T$  des Systems ist.  $T$  ist selbst eine Zustandsgröße; dieser Sachverhalt ist Gegenstand des *nullten Hauptsatzes der Thermodynamik*:

*Zu jedem thermodynamischen System gibt es eine Zustandsgröße  $T$ , ihre absolute Temperatur. Wenn sich zwei Systeme im thermodynamischen Gleichgewicht befinden, ist die absolute Temperatur bei beiden Systemen gleich.*

Multiplikation von  $dQ$  mit  $1/T$  führt auf das vollständige Differential

$$dS = \frac{dQ}{T} = \frac{dU + p dV}{T} \quad (1.6)$$

und damit auf eine neue Zustandsgröße, die *Entropie*  $S$  des betrachteten Systems. Daß es sich dabei tatsächlich um eine Zustandsgröße handelt, kann man zeigen, indem man einen beliebigen reversiblen Kreisprozeß im  $p$ - $V$ -Diagramm in unendlich viele unendlich schmale reversible *Carnot-Prozesse* zerlegt; die vielen kleinen Zacken aus abwechselnd isothermen und adiabatischen Pfaden, die bei  $n$  solchen Prozessen noch auftreten, verschwinden beim Grenzübergang  $n \rightarrow \infty$ . Für reversible Carnot-Prozesse gilt bekanntlich  $Q_1/Q_2 = T_1/T_2$  und im infinitesimalen Fall daher

$$\frac{dQ_1}{T_1} = \frac{dQ_2}{T_2},$$

woraus unmittelbar

$$\oint \frac{dQ_{\text{rev}}}{T} = 0 \quad (1.7)$$

folgt. Der Index bei  $dQ$  betont dabei den reversiblen Charakter des beschriebenen Kreisprozesses. Damit ist bewiesen, daß  $dS = dQ/T$  ein vollständiges Differential und  $S$  eine Zustandsgröße ist, sofern das System ausschließlich reversible Veränderungen durchläuft. Für die Entropie oder genauer gesagt die Entropiedifferenz eines thermodynamischen Systems gilt somit

$$\Delta S = \int \frac{dQ_{\text{rev}}}{T}.$$

Wir berechnen als Beispiel die Entropie eines idealen Gases. Dessen Anfangszustand sei durch die Werte  $S_0$ ,  $V_0$  und  $T_0$  definiert; da sie sich bei der Berechnung von Entropiedifferenzen herausheben, müssen sie nicht festgelegt werden. Für ideale Gase gilt einerseits die Energiegleichung  $U = \nu C_V T + \text{const.}$ , wobei  $C_V$  die molare Wärmekapazität

bei konstantem Volumen ist, und andererseits die Zustandsgleichung<sup>17</sup>  $pV = \nu RT$  mit der allgemeinen Gaskonstanten  $R = k_B N_A = 8,314 \text{ J/mol}\cdot\text{K}$ ;  $N_A = 6,022 \cdot 10^{23}$  ist die Avogadro-Zahl und  $\nu$  die Anzahl der Mole. Eingesetzt in (1.6) liefert das

$$dS = \nu C_V \frac{dT}{T} + \nu R \frac{dV}{V}$$

und nach Integration

$$S - S_0 = \nu C_V \int_{T_0}^T \frac{dT'}{T'} + \nu R \int_{V_0}^V \frac{dV'}{V'} = \nu C_V \ln \frac{T}{T_0} + \nu R \ln \frac{V}{V_0}. \quad (1.8)$$

Als nächstes gilt es, sich davon zu überzeugen, daß obige Definition der Entropie äquivalent zu der in Abschnitt 1 gegebenen statistischen Definition derselben ist. Dazu betrachten wir erst einmal ein einfaches Beispiel, nämlich einen Behälter mit Volumen  $V_1$  und darin die freie Expansion eines Gases vom Volumen  $V_2$  auf das gesamte Volumen  $V_1$ . Stellt man sich den Behälter als in  $n$  gleich große Teile mit Volumen  $V_2$  eingeteilt vor, so gilt  $V_1 = nV_2$ . Für ein Teilchen gibt bezogen auf die Teile des Behälters in diesem  $n$  mögliche Plätze, für zwei Teilchen gibt es folglich  $n^2$  und für  $N$  Teilchen entsprechend  $n^N$  mögliche Anordnungen. Die thermodynamische Wahrscheinlichkeit des Gases im Zustand der vollständigen Expansion in dem Behälter beträgt daher  $W = n^N$ . Wegen  $n = V_1/V_2$  und  $R = k_B N_A = k_B N/\nu$  folgt daraus

$$\ln W = N \ln n = \frac{\nu R}{k_B} \ln \frac{V_1}{V_2}$$

und damit

$$k_B \ln W = \nu R \ln \frac{V_1}{V_2}.$$

Das ist aber nach (1.8) gerade die Entropiedifferenz, die auftritt, wenn sich ein ideales Gas im Vakuum vom Volumen  $V_2$  auf das Volumen  $V_1$  ausdehnt, da die Temperatur bei diesem Prozeß konstant bleibt<sup>18</sup>. Es gilt hier also in der Tat

$$\Delta S = k_B \ln W.$$

Die Äquivalenz des Boltzmannschen Prinzips zur Clausiusschen Definition der Entropie läßt sich nun ganz allgemein zeigen. Dazu betrachten wir ein allgemeines System aus  $N$  Teilchen, von denen jeweils  $N_j$  Stück die Energie  $\varepsilon_j$  besitzen sollen. Die Gesamtenergie beträgt somit  $E = \sum_j N_j \varepsilon_j$ . Das System soll sich im thermischen Gleichgewicht befinden und eine reversible Wärmezufuhr erfahren. Das bedeutet daß die Besetzung der Energieniveaus  $\varepsilon_j$  der Boltzmann-Statistik folgt, es gilt also

$$N_j = \frac{N e^{-\varepsilon_j/k_B T}}{\sum_j e^{-\varepsilon_j/k_B T}}. \quad (1.9)$$

<sup>17</sup>Im Volksmund *ideale Gasgleichung*

<sup>18</sup>Dieser Vorgang ist als *Versuch von Gay-Lussac* bekannt geworden.

Die Änderung der Energie ist gleich der Wärmezufuhr; das liefert

$$dE = dQ = \sum_j N_j d\varepsilon_j + \sum_j \varepsilon_j dn_j. \quad (1.10)$$

Für die Anzahl der Mikrozustände eines beliebigen Zustands des Systems gilt<sup>19</sup>

$$W = \frac{N!}{\prod_j N_j!},$$

was sich mit Hilfe der Stirlingschen Formel abschätzen läßt zu

$$\ln W = N \ln N - \sum_j N_j \ln N_j.$$

Das Boltzmannsche Prinzip liefert folglich für die Entropie des Systems

$$S = k_B \left( N \ln N - \sum_j N_j \ln N_j \right).$$

Die reversible Wärmezufuhr bewirkt kleine Änderungen von  $W$  und damit auch von  $S$ ; für diese gilt wegen (1.4) und der durch  $\sum_j dN_j = 0$  ausgedrückten Erhaltung der Teilchenzahl

$$\begin{aligned} dS &= -k_B \sum_j \ln N_j dN_j - k_B \sum_j \frac{dN_j}{N_j} N_j \\ &= -k_B \sum_j \ln \frac{N}{\sum_j e^{-\varepsilon_j/k_B T}} dN_j - \frac{1}{T} \sum_j \varepsilon_j dN_j \\ &= -\frac{1}{T} \sum_j \varepsilon_j dN_j. \end{aligned}$$

Da die Wärmezufuhr reversibel erfolgen soll, gilt  $\sum_j N_j d\varepsilon_j = 0$ , und mit (1.10) ergibt das

$$dS = \frac{dQ}{T}.$$

Aus der statistischen Definition

$$S = k_B \ln W$$

folgt also die thermodynamische Definition

$$\Delta S = \int \frac{dQ}{T}$$

und umgekehrt.

---

<sup>19</sup>Siehe Abschnitt 2.3

Mit Hilfe der Entropie kann man zwei weitere wichtige Zustandsgrößen thermodynamischer Systeme definieren, und zwar einerseits die *freie Energie*  $F$ , definiert durch

$$F = U - TS,$$

andererseits die *freie Enthalpie*  $G$ , die durch

$$G = H - TS = U + pV - TS$$

erklärt ist.  $U$ ,  $H$ ,  $F$  und  $G$  sind Beispiele für *thermodynamische Potentiale*. Sie sind insbesondere in der Gleichgewichtsthermodynamik von überragender Bedeutung. Ihre Umrechnungsformeln zeigen exemplarisch die Methode der *Legendre-Transformationen*, mit denen in Differentialausdrücken von einer unabhängigen Variable zu deren konjugierter Variable gewechselt werden kann, und die außer in der Thermodynamik auch in der Mechanik und allgemein in der Analysis vielfältige Anwendungen finden.

### 1.2.4.3 Irreversible Prozesse und der Satz von Clausius

Für *reversible Zustandsänderungen* stellt die Entropie eine weitere Zustandsgröße dar, ist also für thermodynamische Berechnung von großer Wichtigkeit. Ihre wahrhaft fundamentale Bedeutung entfaltet diese Größe jedoch erst im Zusammenhang mit *irreversiblen Zustandsänderungen*. Eine genauere Betrachtung solcher Prozesse führt unmittelbar zum zweiten Hauptsatz der Thermodynamik.

Am einfachsten sieht man dies wieder an der oben diskutierten freien Expansion eines idealen Gases. Die dabei berechnete Entropiedifferenz

$$\Delta S = \nu R \ln \frac{V_1}{V_2}$$

ist wegen  $V_2 > V_1$  stets negativ, *die Entropie nimmt also zu*. Dieser Sachverhalt gilt in völliger Allgemeinheit, wovon wir uns sogleich überzeugen werden.

Wir ersetzen dazu den reversiblen Kreisprozeß aus Abschnitt 4.2 durch einen beliebigen irreversiblen Kreisprozeß und denken uns diesen wieder in unendlich viele unendlich schmale Carnot-Prozesse, von denen nun aber einer oder mehrere irreversibel sind. Für irreversible Carnot-Prozesse gilt jedoch  $Q_1/Q_2 < T_1/T_2$ , im infinitesimalen Fall also

$$\frac{dQ_1}{T_1} < \frac{dQ_2}{T_2}, \quad (1.11)$$

so daß bei irreversiblen Kreisprozessen das Gleichheitszeichen in (1.7) durch ein Kleinerzeichen zu ersetzen ist. Für einen beliebigen Kreisprozeß gilt folglich

$$\oint \frac{dQ}{T} \leq 0.$$

Das ist die *Clausiussche Ungleichung* [103], [104]. Sie ist eine spezielle Form des zweiten Hauptsatzes der Thermodynamik.

Um nun die Allgemeingültigkeit der in obigem Spezialfall festgestellten Entropiezunahme zu sehen, betrachten wir ein thermodynamisches System, das durch einen beliebigen irreversiblen Prozeß vom Zustand  $A$  in den Zustand  $B$  überführt werde, und

dazu einen reversiblen Prozeß, der Das System von  $B$  nach  $A$  bringen soll. Zusammen ergibt dies einen irreversiblen Kreisprozeß, für den nach der Clausiusschen Ungleichung

$$\oint \frac{dQ}{T} < 0$$

gilt, ausführlich geschrieben also

$$\int_A^B \frac{dQ_{\text{irr}}}{T} + \int_B^A \frac{dQ_{\text{rev}}}{T} = \int_A^B \frac{dQ_{\text{irr}}}{T} + S_A - S_B < 0.$$

Daraus folgt

$$S_B - S_A > \int_A^B \frac{dQ_{\text{irr}}}{T}. \quad (1.12)$$

Dabei setzt sich  $dQ_{\text{irr}}$  aus der von außen zugeführten Wärme  $dQ_a$  und irgendwelchen irreversiblen Wärmeaustauschprozessen  $dQ_i$  im Inneren des Systems zusammen. Für thermisch abgeschlossene Systeme gilt  $dQ_a = 0$ , und (1.12) wird zu

$$S_B - S_A > \sum_i \int_A^B \frac{dQ_i}{T}. \quad (1.13)$$

Alle Integrale in der Summe auf der rechten Seite von (1.13) sind jedoch wegen (1.11) positiv, und es gilt daher

$$S_B > S_A.$$

Damit haben wir eine weitere Formulierung des *zweiten Hauptsatzes der Thermodynamik*:

*In abgeschlossenen Systemen kann die Entropie bei irreversiblen Vorgängen nur zunehmen.*

Bei rein reversiblen Prozessen bleibt sie gleich, abnehmen kann sie überhaupt nicht. Entropie kann stets nur lokal, also in einem Teilsystem abgesenkt werden, das jedoch auf Kosten der Entropieerhöhung im Restsystem.

Dieses Gesetz ist nicht nur für die Thermodynamik von fundamentaler Bedeutung, denn es zeichnet eine *Zeitrichtung* aus, nämlich diejenige in Richtung größer werdender Entropie. Auf der Ebene der Thermodynamik ist das natürlich eine rein phänomenologische Aussage; sie läßt sich indes mit Hilfe der statistischen Definition der Entropie plausibel machen. Diese ist wie in Abschnitt 2.3 beschrieben ein Maß für die Anzahl der Mikrozustände eines Systems, die zum gleichen Makrozustand führen. Der zweite Hauptsatz sagt demnach, daß sich ein sich selbst überlassenes System stets von Makrozuständen, die durch wenige Mikrozustände repräsentierbar sind, zu solchen mit vielen möglichen Mikrozuständen entwickelt. Genauer gesagt: Bei abgeschlossenen Systemen führen irreversible Prozesse stets zu einem Anwachsen der Anzahl der möglichen Mikrozustände, während diese bei reversiblen Prozessen gleich bleibt. Die Anzahl der zu einem Makrozustand gehörenden Mikrozustände läßt als Maß für die Wahrscheinlichkeit des Auftretens dieses Makrozustandes deuten. Ein Makrozustand ist umso wahrscheinlicher,

durch je mehr Mikrozustände er realisiert wird und umgekehrt, wobei vorausgesetzt wird, daß alle Mikrozustände gleichwahrscheinlich sind. *Das Prinzip der Erhöhung der Entropie besagt demnach nichts anderes, als das sich physikalische Systeme ohne äußeren Zwang stets von unwahrscheinlicheren zu wahrscheinlicheren Makrozuständen entwickeln.*

Die Vorstellung abgeschlossener Systeme ist strenggenommen eine Idealisierung. Genauer gesagt ist das einzige wirklich abgeschlossene System das Universum. In diesem Sinne formulieren wir mit Clausius den ersten und den zweiten Hauptsatz der Thermodynamik wie folgt: *Die Energie des Universums ist konstant. Seine Entropie strebt einem Maximum zu* [103], [104].

Abgesehen von diesen beinahe schon metaphysischen Aussagen erhält der zweite Hauptsatz seine Bedeutung vor allen Dingen auch durch seine Eigenschaft, über das freiwillige oder nicht freiwillige Ablaufen von Naturvorgängen zu entscheiden. Denn offensichtlich läßt sich aus ihm die Aussage ableiten, daß ein Prozeß dann spontan abläuft, wenn die Entropie insgesamt dabei zunimmt. Hiervon macht man beispielsweise in der Chemie Gebrauch, um zu ermitteln, ob eine Reaktion von selbst stattfindet oder nicht. Da es dabei auf die gesamte Entropieänderung ankommt, muß man eigentlich die zu beschreibende Reaktion und deren Umgebung betrachten; wünschenswert wäre es jedoch, allein mit den Eigenschaften des Systems selbst auszukommen, in welchem die Reaktion stattfindet. Man unterteilt hierfür die gesamte Entropieänderung in einen Anteil des Systems und einen der Umgebung und schreibt

$$dS_{\text{ges}} = dS_{\text{Sys}} + dS_{\text{Umg}} \equiv dS + dS_{\text{Umg}}.$$

Bei konstanter Temperatur und konstantem Druck kommt die Entropieänderung der Umgebung allein durch Wärmeaustausch von System und Umgebung zustande. Für die zugehörige Entropieänderung gilt daher

$$dS_{\text{Umg}} = -\frac{dH}{T}$$

und für die gesamte Entropieänderung entsprechend

$$dS_{\text{ges}} = dS - \frac{dH}{T}.$$

Mit der freien Enthalpie  $G = H - TS$  wird daraus

$$dG = -T dS_{\text{ges}}. \quad (1.14)$$

Aufgrund der Voraussetzung  $p, T = \text{konst.}$  ist das gleichbedeutend mit  $dG = dH - T dS$ , das heißt, man hat es tatsächlich nur mit Größen zu tun, die sich auf Änderungen des betrachteten Systems beziehen, so daß eine Betrachtung der Änderungen der Umgebung nicht erforderlich ist. Für praktische Anwendungen kann man im allgemeinen mit Differenzen anstelle von Differentialen rechnen. Die Größe

$$\Delta G = \Delta H - T \Delta S$$

heißt *freie Reaktionsenthalpie*. Aus dem zweiten Hauptsatz der Thermodynamik folgt nun für eine bei konstantem Druck und konstanter Temperatur durchgeführte Reaktion:



1. Ist  $\Delta G < 0$ , läuft die Reaktion freiwillig ab,
2. ist  $\Delta G = 0$ , befindet sich das System im Gleichgewicht,
3. ist  $\Delta G > 0$ , läuft die Reaktion nicht freiwillig ab.

Für die umgekehrte Reaktion gilt dabei jeweils das Gegenteil. Die freie Reaktionsenthalpie berücksichtigt dabei gleich zwei Grundprinzipien des Ablaufs dynamischer Systeme: Jede Reaktion strebt stets erstens ein Energieminimum und zweitens ein Maximum an innerer Unordnung an.

Es sollte nicht unerwähnt bleiben, daß der zweite Hauptsatz genau besehen nicht dieselbe grundlegende Naturgesetzeigenschaft beanspruchen kann wie der erste. Letzterer ist eine spezielle Formulierung des Energieerhaltungssatzes, der wie alle starken Erhaltungssätze aus einer Symmetrie, in diesem Fall aus der Invarianz der Lagrange-Dichte des Systems unter Raum-Zeit-Translationen, folgt und damit aus fundamentalen Naturprinzipien abgeleitet werden kann<sup>20</sup>. Ersterer dagegen ist ein Erfahrungssatz: Die spontane Abnahme der Anzahl der einem System zugänglichen Mikrozustände verstößt nicht notwendigerweise gegen den Energieerhaltungssatz, ist aber im obigen Sinne extrem unwahrscheinlich, so unwahrscheinlich, daß ihr Eintreten für alle praktischen Belange ausgeschlossen werden kann. Damit ist der zweite Hauptsatz zwar gemessen an seiner Position in der Hierarchie grundlegender physikalischer Gesetze nicht gleichberechtigt mit dem ersten, seine Bedeutung für den Ablauf realer physikalischer Prozesse ist aber mindestens genau so groß, da er für alle praktischen Belange festlegt, welche davon tatsächlich stattfinden und welche nicht. Gleichwohl kann hier von völliger Klarheit der Sachlage nach wie vor keine Rede sein; insbesondere die naturphilosophische Debatte um den zweiten Hauptsatz hält unverändert an.

#### 1.2.4.4 Das Nernstsche Wärmetheorem

Bisher haben wir stets *Entropiedifferenzen* betrachtet. Es fehlt noch so etwas wie ein Referenzwert, mit dessen Hilfe sich absolute Werte der Entropie und damit auch anderer Zustandsgrößen thermodynamischer Systeme angeben lassen.

Wir betrachten wieder einen reversiblen Prozeß und dazu die freie Energie  $F = U - TS$  und deren Differential  $dF = dU - T dS - S dT = -p dV - S dT$ . Bei konstantem Volumen  $V$  folgt daraus

$$\left(\frac{\partial F}{\partial T}\right)_V = -S \quad (1.15)$$

und damit weiter

$$F = U + T \left(\frac{\partial F}{\partial T}\right)_V, \quad (1.16)$$

eine Beziehung, die auch für  $\Delta F$  gilt,

$$\Delta F = \Delta U + T \left(\frac{\partial \Delta F}{\partial T}\right)_V. \quad (1.17)$$

---

<sup>20</sup>Im Gegensatz dazu bringen schwache Erhaltungssätze lediglich die Zeitunabhängigkeit physikalischer Größen zum Ausdruck.



(1.16) beziehungsweise (1.17) heißt *Gibbs-Helmholtzsche Gleichung*.  $\Delta F$  ist ein Maß für die *chemische Affinität* von Stoffen, das heißt für die Geschwindigkeit chemischer Reaktionen zwischen ihnen. Die ursprüngliche, auf Thomsen und Berthelot zurückgehende Vermutung,  $\Delta U$  als Maß hierfür verwenden zu können, erweist sich als undurchführbar, was man besonders deutlich bei endothermen Reaktionen erkennt, denn bei diesen ist  $\Delta U$  negativ,  $\Delta F$  jedoch wie bei jeder chemischen Reaktion positiv.

Es stellt sich nun die Frage nach dem Wert der Größe

$$\Delta F - \Delta U = T \left( \frac{\partial \Delta F}{\partial T} \right)_V.$$

Dieser ist erfahrungsgemäß häufig sehr klein und noch dazu umso kleiner, je kleiner die Temperatur ist, was Nernst zu der Vermutung veranlaßte, daß  $\Delta F$  und  $\Delta U$  für  $T \rightarrow 0$  den gleichen Wert annehmen und sich darüberhinaus asymptotisch berühren. Formal bedeutet diese Aussage

$$\lim_{T \rightarrow 0} \left( \frac{\partial \Delta F}{\partial T} \right)_V = \lim_{T \rightarrow 0} \left( \frac{\partial \Delta U}{\partial T} \right)_V = 0.$$

Das ist das *Nernstsche Wärmetheorem*. Man nennt es auch den *dritten Hauptsatz der Thermodynamik*. In dieser Form sagt es aus, daß bei hinreichend tiefen Temperaturen sowohl die Differenz der freien Energie als auch die Differenz der inneren Energie zweier Zustände eines thermodynamischen Systems temperaturunabhängig wird<sup>21</sup>.

Hier ist dem Thema dieses Kapitels entsprechend eine etwas andere Formulierung des dritten Hauptsatzes interessanter. Wir gelangen zu dieser, indem wir die Entropie einbauen. Verwenden wir

$$\Delta F = \Delta U - T \Delta S,$$

so erhalten wir mit Hilfe von (1.15)

$$\left( \frac{\partial \Delta F}{\partial T} \right)_V = -\Delta S.$$

Mit anderen Worten: Es gilt  $\lim_{T \rightarrow 0} \Delta S = 0$ , das heißt *bei hinreichend tiefen Temperaturen nimmt die Entropie thermodynamischer Systeme einen konstanten Wert an*.

Wie dieser Wert nun genau aussieht kann im Rahmen der Thermodynamik nicht bestimmt werden. Allerdings ist in jener generell nur von Entropiedifferenzen die Rede, das heißt die Entropie ist nur bis auf eine additive Konstante definiert. Planck nahm das zum Anlaß, diese Konstante auf den Wert Null festzulegen [411]. Was in der Thermodynamik eine zwar sinnvolle, aber eigentlich willkürliche Annahme darstellt, läßt sich in der Quantenmechanik beweisen. Denn am absoluten Nullpunkt muß sich ein System genau wie jedes beliebige Teilsystem desselben *im Grundzustand befinden*. Damit sind die

<sup>21</sup>Ursprünglich formulierte Nernst sein Theorem nur für feste und flüssige Stoffe [388], [389], da Gase spätestens in der Nähe des absoluten Nullpunktes zu Phasenübergängen neigen und  $\Delta F$  und  $\Delta U$  dann keine stetigen Funktionen der Temperatur mehr sind. Diese Einschränkung setzt aber stillschweigend voraus, daß Zustandsgleichungen wie die ideale Gasgleichung oder die van der Waals-Gleichung auch bei sehr tiefen Temperaturen noch gelten, was jedoch nicht der Fall ist. Stattdessen neigen Gase bei tiefsten Temperaturen dazu zu entarten. Im allgemeinen gilt das Nernstsche Wärmetheorem auch für solche Quantengase, es gibt jedoch Ausnahmen.

thermodynamischen Wahrscheinlichkeiten jedes Teilsystems und folglich auch diejenige des Gesamtsystems gleich Eins, und die statistische Definition der Entropie liefert für letztere unmittelbar den Wert Null. Der dritte Hauptsatz der Thermodynamik läßt sich daher auch wie folgt formulieren:

*Die Entropie physikalischer Systeme besitzt am absoluten Nullpunkt den Wert Null.*

Wir betrachten einige Folgerungen aus dem dritten Hauptsatz. Für den *thermischen Ausdehnungskoeffizienten*  $\alpha$  gilt

$$\alpha = \frac{1}{V} \left( \frac{\partial V}{\partial T} \right)_p,$$

und wegen der thermodynamischen Relation

$$\left( \frac{\partial S}{\partial p} \right)_T = - \left( \frac{\partial V}{\partial T} \right)_p$$

folgt daraus  $\lim_{T \rightarrow 0} \alpha = 0$ : *Der Ausdehnungskoeffizient verschwindet am absoluten Nullpunkt.* Analoges gilt für den *Spannungskoeffizienten*  $\beta$ , definiert durch

$$\beta = \frac{1}{p} \left( \frac{\partial p}{\partial T} \right)_V.$$

Aufgrund von

$$\left( \frac{\partial S}{\partial p} \right)_T = - \left( \frac{\partial V}{\partial T} \right)_p$$

folgt auch hier  $\lim_{T \rightarrow 0} \beta = 0$ : *Der Spannungskoeffizient verschwindet am absoluten Nullpunkt.*

Wichtiger für unsere Interessen hier ist das Verhalten der Wärmekapazitäten. Für diese gilt

$$C_V = T \left( \frac{\partial S}{\partial T} \right)_V, \quad C_p = T \left( \frac{\partial S}{\partial T} \right)_p,$$

und Integration liefert

$$S(V, T) = \int_0^T \frac{C_V}{T} dT, \quad S(p, T) = \int_0^T \frac{C_p}{T} dT,$$

wobei die unteren Integrationsgrenzen direkt aus dem dritten Hauptsatz folgen. Aus diesen Integralen folgt aber unmittelbar  $\lim_{T \rightarrow 0} C_V = \lim_{T \rightarrow 0} C_p = 0$ , weil sie andernfalls an den unteren Grenzen unendlich werden würden. *Die spezifischen und molaren Wärmekapazitäten verschwinden am absoluten Nullpunkt.* Diese Konsequenz aus dem dritten Hauptsatz hat eine weitere zur Folge: *Es ist unmöglich, die Temperatur eines physikalischen Systems exakt auf den absoluten Nullpunkt abzusenken.*

### 1.2.5 Die Gibbsche Fundamentalbeziehung

Thermodynamische Größen beziehen sich stets auf Systeme aus sehr vielen Teilchen, so daß der Größe des Systems, ausgedrückt durch Stoffmenge, Teilchenzahl, Masse oder was auch immer, eine wichtige Bedeutung zukommt. Es werden jedoch nicht alle thermodynamischen Größen davon beeinflußt. Man unterscheidet daher zwischen *extensiven Größen*, die von der Größe des Systems abhängen, und *intensiven Größen*, die davon unabhängig sind. Beispiele für erstere sind trivialerweise Stoffmenge, Masse und Teilchenzahl, aber auch Volumen, innere Energie, Enthalpie, Entropie und so weiter; Beispiele für letztere sind Temperatur und Druck<sup>22</sup>. Die Unterscheidung von extensiven und intensiven Größen soll in den folgenden Abschnitten verdeutlicht werden, indem wir uns etwas mit statistischen Gesamtheiten beschäftigen. Dabei ist es nicht erforderlich, das gesamte Angebot von mikrokanonisch bis beliebig verallgemeinert großkanonisch durchzugehen; zwei Beispiele genügen vollauf. Als Nebenprodukt wird dabei eine statistische Begründung wichtiger thermodynamischer Sachverhalte herauskommen.

#### 1.2.5.1 Die kanonische Gesamtheit

Wir betrachten ein Ensemble identischer stationärer Systeme im thermischen Gleichgewicht mit Hamilton-Operator  $\hat{H}$ . Die Quantenmechanik lehrt, daß der Dichteoperator einer solchen *kanonischen Gesamtheit* die Form

$$\hat{\rho} = \frac{e^{-\hat{H}/k_{\text{B}}T}}{\text{tr } e^{-\hat{H}/k_{\text{B}}T}}.$$

hat. In der statistischen Mechanik schreibt man das üblicherweise

$$\hat{\rho} = \frac{1}{Z} e^{-\hat{H}/k_{\text{B}}T}$$

und definiert damit die *kanonische Zustandssumme*  $Z := \text{tr } e^{-\hat{H}/k_{\text{B}}T}$ . Zustandssummen lassen sich auch für andere Gesamtheiten bilden und stellen die zentralen Größen der Gleichgewichtsthermodynamik dar; aus ihnen kann man alle weiteren thermodynamischen Größen des Gleichgewichts bestimmen.

Für weitere konkrete Berechnungen ist es zweckmäßig, in die Energiebasis zu gehen.  $\{|\varphi(E)\rangle\}$  sei das System der Eigenzustände des Hamilton-Operators, es gelte also jeweils  $\hat{H}|\varphi(E)\rangle = E|\varphi(E)\rangle$ . Die Dichtematrix ist in dieser Basis diagonal; die Diagonalelemente sind<sup>23</sup>

$$\langle\varphi(E)|\hat{\rho}|\varphi(E')\rangle = \frac{e^{-E_n/k_{\text{B}}T}}{Z} \equiv \rho(E_n). \quad (1.18)$$

<sup>22</sup>Im Gegensatz zu diesen Beispielen sind stoffspezifische und molare Größen wie Dichte, Atomgewicht und dergleichen natürlich trivialerweise intensiv, da bei ihnen die Abhängigkeit von der Systemgröße herausgerechnet wird.

<sup>23</sup>Einer weit verbreiteten Konvention folgend verwenden wir in diesem Buch die Bra-Ket-Schreibweise, was bei Matrixelementen unbeschränkter Operatoren eigentlich problematisch ist. Da hier aber ausnahmslos Matrixelemente selbstadjungierter Operatoren auftreten, können keine nicht eindeutig definierte Ausdrücke auftreten.

Für die Zustandssumme folgt mit Hilfe der Vollständigkeitsrelation

$$\begin{aligned}
Z &= \text{tr} e^{-\hat{H}/k_{\text{B}}T} \\
&= \int \langle \varphi(E) | e^{-\hat{H}/k_{\text{B}}T} | \varphi(E) \rangle dE \\
&= \iint \langle \varphi(E) | \varphi(E') \rangle \langle \varphi(E') | e^{-\hat{H}/k_{\text{B}}T} | \varphi(E) \rangle dE dE' \\
&= \iint \delta(E - E') e^{-E'/k_{\text{B}}T} dE dE' \\
&= \int e^{-E/k_{\text{B}}T} dE; \tag{1.19}
\end{aligned}$$

Das Integral gilt für kontinuierliche Energieeigenwerte, für diskrete wird daraus die Summe

$$Z = \sum_{n=0}^{\infty} e^{-E_n/k_{\text{B}}T}.$$

Mit der Zustandssumme können wir den Erwartungswert des Hamilton-Operators berechnen. Dafür gilt zunächst (Vorgehensweise wie oben)

$$\begin{aligned}
\langle E \rangle &= \langle \hat{H} \rangle = \text{tr}(\hat{\rho} \hat{H}) \\
&= \frac{1}{Z} \text{tr}(\hat{H} e^{-\hat{H}/k_{\text{B}}T}) \\
&= \frac{1}{Z} \int \langle \varphi(E) | \hat{H} e^{-\hat{H}/k_{\text{B}}T} | \varphi(E) \rangle dE \\
&= \frac{1}{Z} \int E e^{-E/k_{\text{B}}T} dE.
\end{aligned}$$

Der Term hinter dem Integral erinnert verdächtig an Nachdifferenzieren und Kettenregel; in der Tat liefert (1.19)

$$\frac{\partial Z}{\partial T} = \int \frac{E}{k_{\text{B}}T^2} e^{-E/k_{\text{B}}T} dE$$

und man erhält

$$\langle E \rangle = k_{\text{B}} T^2 \frac{1}{Z} \frac{\partial Z}{\partial T}. \tag{1.20}$$

Wir schreiben von nun an dafür  $\langle E \rangle \equiv E$ .

Um zu einer thermodynamischen Formulierung des Energiemittelwerts zu gelangen, müssen wir diesen durch rein makroskopische Größen ausdrücken. Nach einer Idee von J. W. Gibbs kommen dabei auch gleich der erste und zweite Hauptsatz der Thermodynamik gewissermaßen als Nebenprodukte heraus [190]. Zunächst bringen wir die über das Boltzmannsche Prinzip

$$S = k_{\text{B}} \ln W$$

definierte *Entropie*  $S$  ins Spiel; dabei ist  $W$  die Anzahl der Mikrozustände, die einen bestimmten Makrozustand des Ensembles realisieren. Damit ergibt sich unmittelbar

$$Z = W e^{-E/k_{\text{B}}T};$$

es folgt mit Hilfe von (1.18)

$$S = k_B \ln Z + \frac{E}{T} = \frac{E - E_n}{T} - k_B \ln \rho(E_n).$$

und nach Multiplizieren mit  $\rho(E_n)$  und Summieren über  $n$  weiter

$$\sum_{n=0}^{\infty} S \rho(E_n) = \frac{1}{T} \left( \sum_{n=0}^{\infty} E \rho(E_n) - k_B \sum_{n=0}^{\infty} E_n \rho(E_n) \right) - \sum_{n=0}^{\infty} \rho(E_n) \ln \rho(E_n).$$

Wegen  $\sum_{n=0}^{\infty} \rho(E_n) = 1$  und  $\sum_{n=0}^{\infty} E_n \rho(E_n) = E$  erhält man daraus für die Entropie

$$S = -k_B \sum_{n=0}^{\infty} \rho(E_n) \ln \rho(E_n) = -k_B \langle \ln \rho \rangle. \quad (1.21)$$

Beachtet man nun, daß die Zustandssumme nicht nur von der absoluten Temperatur  $T$ , sondern mindestens auch vom Volumen  $V$  abhängig ist, kann man mit Hilfe der Relation

$$P_n = -\frac{\partial E_n}{\partial V},$$

die den *Druck* des Systems im Zustand  $n$  liefert, den mittleren Druck der kanonischen Gesamtheit berechnen. Man findet analog zu (1.20)

$$\langle P \rangle = \sum_{n=0}^{\infty} P_n \rho(E_n) = k_B T \frac{1}{Z} \frac{\partial Z}{\partial V}. \quad (1.22)$$

Auch hier schreiben wir abkürzend  $\langle P \rangle \equiv P$ . Mit (1.20) und (1.22) ergibt sich zusammengefaßt

$$d \ln Z = \frac{dZ}{Z} = \frac{1}{Z} \left( \frac{\partial Z}{\partial T} dT + \frac{\partial Z}{\partial V} dV \right) = \frac{1}{k_B T} \left( \frac{E}{T} dT + P dV \right).$$

Verwendet man außerdem  $\ln Z = S/k_B - E/k_B T$ , so erhält man

$$T dS - dE + \frac{E}{T} dT = \frac{E}{T} dT + P dV$$

und damit

$$dE = T dS - P dV. \quad (1.23)$$

Das ist die *Gibbssche Fundamentalbeziehung*<sup>24</sup>. Damit haben wir gleichzeitig die Relation (1.6) wiedergefunden, das heißt, Gleichung (1.23) beinhaltet für die hier betrachteten kanonischen Gesamtheiten den ersten Hauptsatz, indem sie die Relation  $dE = \delta Q - P dV$  enthält, die aussagt, daß die innere Energie eines Systems die Differenz aus zugeführter Wärme und verrichteter Arbeit ist, und die Gleichgewichtsversion des zweiten Hauptsatzes, indem sie die Relation  $\delta Q = dS/T$  etabliert. Insbesondere aber ist die Herleitung

<sup>24</sup>Da man (1.23) wie bereits erwähnt im Sinne des äußeren Differentialformenkalküls auffassen kann, spricht man häufig auch von der *Gibbsschen Fundamentalform*.

der Gleichung (1.23) die gewünschte statistische Begründung der thermodynamisch-phänomenologischen Beschreibung der Energie in Form eines vollständigen Differentialausdrucks.

Die Gibbssche Fundamentalbeziehung verknüpft außerdem exemplarisch extensive mit intensiven Größen. Sie stellt selbst das Differential einer extensiven Größe dar und enthält auf der rechten Seite eine Summe aus Produkten aus je einer intensiven Größe und dem Differential einer extensiven Größe. Auf diese Struktur werden wir wieder zurückkommen.

### 1.2.5.2 Die verallgemeinerte großkanonische Gesamtheit

Die Beschränkung auf Systeme, die sich nur in thermischem Kontakt mit ihrer Umgebung befinden, ist natürlich zu eng. Wir verallgemeinern das jetzt auf offene Systeme, denen sowohl durch thermischen Kontakt, als auch durch Teilchenaustausch und insbesondere durch Arbeitsverrichtung aller Art Energie zugeführt oder entzogen werden kann. Man spricht bei Ensembles identischer solcher Systeme von *verallgemeinerten großkanonischen Gesamtheiten*. Die Beschreibung des thermischen Energieaustauschs erfolgt wie im vorigen Abschnitt. Der Teilchenaustausch führt zu statistischen Schwankungen der Teilchenzahl um ihren Mittelwert  $\langle N \rangle \equiv N$ , so daß eine zusätzliche Wahrscheinlichkeitsverteilung der von der Energie  $E$  abhängigen Teilchenzahl  $N(E)$  existiert. Der dabei auftretende Energieaustausch wird im Hamilton-Operator durch Hinzufügen eines Terms

$$\hat{H}_N = -\mu N \quad (1.24)$$

berücksichtigt; dabei ist  $\mu$  das sogenannte *chemische Potential*<sup>25</sup>. Besteht das System aus  $k$  Systemkomponenten, so gehört zu jeder ein eigenes chemisches Potential, so daß (1.24) zu

$$\hat{H}_N = - \sum_{i=1}^k \mu_i N_i$$

wird. Das chemische Potential beschreibt die Fähigkeit des Systems, durch Änderung der Teilchenzahl Arbeit zu verrichten, oder anders formuliert, die Abhängigkeit der inneren Energie des Systems von der Teilchenzahl<sup>26</sup>. Alle weiteren Arbeitsverrichtungen geschehen in der Regel dadurch, daß an irgendwelchen Größen  $q_j$  verallgemeinerte äussere Kräfte  $f_j$  angreifen; diesen Effekten wird im Hamilton-Operator durch Hinzufügen eines Terms

$$H_f = - \sum_{j=1}^l f_j q_j$$

Rechnung getragen. Der gesamte Hamilton-Operator hat dann die Gestalt

$$\hat{H}_{\text{ges}} = \hat{H} - \sum_{i=1}^k \mu_i N_i - \sum_{j=1}^l f_j q_j$$

<sup>25</sup>Der Name erklärt sich dadurch, daß chemische Reaktionen exemplarische Vorgänge sind, bei denen sich Teilchenzahlen von offenen Systemen ändern.

<sup>26</sup>Anschaulich gesprochen ist das beispielsweise die Neigung des Systems, sich in und mit anderen Substanzen umzuwandeln, sich in andere Aggregatzustände zu begeben oder auch sich im Raum umzuverteilen.

oder, wenn wir  $\mu_i \equiv f_i$  und  $N_i \equiv q_i$  für  $i = 1, 2, \dots, k$  setzen und die Summe bis  $L \equiv k + l$  laufen lassen,

$$\hat{H}_{\text{ges}} = \hat{H} - \sum_{j=1}^L f_j q_j.$$

Entsprechend lautet der Dichteoperator eines solchen allgemeinen offenen Systems

$$\hat{\rho} = \frac{1}{\Upsilon} \exp \left[ -\frac{1}{k_B T} \left( \hat{H} - \sum_{j=1}^L f_j q_j \right) \right],$$

mit der *verallgemeinerten großkanonischen Zustandssumme*

$$\Upsilon = \text{tr} \exp \left[ -\frac{1}{k_B T} \left( \hat{H} - \sum_{j=1}^L f_j q_j \right) \right].$$

Die Berechnung thermodynamischer Mittelwerte geschieht nun in völliger Analogie zum vorigen Abschnitt. Hier wie dort erhält man den Mittelwert einer durch einen selbst-adjungierten Operator  $\hat{A}$  repräsentierten Größe allgemein mit Hilfe des Dichteoperators gemäß der Relation

$$\langle \hat{A} \rangle = \text{tr} (\hat{A} \hat{\rho}).$$

Wieder schreiben wir abkürzend  $\langle q_j \rangle \equiv Q_j$  und erhalten wie oben

$$Q_j = -k_B T \frac{1}{\Upsilon} \frac{\partial \Upsilon}{\partial f_j} = -k_B T \frac{\partial}{\partial f_j} \ln \Upsilon$$

und

$$\langle \hat{H}_{\text{ges}} \rangle = E - \sum_{j=1}^L f_j Q_j = -k_B T^2 \frac{1}{\Upsilon} \frac{\partial \Upsilon}{\partial T}.$$

Zur Berechnung des Differentialen von  $E$  ist es nützlich, die *verallgemeinerte freie Enthalpie*

$$K = -k_B T \ln \Upsilon;$$

einzuführen. Für sie gilt ersichtlicherweise

$$\Upsilon = e^{-K/k_B T}.$$

Berechnen wir deren partielle Ableitungen nach  $T$  und  $f_j$ , so erhalten wir einerseits unmittelbar

$$\frac{\partial K}{\partial f_i} = Q_i,$$

andererseits mit Hilfe der Produktregel sowie (1.21)

$$\begin{aligned} \frac{\partial K}{\partial T} &= -k_B \ln \Upsilon - k_B T \frac{1}{\Upsilon} \frac{\partial \Upsilon}{\partial T} = \frac{1}{T} \left( K + E - \sum_{j=1}^L f_j Q_j \right) \\ &= -k_B \langle \ln \rho \rangle = -S. \end{aligned} \tag{1.25}$$

Für das Differential von  $K$  ergibt sich damit

$$dK = \frac{\partial K}{\partial T} dT + \frac{\partial K}{\partial f_i} df_i = -S dT - \sum_{j=1}^L Q_j df_j. \quad (1.26)$$

$E$  erhält man aus  $K$  nach (1.25) durch die Legendre-Transformation

$$E = K + TS + \sum_{j=1}^L f_j Q_j;$$

diese liefert

$$dE = dK + T dS + S dT + \sum_{j=1}^L f_j dQ_j + \sum_{j=1}^L Q_j df_j,$$

und das gewünschte Differential von  $E$  lautet wegen (1.26)

$$dE = T dS + \sum_{j=1}^L f_j dQ_j. \quad (1.27)$$

Diese Gleichung ist die Verallgemeinerung der Gibbsschen Fundamentalbeziehung für beliebige offene Systeme<sup>27</sup>. Sie ist wieder eine Summe aus Produkten, die jeweils aus einer intensiven Größe und dem Differential einer extensiven Größe bestehen; mit  $T \equiv f_0$  und  $S \equiv Q_0$  wird daraus

$$dE = \sum_{j=0}^L f_j dQ_j. \quad (1.28)$$

Dabei gilt

$$f_j = \frac{\partial E}{\partial Q_j}, \quad j = 0, 1, \dots, L.$$

Die Relation (1.27) läßt sich unmittelbar physikalisch interpretieren. Sie besagt im Sinne des ersten Hauptsatzes, daß die innere Energie des betrachteten Systems stets die Differenz aus zugeführter Wärme und verrichteter Arbeit ist. Der erste Summand auf der rechten Seite von (1.27) ist die die Wärmemenge  $\delta Q = T dS$ , was gleichzeitig den zweiten Hauptsatz beinhaltet; die restliche Summe enthält neben den chemischen Potentialen und den zugehörigen Teilchenzahlen weitere Paare aus extensiven und intensiven Größen, die jeweils Prozessen der Arbeitsverrichtung und den entsprechenden Energieformen zugeordnet werden können. Das kann beispielsweise die Form

$$\begin{aligned} dE = & T dS - p dV + \mu dN + \vec{F} d\vec{r} + \vec{v} d\vec{p} + \vec{\omega} d\vec{L} + D ds + \sigma dA \\ & + \phi_G dm + \phi_{el} dQ - \mu_m dB + \dots \end{aligned}$$

annehmen<sup>28</sup>. Dabei bedeuten im Einzelnen

<sup>27</sup>Wenn diese Gleichung einen Term mit chemischem Potential enthält, spricht man auch von der *Gibbs-Duhem-Gleichung*.

<sup>28</sup>Der vierte und der fünfte Summand sind natürlich physikalisch identisch und tauchen nur dann beide auf, wenn sie von unterschiedlichen Kräften kommen.



- $p$  der Druck und  $dV$  die Volumenänderung bei Ausdehnung beziehungsweise Kompression,
- $\mu$  das chemische Potential und  $dN$  die Änderung der Teilchenzahl,
- $\vec{F}$  eine äußere Kraft und  $d\vec{r}$  die dadurch bewirkte Verrückung,
- $\vec{\omega}$  der Axialvektor der Winkelgeschwindigkeit und  $d\vec{L}$  die Änderung des Drehimpulses,
- $D$  die Federhärte und  $ds$  die Längenänderung,
- $\sigma$  die Oberflächenspannung und  $dA$  die Oberflächenänderung,
- $\phi_G$  das Gravitationspotential und  $dm$  Massenzu- oder -abnahme,
- $\phi_{el}$  das elektrische Potential und  $dQ$  die Zu- oder Abnahme der elektrischen Ladung,
- $\mu_m$  das magnetische Moment und  $dB$  die Änderung der magnetischen Induktion,

und so weiter. Die Differentiale sind mit Ausnahme von dem der Entropie vorzeichenbehaftet; sie sind negativ, wenn sie zur Abnahme und positiv, wenn sie zur Zunahme der inneren Energie des Systems führen. Im ersten Fall wird am System Arbeit verrichtet, im zweiten verrichtet das System selbst Arbeit. Die zugehörigen Energieformen, die dabei ab- oder zugeführt werden, sind

- $E_V = p \Delta V$  (Kompressions- beziehungsweise Expansionsenergie<sup>29</sup>),
- $E_\mu = \mu N$  (Energie, die bei Änderung der Teilchenzahl ausgetauscht wird<sup>30</sup>),
- $E_F = \int \vec{F} d\vec{r}$  (Energie, die beim Verrichten von Arbeit gegen  $\vec{F}$  zugeführt wird),
- $E_{kin} = \frac{1}{2} m v^2$  (kinetische Energie),
- $E_{rot} = \frac{1}{2} I \omega^2$  (Rotationsenergie, dabei ist  $I = \int r^2 dm$  das Trägheitsmoment),
- $E_{Spann} = \frac{1}{2} D s^2$  (Spannenergie bei elastischen Systemen),
- $E_{Ob} = \sigma \Delta A$  (Oberflächenenergie),
- $E_G = m \phi_G$  (Lageenergie),
- $E_{el} = Q \phi_{el}$  (elektrische Energie),
- $E_{magn} = \mu_m B$  (magnetische Energie),

und so weiter.

<sup>29</sup>Hierbei handelt es sich natürlich um nichts anderes als um kinetische Energie der Teilchen, aus denen sich das System zusammensetzt.

<sup>30</sup>Das trifft auf alle Prozesse zu, bei denen sich die Stoffzusammensetzung ändert. Typische Beispiele hierfür sind chemische Reaktionen.

### 1.2.6 Anleihen aus der Strömungsmechanik

Eine Änderung der inneren Energie  $E$  des Systems tritt wie gesagt genau dann ein, wenn an dem System Arbeit verrichtet wird oder das System selbst Arbeit verrichtet. Entsprechend ist

$$P \equiv \frac{dE}{dt}$$

die *Leistung*, die dabei erbracht wird. Unter der Voraussetzung der zeitlichen Konstanz der Größen  $f_j$  gilt dabei gemäß der Kettenregel

$$\frac{dE}{dt} = \sum_{j=0}^L \frac{\partial E}{\partial Q_j} \frac{dQ_j}{dt}.$$

Die partiellen Ableitungen in diesem Ausdruck sind gerade die  $f_j$ , und somit liefert ein Vergleich mit (1.28)

$$\frac{dE}{dt} = \sum_{j=0}^L f_j \frac{dQ_j}{dt}.$$

Damit läßt sich die Gesamtleistung auf die zeitlichen Änderungen der einzelnen das System beschreibenden extensiven Größen zurückführen. Für obige Beispiele etwa ergibt sich daraus für die Gesamtleistung

$$P = T \frac{dS}{dt} - p \frac{dV}{dt} + \vec{F} \vec{v}_r + \vec{v} \vec{F}_p + \vec{\omega} \vec{M} + D v_s + \sigma \frac{dA}{dt} + \phi_G \frac{dm}{dt} + \phi_{el} \frac{dQ}{dt} + \mu_m \frac{dB}{dt} + \dots$$

Dabei ist  $\vec{F}_p = d\vec{p}/dt$  das zweite Newtonsche Axiom;  $\vec{M} = d\vec{L}/dt$  heißt Drehmoment.

Wenn sich extensive Größen zeitlich ändern, kann eine Änderung der Systemgröße der Auslöser sein, muß aber nicht. Es kann auch daran liegen, daß es sich nicht um eine Erhaltungsgröße handelt, oder aber, daß die Größe über eine Systemgrenze hinweg auf ein anderes System übertragen wird. Das deutet darauf hin, daß man sich extensive Größen in der thermodynamischen Näherung als *kontinuierlich über den Raum verteilt* vorstellen kann. Damit kann man ihnen im Sinne der Hydromechanik ein Geschwindigkeitsfeld  $\vec{v}(\vec{r}, t)$  zuordnen, mit dem sich ihre Verteilung im Raum ändert. Entsprechend kann man die zu einer extensiven Größe  $X$  gehörende Raumdichte

$$\rho_X = \lim_{\Delta V \rightarrow 0} \frac{\Delta X}{\Delta V} = \frac{dX}{dV}$$

und Stromdichte

$$\vec{j}_X = \rho_X \vec{v}_X$$

definieren<sup>31</sup>. Für den „Bestand“ von  $X$  in einer beliebigen Lebesgue-meßbaren Menge  $M \subseteq \mathbb{R}^3$  gilt dann

$$X = \iiint_M \rho_X dV$$

<sup>31</sup>Der Limes erinnert nachdrücklich daran, daß hier die Näherung einer kontinuierlichen Verteilung der Größe  $X$  verwendet wird.

und für die Stromstärke

$$I_X = \iint_{\partial G} \vec{j}_X d\vec{A}.$$

Ist  $X$  eine Erhaltungsgröße, so folgt für jedes geschlossene Raumgebiet  $G \subset \mathbb{E}^3$

$$\oiint_{\partial G} \vec{v}_X d\vec{A} = 0;$$

ist  $X$  keine Erhaltungsgröße, so ist die Schüttung aller Quellen und Senken innerhalb von  $G$  gegeben durch

$$\Sigma_X(G) = \oiint_{\partial G} \vec{v}_X d\vec{A},$$

und die lokale Quelldichte ist der Grenzwert

$$\sigma_X(\vec{r}) = \lim_{\Delta V \rightarrow 0} \frac{\Delta \Sigma_X}{\Delta V} = \lim_{\Delta V \rightarrow 0} \frac{1}{\Delta V} \oiint_{\partial V} \vec{v}_X d\vec{A} \equiv \operatorname{div} \vec{v}_X.$$

Der Bestand von  $X$  im (beliebigen) geschlossenen Gebiet  $G$  wird durch  $I_X$  und  $\Sigma_X$  bestimmt; eine Bilanz liefert folglich

$$\frac{dX}{dt} = I_{X,\text{raus}} - I_{X,\text{rein}} + \frac{d\Sigma_X}{dt}$$

oder (Vertauschen von Differentiationen und Integrationen ist bei stetigen Integranden erlaubt)

$$\iiint_G \frac{\partial \rho_X}{\partial t} dV = \oiint_{\partial G} \vec{j}_X d\vec{A} + \oiint_{\partial G} \vec{v}_X d\vec{A}$$

und nach Anwendung des Gaußschen Integralsatz auf die rechte Seite

$$\iiint_G \frac{\partial \rho_X}{\partial t} dV = \iiint_G \operatorname{div} \vec{j}_X dV + \iiint_G \sigma_X dV.$$

Daraus folgt die *Kontinuitätsgleichung*

$$\frac{\partial \rho_X}{\partial t} = \operatorname{div} \vec{j}_X + \sigma_X.$$

Nimmt man die elektrische Ladung als Beispiel, so erhält man die bekannten Gleichungen

$$\frac{\partial \rho}{\partial t} = \operatorname{div} \vec{j} \quad \text{und} \quad \frac{dQ}{dt} = I.$$

Man kann zeigen<sup>32</sup>, daß stets dann, wenn der zu einer extensiven Größe  $X$  gehörende Strom  $I_X$  dissipativ, also entropieerzeugend wirkt, dieser mit dem Gradienten der zugehörigen intensiven Größe  $\xi$  verbunden ist. Im einfachsten Fall gilt dabei

$$\vec{j}_X \propto \operatorname{grad} \xi; \tag{1.29}$$

<sup>32</sup>Siehe beispielsweise [488].

das bedeutet, daß die räumliche Richtung des steilsten Anstiegs beziehungsweise Abstiegs von  $\xi$  die Richtung des Stromdichtevektors festlegt.

Man darf dabei jedoch nicht vergessen, daß die beschriebene hydromechanische Analogie und die damit untrennbar verbundene Annahme einer kontinuierlich räumlich verteilten Natur der extensiven Größen eine rein formale Angelegenheit ist und insbesondere eine *Näherung* und damit ein Modell darstellt, daß man auf keinen Fall überbeanspruchen darf.

Es folgt nun eine kritische Bestandsaufnahme dessen, was im Rahmen des Karlsruher Physikkurses aus den geschilderten physikalischen Sachverhalten gemacht wird. Dabei wird sich zeigen, daß die Probleme, die hier auftreten, ausnahmslos in der dabei verwendeten Interpretation und didaktischen Umsetzung für den Schulunterricht liegen; die physikalische Richtigkeit der oben beschriebenen thermodynamischen Sachverhalte steht nicht zur Diskussion.

## 1.3 Zwölf Argumente gegen den Karlsruher Physikkurs

Die Schwierigkeiten, die wir im Karlsruher Physikkurs sehen, lassen sich in drei Kategorien einteilen, eine physikalische, eine wissenschaftsphilosophische und eine didaktische. Die erste Kategorie beinhaltet im wesentlichen eine Überstrapazierung der Gleichgewichtsthermodynamik weit über deren Anwendbarkeitsbereich hinaus sowie weitere überzogene Verallgemeinerungen, die zweite eine Auffassung von Aufbau und Interpretation der Physik, die aus unserer Sicht nicht haltbar ist, und die dritte eine fragwürdige Zielsetzung und eine ebensolche Strategie des Physikunterrichts. Anstelle einer Einführung in die Grundideen des Karlsruher Physikkurses stellen wir dieselben in den folgenden Abschnitten kurz vor<sup>33</sup>, um sie danach jeweils gleich an Ort und Stelle kritisch zu hinterfragen.

### 1.3.1 Physikalische Aspekte

#### 1.3.1.1 Die Bedeutung der Thermodynamik

Unser erstes Argument betrifft wie bereits angedeutet den Gültigkeitsbereich der Grundlage des Karlsruher Physikkurses.

***Argument 1: Der Karlsruher Physikkurs verwendet die Gleichgewichtsthermodynamik in einer weit über deren Gültigkeitsbereich hinausgehenden Art und Weise.***

Die Grundlage des Karlsruher Physikkurses läßt sich grob in folgender Aussage subsumieren: Die Thermodynamik ist die Fundamentaldisziplin der Physik und bestimmt das dynamische Verhalten realer Systeme. Als Rechtfertigung dafür wird stets die verallgemeinerte Gibbssche Fundamentalbeziehung herangezogen. Deren rechte Seite macht in

---

<sup>33</sup>Lehrbücher zum Karlsruher Physikkurs für die Sekundarstufe I und II sind bei Aulis erschienen [225] - [227], [229] - [232]; vergleiche auch [217]. Eine kurze Darstellung des Konzepts liefern [224] und [228], weitere Informationen findet man in [413].

der Tat einen absolut symmetrischen Eindruck und scheint eine einheitliche Betrachtung völlig unterschiedlicher Teildisziplinen der Physik zu ermöglichen. Die Welt scheint komplett in extensive und intensive Größen einteilbar zu sein. Zur Beschreibung der Dynamik physikalischer Systeme wird anschließend die in Abschnitt 1.2.6 beschriebene hydromechanische Analogie herangezogen; alles, was physikalisch passieren kann, ist demnach beschreibbar als Strömung einer oder mehrerer extensiver Größen, wobei diese Strömung stets längs des Gefälles der zugehörigen intensiven Größe zu erfolgen hat. Die intensiven Größen werden folglich durchweg als Potentiale gedeutet, die die Strömungen bestimmen. Aufgabe der Physik ist es dann lediglich, Bilanzgleichungen für alle möglichen Stromstärken aufzustellen.

Doch die Harmonie täuscht. Bei aller scheinbaren Eleganz und Geschlossenheit dieses Modells dürfen nämlich seine Grenzen nicht übersehen werden. Zuerst einmal beschreibt es nur Systeme, die sich *im thermischen Gleichgewicht* mit ihrer Umgebung befinden. Nur dann sind die angeblich fundamentalen Größen  $S$  und  $T$  thermodynamisch überhaupt definiert, und nur dann ist die Energie eine eindeutige Funktion der extensiven Größen  $S, V, N$  und dergleichen, was für die Gültigkeit der Gibbschen Fundamentalbeziehung zwingend erforderlich ist, denn für Nichtgleichgewichtszustände und irreversible Prozesse kann auch  $T dS > \delta Q$  gelten. Vor allen Dingen aber handelt es sich dabei um eine *thermodynamische* Betrachtungsweise, das heißt, sämtliche dabei auftauchende extensive Größen sind *makroskopische Mittelwerte* über Größen, deren mikroskopisches Verhalten sehr viel komplizierter ist, im allgemeinen zu kompliziert für eine analytische Beschreibung, weswegen man überhaupt auf die Verlegenheitslösung der Mittelwertbildung angewiesen ist. Außerdem setzt die Verwendung von Differentialen eigentlich eine kontinuierliche Struktur der zu differenzierenden Größen voraus, die keineswegs stets vorliegt. Die wichtigsten Gegenbeispiele bilden Masse, elektrische Ladung und Teilchenzahl, und auch sonst ist die kontinuierliche Natur häufig nur als klassische Näherung gegeben. Dieser letzte Aspekt geht in seiner Bedeutung über die durch die Thermodynamik vorgegebenen engen Grenzen des Modells weit hinaus und wird gleich noch in einem eigenen Abschnitt weiter präzisiert. Doch selbst wenn man, berechtigt oder nicht, eine solche kontinuierliche Struktur einfach voraussetzt, bleibt das Problem, daß bei der Herleitung der Gibbschen Fundamentalbeziehung natürlich stillschweigend ausreichende Differenzierbarkeitseigenschaften der dabei auftretenden extensiven Größen vorausgesetzt werden. Davon kann jedoch in den Bereichen, in welchen der Karlsruher Physikkurs die Gibbsche Relation anwenden will, häufig nicht die Rede sein. Oft sind die betrachteten Funktionen nicht einmal stetig; man denke etwa an den Impuls bei mechanischen Reflexionen oder generell an extensive Größen bei Phasenübergängen erster Art. Damit verliert der Karlsruher Physikkurs in solchen Fällen seine mathematische Grundlage, die zwar sicherlich für den gymnasialen Unterricht keine direkte Bedeutung hat, aber dennoch stets gewährleistet sein muß.

Unter diesen Gesichtspunkten erscheint die Beförderung der Thermodynamik zur Fundamentaldisziplin als gewaltsames Hineinpressen aller physikalischer Teildisziplinen in einen Rahmen, der hierfür absolut ungeeignet und viel zu beschränkt ist. Der Versuch, die gesamte Physik auf einer Grundlage aufzubauen, die sich bei näherer Betrachtung als phänomenologische Näherung zur Umgehung der für praktische Zwecke unhandlichen Beschreibung der fundamentalen mikroskopischen Prozesse erweist, muß von vorn

herein schiefehen<sup>34</sup>. Das ist, wie wir noch sehen werden, schon bei Beschränkung auf elementare Mittelstufenphysik unangemessen, obwohl dort Konzepte der modernen Physik weitgehend ohne Bedeutung sind<sup>35</sup>. Was dabei überdies völlig unter den Tisch fällt, ist der mit der Quantenmechanik und mehr noch der Quantenfeldtheorie sowie deren offensichtlichen Universalität nicht mehr wegzudiskutierende Vorrang der *mikroskopischen Beschreibung* der Natur durch die Physik und die damit einhergehende, im Detail nach wie vor unbeantwortete Frage, wie man von dieser zu einer konsistenten makroskopischen Beschreibung gelangen kann. Der Karlsruher Physikkurs ignoriert das und schließt somit wesentliche Teile der Physik von vorneherein aus. Inwieweit das trotz der inzwischen erfolgten Übernahme der Quantenmechanik in dessen Stoffplan dennoch weiterhin der Fall ist, wird ebenfalls in einem eigenen Abschnitt zu erörtern sein.

### 1.3.1.2 Energieträger oder Energieformen?

Für besonders kritikwürdig halten wir den Umgang des Karlsruher Physikkurses mit dem Begriff der Energie.

***Argument 2: Der Karlsruher Physikkurs vermittelt die Vorstellung der Energie als einer Substanz, die einer Flüssigkeit gleich durch die Gegenstände fließt und außerdem dabei von irgendwelchen Trägern huckepack transportiert wird.***

Eine auf den ersten Blick nur geringfügige Abweichung des Karlsruher Physikkurses vom herkömmlichen Zugang erweist sich auf den zweiten Blick schnell als äußerst weitgehend. Es soll davon Abstand genommen werden, Energie als abgeleitete Größe zu betrachten, die in unterschiedlichen *Energieformen* vorkommt, welche sich ineinander umwandeln können; stattdessen wird die Energie als nicht weiter hinterfragbare *Grundgröße* betrachtet, als ein „Etwas“, das zum Bewegen von Gegenständen, zum Erhöhen von Temperaturen, zum Betrieb elektrischer Geräte, kurz gesagt eigentlich für so ziemlich alles notwendig ist. Dabei wird betont, daß es sich bei diesem „Etwas“ in allen denkbaren Situationen um das selbe handelt. Gerechtfertigt wird das mit der angeblich bei Schülerinnen und Schülern intuitiv bereits vorhandenen Vorstellung von diesem „Etwas“. Die unterschiedlichen Erscheinungsformen der Energie werden mit Hilfe eines Energieträger-Konzepts erklärt [160], [450], [414], [418], und was bisher als Umwandlung von einer Energieform in eine andere gedeutet wurde, erfährt nun seine Interpretation als Umladvorgang von einem Energieträger auf einen anderen im Rahmen von Strömungen der Energieträger zusammen mit ihrer Fracht, der Energie. Dabei wird betont, daß ein solches naives Rucksackmodell zunächst unwidersprochen zu akzeptieren ist und erst nach und nach verfeinert werden soll [418]. Von Anfang an steht dabei die Vorstellung im Mittelpunkt, daß die Energie ein „Etwas“ ist, das strömt. Entsprechend wird dann die *Energiestromstärke* eingeführt als vorbeiströmende Energiemenge pro Zeit. Daß damit natürlich nichts anderes als der innerhalb wie außerhalb der physikalischen Sprechweise etablierte Begriff der *Leistung* gemeint ist, wird nicht weiter thematisiert.

<sup>34</sup>Janez Strnads Vergleich dieser Vorgehensweise mit derjenigen des Prokrustes trifft die Sache ziemlich genau [498].

<sup>35</sup>Eine Ausnahme bildet hier das Teilchenmodell. Siehe dazu Abschnitt 1.3.3.1 weiter unten.



Die Vorstellungen von Energieträgern und strömender Energie haben zusätzlich zu allem, was im vorigen Abschnitt bereits beschrieben wurde, die schwerwiegende Konsequenz, daß die Energie den Charakter einer diffus bis gar nicht definierten *Substanz* bekommt<sup>36</sup>. Daß so etwas dem abstrakten Begriff der Energie nicht nur nicht gerecht wird, sondern diesem geradezu entgegensteht, ist offensichtlich. Einige Beispiele mögen das verdeutlichen.

Schon einfachste mechanische Konfigurationen sind mit einer Substanzvorstellung der Energie unvereinbar. Man denke etwa an einen Körper, der in einem Gravitationsfeld angehoben wird. Dabei steigt seine Lageenergie<sup>37</sup>, aber es ist nichts, aber auch gar nichts substantielles in den Körper eingedrungen, und er unterscheidet sich in keiner Weise von einem identischen Körper auf geringerer Höhe. Er kann an Ort und Stelle mit seiner neu gewonnenen Energie auch überhaupt nichts anfangen. Die Aussage, seine Lageenergie sei größer geworden, bezieht sich lediglich auf die möglichen Ereignisse, die passieren können, wenn er wieder herunterfällt. Von einer in den Körper geströmten Substanz kann keinerlei Rede sein. Insbesondere ist das auch nicht mit dem Sachverhalt der völlig willkürlichen Festlegbarkeit des Nullpunkts der Lageenergie in Einklang zu bringen. Wichtiger noch ist die Tatsache, daß die kinetische Energie ebenfalls bezugssystemabhängig ist. Ein Körper kann je nach verwendetem Inertialsystem keine oder irgendeine nichtverschwindende kinetische Energie besitzen<sup>38</sup>.

Ein weiteres Beispiel ist eine Flüssigkeit unter Druck. Ist Energie eine Substanz, die von Energieträgern transportiert wird, dann führt das zur Schlußfolgerung, daß Wasser, welches unter Druck gesetzt wird, mit Energie „vollgepumpt“ wird und das Volumen der Energieträger ist [450]. Das ist jedoch ein Trugschluß; Wasser unter Druck trägt für sich allein keine Energie, denn im Ausdruck  $dE = T dS - p dV$  steht der Summand  $p dV$  für die zur Druckerhöhung erforderliche mechanische Arbeit, also beispielsweise bei einem Kolben mit Stempelfläche  $A$  und einer *Verschiebung* des Stempels um  $ds$  für

$$dW = p dV = p A ds,$$

womit deutlich wird, daß dabei der *Prozeß* der Volumenänderung und damit derjenige der Druckerhöhung im Vordergrund steht und nicht der unter anderem durch Volumen

<sup>36</sup>Diese Vorstellung ist keine etwaige Fehlinterpretation des Energiebegriffs des Karlsruher Physikkurses, sondern wird dort mindestens in Kauf genommen, wenn nicht sogar angestrebt; siehe beispielsweise [450] oder [418]. In [418] wird ein solcher Energiebegriff gleichzeitig zur Grundlage einer anzustrebenden gemeinsamen Fachsprache aller Naturwissenschaften im gymnasialen Unterricht erklärt. Ähnliches ist vom Kultusministerium in Stuttgart in einer Veröffentlichung des Landesinstituts für Schulentwicklung zu lesen [45], wobei dort interessanterweise auf die Gefahr eines naiven Rucksackmodells hingewiesen und die herkömmliche Vorgehensweise als Alternative offen gehalten wird.

<sup>37</sup>Hierfür scheint sich im gymnasialen Physikunterricht gegenwärtig der absurde Ausdruck „Höhenenergie“ breitzumachen. Dafür ist der Karlsruher Physikkurs sicherlich nicht verantwortlich zu machen; dort gibt es ja auch keine *Energieformen*. Es handelt sich dabei jedoch um ein weiteres inakzeptables Beispiel der Erfindung neuer, willkürlicher Begriffe durch die Schulphysik.

<sup>38</sup>Ein elegantes Beispiel hierfür stammt von J. W. Warren [514]. Er betrachtet zwei Körper A und B, die sich relativ zueinander gleichförmig mit 2 m/s bewegen, und einen dritten Körper C der Masse 2 kg, der sich von A und B gleichförmig mit 1 m/s wegbewegt und folglich für beide die kinetische Energie 1 J besitzt. Wird C durch Wechselwirkung mit einem weiteren Körper D bezüglich A zur Ruhe gebracht, ist seine kinetische Energie von A aus gesehen 0 J, von B aus gesehen jedoch 4 J. Energie ist eben keine Substanz, sondern eine abstrakte Größe, die in diesem Fall erlaubt, die Arbeit zu berechnen, die jeweils notwendig ist, um C in Bezug auf A oder B zum Anhalten zu bringen.

und Druck festgelegte *Zustand* des Wassers. Erst wenn das Wasserreservoir in die Lage versetzt wird, selbst Arbeit zu verrichten, beispielsweise wenn man es von links mit einem Stempel kleiner Fläche komprimiert und es dadurch nach rechts einen Stempel großer Fläche verschieben kann, wird die Energie abrufbar<sup>39</sup>. Da das an jeder Stelle möglich ist, ist die Energie im Wasser nicht lokalisierbar, wie es bei einer Substanz der Fall sein müßte. Vergleichbares gilt erst recht für Gase. Die Energie-Substanz- und die Energieträger-Vorstellung des Karlsruher Physikkurses suggerieren, daß ein Gas, das unter Druck gesetzt wird, dadurch gleichzeitig mit der Substanz Energie vollgepumpt wird. Das ist jedoch eine physikalisch sinnlose Aussage, da die innere Energie eines idealen Gases

$$E = \frac{3}{2} N T$$

unabhängig von dessen Druck ist. Bei realen Gasen wird es noch deutlicher; um das zu sehen, berechnen wir die innere Energie eines Van der Waals-Gases, indem wir die partiellen Ableitungen derselben nach  $T$  und nach  $V$  bestimmen und anschließend integrieren. Das geschieht am einfachsten über den Umweg der partiellen Ableitungen der Entropie. Wir betrachten dazu die Van der Waalssche Zustandsgleichung

$$\left(p + \frac{a}{V^2}\right) (V - b) = N k_B T$$

( $a$  ist der Kohäsionsdruck und  $b$  das Eigenvolumen des Gases). Daraus folgt mit der Gibbsschen Fundamentalbeziehung

$$\begin{aligned} dS &= \frac{1}{T} \left[ dE + \left( \frac{N k_B T}{V - b} - \frac{a}{V^2} \right) dV \right] \\ &= \frac{1}{T} \left[ \left( \frac{\partial E}{\partial T} \right)_V dT + \left( \frac{\partial E}{\partial V} \right)_T dV + \left( \frac{N k_B T}{V - b} - \frac{a}{V^2} \right) dV \right] \end{aligned}$$

und weiter durch Vergleich mit

$$dS = \left( \frac{\partial S}{\partial T} \right)_V dT + \left( \frac{\partial S}{\partial V} \right)_T dV$$

für die partiellen Ableitungen der Entropie

$$\begin{aligned} \left( \frac{\partial S}{\partial T} \right)_V &= \frac{1}{T} \left( \frac{\partial E}{\partial T} \right)_V, \\ \left( \frac{\partial S}{\partial V} \right)_T &= \frac{1}{T} \left( \frac{\partial E}{\partial V} \right)_T + \frac{N k_B}{V - b} - \frac{a}{T V^2}. \end{aligned}$$

Die zweiten Ableitungen von  $S$  sind

$$\begin{aligned} \frac{\partial^2 S}{\partial T \partial V} &= \frac{1}{T} \frac{\partial^2 E}{\partial T \partial V}, \\ \frac{\partial^2 S}{\partial V \partial T} &= \frac{1}{T} \frac{\partial^2 E}{\partial V \partial T} - \frac{1}{T^2} \left( \frac{\partial E}{\partial V} \right)_T + \frac{a}{T^2 V^2}, \end{aligned}$$

<sup>39</sup>Das läßt sich mit einem ebenfalls von Warren beschriebenen Zahlenbeispiel belegen [514]. Wasser hat ein Kompressionsmodul von circa  $2,08 \cdot 10^9$  Pa, folglich ist zur Kompression von  $1 \text{ m}^3$  Wasser von 0 auf 1 bar Arbeit vom Betrag 2,5 J erforderlich. Beim Durchlaufen einer Druckdifferenz von 1 bar in einer hydraulischen Maschine liefert die gleiche Wassermenge mit  $10^5$  J die vierzigtausendfache Energie, was die Vorstellung von Wasser als Energieträger ad absurdum führt.



und wegen

$$\frac{\partial^2 S}{\partial T \partial V} = \frac{\partial^2 S}{\partial V \partial T}$$

ergibt das einerseits

$$\left(\frac{\partial E}{\partial V}\right)_T = \frac{a}{V^2}. \quad (1.30)$$

Andererseits gilt für ein ideales Gas und damit auch für ein Van der Waals-Gas

$$\left(\frac{\partial E}{\partial T}\right)_V = \frac{3}{2}N. \quad (1.31)$$

Die innere Energie erhält man durch Integration, genauer gesagt gilt

$$E = \int \left(\frac{\partial E}{\partial V}\right)_T dV + \int \left(\frac{\partial E}{\partial T}\right)_V dT,$$

und mit (1.30) und (1.31) liefert das

$$E = \frac{3}{2}NT - \frac{a}{V}.$$

Daraus liest man ab, daß ein reales Gas unter höherem Druck sogar etwas weniger innere Energie enthält als das selbe Gas unter geringerem Druck. Das läßt keinerlei Spielraum für irgendwelche Substanzinterpretationen. Die innere Energie ist nicht mehr und nicht weniger als ein Maß für die gesamte kinetische Energie der ungeordneten Bewegungen der Gasmoleküle und damit ein Ausdruck der mikroskopischen Dynamik des Systems.

Sehr schön sieht man das auch an einem zuerst von Robert Emden beschriebenen und später von Arnold Sommerfeld präzisierten Beispiel von unmittelbarer praktischer Relevanz, nämlich dem Betrieb einer Heizung in einem Zimmer [143], [488]. Unter der Annahme gleichen Luftdrucks innen und außen gilt für die Zunahme der inneren Energie der Zimmerluft bei Erhöhung der Temperatur von  $T_0$  auf  $T$

$$U - U_0 = c_v m (T - T_0),$$

wobei  $m$  die Masse und  $c_v$  die spezifische Wärmekapazität der Luft bei konstantem Volumen ist. Unter Verwendung der Zustandsgleichung  $pV = NRT$  und der mittleren Molmasse  $\mu$  der Luft erhält man daraus für die Zunahme der inneren Energie beim Heizen

$$U - U_0 = \frac{c_v \mu p V}{R} - m c_v T_0$$

Ist  $\rho$  die Massendichte der Luft, so folgt für die Energiedichte im geheizten Zimmer

$$u = \frac{c_v \mu p}{R} + \rho \left( \frac{U_0}{m} - c_v T_0 \right),$$

und erneutes Einsetzen der Zustandsgleichung, diesmal in der Form  $\rho T = \mu p / R$ , liefert

$$u = \frac{\mu p}{R} \left[ c_v + \frac{1}{T} \left( \frac{U_0}{m} - c_v T_0 \right) \right].$$

Der Klammerausdruck ist aufgrund der großen Verdampfungswärme der Luft positiv, und das bedeutet, daß die Energiedichte durch die Heizung abnimmt. Auch hier versagt die Vorstellung, man pumpe etwas in das Zimmer hinein.

Ist die Interpretation der Energie als einer Substanz inkonsistent, so funktioniert auch das Energieträger-Konzept nicht mehr, denn abstrakte Begriffe können nicht transportiert werden. Aber selbst wenn man davon absieht, findet man sich mit einem unlösbaren Problem konfrontiert. Denn häufig wandern Energieträger und Energie mit unterschiedlicher Geschwindigkeit; wir werden im nächsten Abschnitt ein konkretes Beispiel dafür kennenlernen. Damit erscheint die Interpretation der extensiven Größen als Energieträger völlig willkürlich. Vor allem aber gibt es Situationen, in denen das Bild der von Energieträgern transportierten Energie-Substanz überhaupt nicht funktioniert. Das sieht man exemplarisch beim vollständig inelastischen Stoß. Nehmen wir an, ein Körper  $A$  mit Geschwindigkeit  $v_A$  und Masse  $m_A$  trifft auf einen Körper  $B$  mit Geschwindigkeit  $v_B$  und Masse  $m_B$ , dann bewegen sich diese aufgrund der Inelastizität nach dem Stoß mit einer gemeinsamen Geschwindigkeit  $u$ , für die aus dem Impulserhaltungssatz

$$m_A v_A + m_B v_B = (m_A + m_B) u$$

die Relation

$$u = \frac{m_A v_A + m_B v_B}{m_A + m_B}$$

folgt. Für die kinetische Energie vorher und nachher gilt

$$\frac{1}{2} (m_A v_A^2 + m_B v_B^2) > \frac{1}{2} (m_A + m_B) u^2,$$

da bei dem inelastischen Stoßprozeß Bewegungsenergie dissipiert wird. Interpretiert man nun den Impuls als Träger der Energie, so muß man zu dem im Rahmen dieses Bildes absurden Schluß kommen, daß er sich zwar nach Maßgabe des Massenverhältnisses der beiden Körper verlustfrei auf diese verteilt, dabei aber gleichzeitig einen Teil der von ihm getragenen Energie verliert. Warren beschreibt in [514] ein weiteres sehr überzeugendes Beispiel in Gestalt eines Permanentmagneten. Wäre Energie eine Substanz, die transportiert wird, so müßte die Wirkung des Magneten nachlassen, wenn er Kräfte auf ferromagnetische Körper ausübt und diese dadurch in Bewegung versetzt, da dann scheinbar Energie vom Magneten auf diese Körper umgeladen wird.

All das bestätigt nachdrücklich, was eigentlich schon von vorn herein klar ist: Die Energie eben gerade nicht das einfache begriffliche Konzept, das von Schülerinnen und Schülern intuitiv schon mitgebracht wird, wie es von den Vertretern des Karlsruher Physikkurses dargestellt wird. Wie wäre es sonst zu erklären, daß ein korrektes Verständnis dieser Größe bis zu Beginn des 20. Jahrhundert noch nicht vorlag, weil ein solches in einem jahrhundertelangen Prozeß unter Beteiligung der jeweils bedeutendsten Physiker der jeweiligen Zeit erst langsam entstehen mußte?<sup>40</sup> Wir plädieren daher nachdrücklich für eine Beibehaltung des Konzepts der unterschiedlichen Energieformen, die sich ineinander umwandeln können. Der Begriff der Energie kann dabei wie beschrieben mit Hilfe der mechanischen Arbeit, die (nicht nur) dazu natürlich integraler Bestandteil des Physikunterrichts bleiben muß, operational *definiert* werden, und Energie bleibt die abstrakte

<sup>40</sup>Die Verwirrung der Begriffe der Energie und der Kraft, die man noch im späten 19. Jahrhundert selbst in Lehrbüchern findet, ist eine bekannte Folge dieses Prozesses.

Hilfsgröße, die sie ist, ohne den falschen Anschein einer realen Substanz zu gewinnen, die darüber hinaus rätselhaft bleibt und nicht oder nur schemenhaft erklärt wird. Insbesondere erspart man sich dabei die Mühe, sich zwangsweise für jede Energieform mit einem teilweise äußerst abstrakten und schwer verständlichen Energieträger herumschlagen zu müssen, der häufig bei genauerem Hinsehen, als es die Schulphysik möglich macht, physikalisch nicht haltbar ist.

Ein weiterer, mehr methodischer Aspekt sollte nicht unerwähnt bleiben. Dieser betrifft das Durcheinander von konkreten und abstrakten Ideen, das mit der Idee der Energieträger unvermeidlicherweise einher geht. Denn die vermeintlichen Energieträger sind in bunter Mischung real existierende Objekte (Substanzen!) und abstrakte Begriffe, Skalare, Vektoren und dergleichen. Die Folge ist unweigerlich, daß die Beschreibung der einfachsten Fälle, wie etwa strömendes Wasser, auf andere beliebig abstraktere Situationen so dauer- wie fehlerhaft übertragen wird. Unterschätzt man dieses Problem, so überschätzt man damit gleichzeitig das Abstraktionsvermögen von Schülerinnen und Schülern in der Mittelstufe, und das gilt nicht nur für die Energie, die nur ein Beispiel einer abstrakten physikalischen Größe ist, die im Karlsruher Physikkurs eine Interpretation als Substanz verkraften muß. Wir kommen ausführlich darauf zurück.

### 1.3.1.3 Fehlvorstellungen zur Entropie

Große Bedeutung weist der Karlsruher Physikkurs einer möglichst frühzeitigen Einführung der Entropie zu, was nicht ohne Folgen bleibt.

***Argument 3: Der Karlsruher Physikkurs vermittelt eine Vorstellung des Begriffs der Entropie als einer Substanz und nimmt dabei in Kauf, daß Entropie für Wärme gehalten wird.***

Hier haben wir es genaugenommen mit einem speziellen Beispiel des im vorigen Abschnitt diskutierten Sachverhalts zu tun, mit einem sehr bedeutsamen Beispiel allerdings, so bedeutsam, daß das einen eigenen Abschnitt rechtfertigt. Wir wollen dabei die sehr berechnete Frage nach der Sinnfälligkeit der Behandlung eines Begriffes wie dem der Entropie schon in der Unter- oder Mittelstufe des Gymnasiums aussparen und hierzu nur die Bemerkung anfügen, daß die Einführung der Entropie in die Physik und deren grundlegende Erforschung durch Clausius, Boltzmann und Nernst sicher nicht zufällig eine der letzten großen klassischen Errungenschaften vor dem Auftauchen der Quantenmechanik darstellt und ihre Bedeutung erst im Rahmen letzterer voll erfaßt werden konnte.

Die Vorstellung der Entropie als Energieträger führt zu all den Problemen, die auch für den Energiebegriff auftreten. Entropie erscheint als eine geheimnisvolle Substanz, welche durch alle Objekte wabert; ihr Charakter als abstrakte Größe, der in den Abschnitten 1.2.3 und 1.2.4 beschrieben wurde, bleibt völlig auf der Strecke. Die Schwierigkeiten sind hierbei schon in der unzutreffenden Voraussetzung angelegt, daß die statistische und die thermodynamische Definition der Entropie völlig äquivalent seien, womit unter anderem die *rein thermodynamische* Behandlung des Entropiebegriffs durch den Karlsruher Physikkurs gerechtfertigt wird. Davon kann jedoch keine Rede sein, denn die statistische Definition ist allgemeiner als die thermodynamische. Letztere setzt den Begriff der Wärme voraus und ist nur in Systemen im (oder nahe beim) Gleichgewicht sinnvoll. Erstere

ist mikroskopisch deutbar, kommt ohne den Wärmebegriff aus und ist in gewissen Grenzen auch für Systeme, die sich nicht im Gleichgewicht befinden, eine sinnvoll Größe. Nur wenn der statistische Entropiebegriff im Vordergrund steht, kann man der fundamentalen Bedeutung, die der Entropie in der Physik makroskopischer Objekte zukommt, auch im Unterricht gerecht werden. Eine wirklich allgemeine Entropiedefinition für Systeme fern vom Gleichgewicht ist im übrigen bis jetzt noch unbekannt.

Unabhängig davon sollte jedoch ganz allgemein unstrittig sein, daß eine Interpretation der Entropie als einer strömenden Substanz selbst im Sinne eines Modells angesichts des statistischen Charakters dieser Größe inakzeptabel ist. Auf die Fehlvorstellungen, die man ganz allgemein mit solchen Strömungs- und damit Kontinuumsphysikkonzepten bei Schülerinnen und Schülern provoziert, kommen wir später ausführlich zurück. Darüberhinaus ist das Konzept des Energieträgers Entropie auch physikalisch überhaupt nicht durchzuhalten. Erstens macht das Bild der Entropieströmungen wesentlich von einer Erhaltungseigenschaft der Entropie Gebrauch, die jedoch nur für reversible Energietransport- und Energieumwandlungsprozesse gegeben ist; bei irreversiblen Prozessen ist die Entropie, wie wir gesehen haben, keine Erhaltungsgröße<sup>41</sup>. Zweitens führt der Plan, die Entropie zum Träger der thermischen Energie zu ernennen, schnell auf Widersprüche. Ein simples Beispiel illustriert das sehr deutlich, nämlich das der Wärmeleitung in Festkörpern. Bei dieser gilt für die Leistung und die Entropiestromstärke die allgemeine Relation

$$P = T I_S.$$

Wärmeleitung erfolgt natürlich unter Beachtung des Energieerhaltungssatzes; da sie längs eines Temperaturgefälles  $\Delta T = T_2 - T_1$  erfolgt, muß die Entropiestromstärke immer größer werden. Die hierbei erbrachte Leistung ist

$$P = -\Delta T I_S(T_1)$$

und die Entropieproduktionsrate folglich

$$\Sigma_S = \frac{1}{T_2} \frac{dQ}{dt} = \frac{P}{T_2} = \frac{(T_1 - T_2) I_S(T_1)}{T_2}.$$

Die Entropiestromstärke vergrößert sich somit auf

$$I_S(T_2) = I_S(T_1) + \Sigma_S = \frac{T_1 I_S(T_1)}{T_2}.$$

Für die Leistung (die „Energiestromstärke“) gilt gleichzeitig

$$P(T_2) = T_2 I_S(T_2) = T_1 I_S(T_1) = P(T_1),$$

das heißt, diese ist konstant. Wäre die Entropie der Träger der Energie, würde sie ihrem Transportgut offensichtlich davonlaufen. Das vermeintlich so anschauliche Bild des mit Energie beladenen Entropiestroms funktioniert hier nicht.

Einige Protagonisten des Karlsruher Physikkurses gehen wie gesagt noch einen drastischen Schritt weiter und empfehlen, Entropie kurzerhand mit Wärme zu identifizieren.

<sup>41</sup>Auf diese Problematik des Karlsruher Entropie-Begriffs haben Backhaus und Schlichting bereits in den achtziger Jahren hingewiesen, fanden aber leider nur wenig Beachtung. Siehe dazu [35].

Daß das aus physikalischer Sicht Unsinn ist, braucht eigentlich gar nicht erwähnt zu werden. Man sieht das beispielsweise an einer Wärmekraftmaschine, die mechanische Arbeit verrichtet. Die nutzbare mechanische Arbeit ist  $W = Q_1 - Q_2 = S(T_1 - T_2)$ . Die Entropie bleibt erhalten oder nimmt höchstens zu, die verfügbare Wärme nimmt ab, da sie teilweise in mechanische Arbeit umgewandelt wird. Ein sich irreduzibel ins Vakuum ausdehnendes Gas liefert ein weiteres Beispiel. Nach Gibbs gilt dabei

$$dE = T dS - p dV$$

und folglich wegen  $dE = 0$

$$dS = \frac{p}{T} dV > 0,$$

das heißt, seine Entropie nimmt zu, die Temperatur und damit die Wärmemenge bleiben jedoch gleich. Auch der Versuch, die Sache zu retten, indem man stattdessen den Ausdruck  $\int T dS$  mit Wärme gleichsetzt, funktioniert nicht, wie man am selben Beispiel sehen kann. Denn dort gilt

$$T dS = 0,$$

und das würde bei einer Interpretation von  $T dS$  als Wärme wegen der Expansion des Gases auf  $dE = -p dV < 0$  führen, im Widerspruch zur Thermodynamik.

Eine besonders wichtige Folgerung hieraus ist die Feststellung, daß die Entropie als Basisgröße ungeeignet ist. Genau wie die Energie wird sie im Karlsruher Physikkurs wie gesehen als nicht weiter zurückführbarer Grundbegriff behandelt, und genau wie das dort schiefeht, funktioniert es auch hier nicht<sup>42</sup>. In beiden Fällen handelt man sich nicht nur das Auftreten der beschriebenen physikalischen Widersprüche ein, sondern auch den Verzicht auf die Möglichkeit, diese Größen zu definieren, so daß als Resultat bei den Schülerinnen und Schülern ein diffuses Bild merkwürdiger Substanzen herauskommen muß, die irgendwo zwischen wahrnehmbarer Realität und gedachter Hilfovstellungen umherschweben. Wir kommen darauf anhand weiterer Beispiele wie auch ganz allgemein wieder zurück.

Wir plädieren aus all diesen Gründen für eine Einführung der Entropie im gymnasialen Physikunterricht, die sich an herkömmlichen Begriffsbildungen orientiert und insbesondere mit der statistischen Interpretation der Entropie beginnt, natürlich nur auf qualitativer Ebene. Sie allein ermöglicht eine einigermaßen anschauliche Vorstellung dieses Begriffes. Dadurch wird insbesondere und von Beginn an vermieden, daß sich bei den Schülerinnen und Schülern Vorstellungen der Entropie als irgendeiner Substanz oder sonst etwas geheimnisvollem einschleichen oder diese mit dem Begriff der Wärme verwechselt wird. Jeder anfänglich rein thermodynamische Zugang zur Entropie provoziert unweigerlich genau solche Fehlinterpretationen. Erst anschließend und auf die qualitative statistische Begriffserklärung aufbauend ist eine elementare thermodynamische Definition der Entropie sinnvoll, wobei weiterhin stets ihr eigentlich statistischer Charakter betont werden muß.

---

<sup>42</sup>Ein weiterer, ganz anderer Vorschlag zur Einführung der Entropie als Basisgröße stammt von Backhaus und Schlichting [32], [33], [34], [35], [446], [447], [448]. Deren Konzept steht nicht nur auf einem ganz anderen fachlichen Fundament, es bewegt sich auch völlig innerhalb der üblichen Terminologie und Begriffsbildung und führt damit gerade nicht zu den beschriebenen Fehlvorstellungen. Das liegt jedoch nicht zuletzt daran, daß die Entropie bei genauerer Betrachtung auch hier keine echte Basisgröße ist, sondern aus dem Begriff der Irreversibilität physikalischer Prozesse abgeleitet wird.

#### 1.3.1.4 Quantenmechanik

Mit der Fixierung auf klassische thermodynamische Konzepte versieht sich der Karlsruher Physikkurs selbst mit einer entscheidenden Einschränkung.

***Argument 4: Der Karlsruher Physikkurs ist nicht geeignet, den grundsätzlichen Unterschied zwischen klassischer Physik und Quantenmechanik zu verdeutlichen.***

Der Umgang der gymnasialen Oberstufenphysik mit der Quantenmechanik stellt leider derzeit ganz allgemein ein Problem dar, unter anderem weil hier mit dem sogenannten Zeigerformalismus ein weiteres höchst problematisches Konzept vorliegt. Da dieses im nächsten Kapitel im Zentrum des Interesses steht, beschäftigen wir uns hier nicht weiter damit. Dennoch muß die Quantenmechanik schon in diesem Kapitel zur Sprache kommen, da sie wie oben bereits angedeutet vom Karlsruher Physikkurs nicht unangetastet bleibt.

Die Behandlung der Quantenmechanik erfolgt beim Karlsruher Physikkurs auf sehr viel konventionellere Art und Weise als diejenige der klassischen physikalischen Teilbereiche – mit einer entscheidenden Ausnahme. Bei dieser handelt es sich wieder um die Erfindung einer neuen Substanz, die den Namen *Elektronium* erhält [229]. Das soll derjenige Stoff sein, dessen Quanten die Elektronen sind. Folglich, so wird erklärt, bestehen die Hüllen der Atome aus Elektronium. Dieses Bild wird nun höchst anschaulich ausgebaut; Kreisströmungen des Elektroniums stehen für Zustände mit Drehimpuls, Übergänge werden als Schwingungen im Elektronium gedeutet, und dergleichen mehr<sup>43</sup>.

Das muß man erst einmal verkraften. Man ist zunächst geneigt, Modelle, die das Bohrsche Atommodell mit seinen schwer zu reparierenden Folgeschäden überflüssig machen, generell zu begrüßen, aber die Beschreibung von Atomhüllenelektronen als einer Substanz ist so beliebig falsch, daß man sich fragen muß, ob man nicht besser bei ersterem geblieben wäre. Was immer Elektronen in der Atomhülle machen, kein Modell wird dem weniger gerecht als die Vorstellung eines Stoffes, der die Atomhülle ausfüllt. Eine der grundlegendsten Eigenschaften der Quantenmechanik ist die, den Rahmen für eine notwendigerweise unanschauliche mikroskopische Beschreibung der Natur zu liefern. Diese bedeutende, auch am Gymnasium unbedingt zu vermittelnde naturphilosophische Erkenntnis wird mit dem Elektroniummodell vollständig aufgegeben. Stattdessen entwickeln die Schülerinnen und Schüler eine trügerische, anschauliche, scheinbar klassische Vorstellung. Sie erhalten keine Chance, auch nur die Spur einer Ahnung des tiefgreifenden Unterschieds zwischen klassischer Physik und Quantenmechanik zu entwickeln. Eine Darstellung der Atomphysik muß klar deutlich machen, daß man das Verhalten der Elektronen in der Atomhülle gar nicht anschaulich beschreiben *kann*. Man kann lediglich räumliche Bereiche angeben, in denen die Wahrscheinlichkeit, das Elektron anzutreffen, beispielsweise größer als 99 % ist; diese räumlichen Bereiche nennt man dann Aufenthaltsräume oder auch *Orbitale*, obwohl das strengenommen nicht ganz korrekt ist<sup>44</sup>. Hat man irgendwo innerhalb eines solchen Orbitals durch eine geeignete experimentelle Anordnung ein Elektron erwischt, so heißt das erstens keineswegs, daß das

---

<sup>43</sup>Hier ist, wie J. Strnad vor einigen Jahren sehr treffend bemerkte [498], nicht nur das Konzept, sondern auch der Name irreführend, da dieser an Positronium erinnert, wobei es sich um ein gebundenes System aus einem Elektron und einem Positron handelt.

<sup>44</sup>Ein Orbital ist definitionsgemäß eine quadratisch Lebesgue-integrierbare Wellenfunktion.



Elektron unmittelbar davor schon dort war, denn bis unmittelbar vor der Messung hatte es die Eigenschaft, einen Ort zu besitzen, gar nicht, genausowenig wie die, einen Impuls oder potentielle oder kinetische Energie zu besitzen. Es hatte nur die Eigenschaft, über definierte Werte der Gesamtenergie sowie von Drehimpuls und Spin zu verfügen. Zweitens ist der Zustand des Elektrons nach der Messung verändert und nicht mehr durch ein Orbital beschreibbar. Dieser zweite Sachverhalt läßt sich zwar gerade noch ad hoc mit dem Elektroniumbild modellieren, aber dadurch verliert es auf einen Schlag seine gesamte vermeintliche Überzeugungskraft. Wer glaubt einem (künstlich möglichst klassisch konstruierten) elastischen Medium, sich bei Bedarf (das heißt bei einer Ortsmessung) spontan praktisch auf einen Punkt zusammenziehen zu können? Der einzig vernünftige Weg ist hier, von vorneherein dem Elektron eine *vollkommen nichtklassische Natur* zuzugestehen; nur so läßt sich der erste Sachverhalt vermitteln, für den es im Elektroniummodell keine Chance gibt. Jeder Versuch, Quantenmechanik mit Hilfe von klassischen Bildern zu vermitteln, ist von vorneherein zum Scheitern verurteilt. Das sollte nicht zuletzt auch im Physikunterricht vermittelt werden.

Wir haben hier wieder ein typisches Beispiel eines von der Schulphysik eigenmächtig eingeführten, in der etablierten Physik unbekanntes Begriffes. Dadurch kommt zu den beschriebenen Schwierigkeiten eine weitere erhebliche dazu: Man entläßt die Schülerinnen und Schüler mit einer Terminologie von der Schule, mit der sie nicht nur gravierende Fehlvorstellungen mitnehmen, sondern in der Fachwelt, sofern sie mit dieser in Kontakt kommen, im günstigsten Fall Verwunderung auslösen. Die Oberstufenvariante des Karlsruher Physikkurses dürfte allerdings bislang noch deutlich weniger verbreitet sein als diejenige für die Mittelstufe. Daher hat es nichts mit den oben diskutierten Problemen zu tun, daß sich ganz generell bei der Vermittlung der Quantenmechanik im Gymnasium einiges ändern sollte. Nur soviel sei an dieser Stelle dazu gesagt: Dem tiefen Bruch, den die Quantenmechanik für das physikalische Weltbild bringt, kann nur durch einen axiomatischen Zugang Rechnung getragen werden, der jegliche Anleihen an klassische Begriffe, die *immer* zu Fehlvorstellungen führen müssen, konsequent vermeidet. Damit wird auch der axiomatischen Vorgehensweise, mit der die Quantenmechanik in mathematischer Strenge aufgezo-gen werden kann, Rechnung getragen. Ein solches axiomatisches Konzept lag in für die gymnasiale Oberstufe genügend elementarisierten Form bis jetzt noch nicht vor; ein entsprechender Unterrichtsvorschlag findet sich im vierten Kapitel.

## 1.3.2 Wissenschaftstheoretische Probleme

### 1.3.2.1 Was sind die physikalischen Grundgrößen?

Wir beginnen diesen Abschnitt mit einem Aspekt, der eng mit den Abschnitten 1.3.1.2 und 1.3.1.3 zusammenhängt und auf die dort auf der Strecke gebliebenen Begriffe zurückkommt.

***Argument 5: Der Karlsruher Physikkurs erklärt gewisse physikalische Grundgrößen zu Basisgrößen, die diesen Anspruch nicht erfüllen können.***

Energie und Entropie sind nicht die einzigen Begriffe, die im Rahmen des Karlsruher Physikkurses gewissermaßen von der zweiten in die erste Reihe geholt werden. Eine ganze

Reihe von Größen, die normalerweise mit guten Gründen als abgeleitete Größen betrachtet werden, finden sich plötzlich als Basisgrößen wieder. Dazu gehören neben Energie und Entropie Impuls, Drehimpuls, Druck, elektrisches Potential und einige weitere. Konsequenterweise müssen dafür andere Größen zurückstecken; insbesondere die Kraft büßt ihre grundlegende Bedeutung ein und taucht, zumindest unter diesem Namen, praktisch gar nicht mehr auf. In all diesen Fällen ist nach unserer Ansicht eine solche Vorgehensweise nicht gerechtfertigt.

Exemplarisch deutlich wird das bei der Behandlung der elementaren Mechanik nach dem Karlsruher Physikkurs [159], [416]. Unter Berufung auf die Mengenartigkeit des Impulses wird dieser schon ganz zu Beginn als Basisgröße eingeführt, umgangssprachlich als „Schwung“ oder „Wucht“ und später qualitativ durch die Aussage, daß ein Körper um so mehr Impuls hat, je höher seine Geschwindigkeit ist und je mehr Masse er hat. Von da an wird alles über Impulsströme beschrieben. Fließt in einen Körper Impuls, wird er schneller, wenn er langsamer wird, fließt Impuls aus ihm ab. Beispielsweise wird ein beschleunigendes Auto beschrieben, indem man sagt, der Motor pumpt Impuls aus der Erde in das Auto. Umgekehrt fließt beim Abbremsen Impuls aus dem Auto heraus und in die Erde hinein. Ob solche Vorgänge schneller oder langsamer ablaufen, wird durch die *Impulsstromstärke* bestimmt, eine Größe, hinter dem sich natürlich nichts anderes als der konventionelle Begriff der *Kraft* verbirgt. Diese übliche Bezeichnung wird im entsprechenden Lehrbuch [225] des Karlsruher Physikkurses nur ganz kurz beiläufig erwähnt und kommt ansonsten nicht vor. Gleichzeitig wird der Impuls zum Energieträger, denn zusammen mit ihm fließt auch Energie in ein System hinein oder aus einem System hinaus.

Wieder haben wir ein anscheinend schönes, zu anderen Strömungsvorgängen analoges Modell, das sich auf den zweiten Blick schnell als nicht nur äußerst problematisch, sondern schlicht als falsch erweist. Die Schwierigkeiten, die man sich mit dem Impuls als Energieträger einhandelt, kamen in Abschnitt 1.3.1.2 bereits zur Sprache, das Modell selbst hat jedoch noch schwerwiegendere Mängel. Zunächst wird hier genau die selbe Fehlinterpretation wie beim Energie- und beim Entropiebegriff provoziert, da wieder die Vorstellung des Strömens einer abstrakten Größe wie des Impulses diesen als eine Substanz erscheinen läßt. Impuls ist aber kein Stoff, sondern ähnlich wie die Energie eine Erhaltungsgröße, die aus einer Symmetrie, in diesem Fall der Translationssymmetrie des betrachteten Problems folgt. Stellt man sich ihn als strömendes „Etwas“ vor, so führt das überdies zu absurden Konsequenzen, beispielsweise bei der einfachen Situation eines Wagens, den man zieht. Nach dem Impulsstrom-Modell fließt Impuls (und auch Energie) *entgegen der Bewegungsrichtung* über die Deichsel in den Wagen und von dort weiter in die Erde; der Impuls des Systems zeigt aber *in die Bewegungsrichtung*, also entgegengesetzt zum Impulsstrom<sup>45</sup>.

Die didaktische Problematik wird jedoch durch eine physikalische noch weit übertroffen, die sofort zutage tritt, wenn man das Differential

$$dp = v dm + m dv$$

der Größe  $p = mv$  betrachtet<sup>46</sup>. Hier beschreibt der erste Summand die Impulsänderung aufgrund der Änderung der Masse, was im nichtrelativistischen Fall als Massestrom

<sup>45</sup>Dieses anschauliche Beispiel stammt von Franz Bader [40].

<sup>46</sup>Vergleiche auch [235].



interpretiert werden kann. Der zweite Term jedoch kommt von der Änderung der Geschwindigkeit, also von der Beschleunigung. Geschwindigkeit kann nicht strömen, sodaß im Fall konstanter Masse bei der Größe  $F = dp/dt$  von einer Stromstärke keine Rede sein kann. Folglich ist der Begriff der Impulsstromstärke insbesondere beim Auftreten von Beschleunigungen *sinnlos* und eine Interpretation der Kraft in diesem Sinne unzulässig; wir kommen darauf gleich zurück.

Natürlich ist die Definition des Impulses als der aus Translationssymmetrie folgenden Erhaltungsgröße keine angemessene Beschreibung für die Schulphysik. Dort genügt es völlig, den Impuls als Produkt aus Masse und Geschwindigkeit und damit als *abgeleitete Größe* zu definieren<sup>47</sup>. Insbesondere sollte das nicht zu Beginn der Behandlung der Mechanik geschehen. Damit kommt man gar nicht erst in Versuchung, sich etwas materielles darunter vorzustellen.

Mindestens genauso folgenschwer ist der aus der Beförderung des Impulses zur Grundgröße folgende weitgehende Verzicht auf den Kraftbegriff, einerseits, weil damit ein grundlegender Begriff der Physik nicht systematisch und standardmäßig eingeführt, sondern durch einen völlig unüblichen neuen Begriff ersetzt wird, der zudem wie gesehen in der Punktmechanik und der Mechanik des starren Körpers gar nicht funktioniert, und andererseits, weil man selbst dann, wenn man über die Inkosistenz der „Impulsstromstärke“ hinwegsieht und auf deren reinen Analogiecharakter zu materiellen Strömungen verweist, auf einem Problem sitzenbleibt. Denn auf diese Weise wird eine wichtige Eigenschaft der Kraft unhandlich, nämlich ihr Vektorcharakter. Die genauen Wege der „Impulsströme“ durch die Gegenstände sind natürlich unbekannt und müssen willkürlich festgelegt werden. Daher geht das einfache Konzept der Kräftezerlegung verloren. Drittens werden die Newtonschen Axiome durch die alleinige Betrachtung von Impulsströmen zu Tautologien. Sie sind jedoch, anders als von den Autoren des Karlsruher Physikkurses behauptet,

<sup>47</sup>Gelegentlich wird darauf verwiesen, daß der Impuls genau wie die Energie im Energie-Impuls-Tensor auf der rechten Seite der Einsteinschen Feldgleichungen

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} + \Lambda g_{\mu\mu} = -\frac{8\pi G}{c^4} T_{\mu\nu},$$

den Grundgleichungen der Gravitation, vorkommt und ihm deshalb gleichermaßen die Bedeutung einer Basisgröße zugestanden werden muß. Der erste Teil dieser Aussage ist natürlich vollständig korrekt, der zweite jedoch falsch, was sich schon dadurch andeutet, daß dann auch Größen wie Torsion und Scherung als Grundgrößen betrachtet werden müßten, da diese ebenfalls im Energie-Impuls-Tensor vorkommen. Der eigentliche Fehler dieser Argumentation ist jedoch sehr viel schwerwiegender. In der allgemeinen Relativitätstheorie wird jeder Vorgang, der zu einer Erhöhung der Energie eines Systems beiträgt, gleichzeitig gravitativ wirksam und muß folglich im Energie-Impuls-Tensor auftreten. Allerdings taucht dort der Anteil der Gravitation selbst nicht auf; dieser steht komplett auf der linken Seite der Feldgleichungen der Gravitation. Man kann zwar durch eine Umformulierung derselben einen sogenannten Energie-Impuls-Pseudotensor definieren, dieser ist jedoch kein Tensor und damit nicht relativistisch kovariant. Die Folge ist, daß die Begriffe der Energie und des Impulses in der allgemeinen Relativitätstheorie gar nicht global kovariant definierbar sind, weswegen es dort im allgemeinen *weder einen Energie- noch einen Impulserhaltungssatz gibt*. Das ist auch schon daran erkennbar, daß diese Erhaltungssätze aus Symmetrien hervorgehen, deren zugehörige Größen Zeit und Ort in der allgemeinen Relativitätstheorie selbst zu dynamischen Variablen und damit nicht zu absoluten Größen werden. Lediglich wenn die betrachteten Raum-Zeiten selbst spezielle Symmetrien aufweisen, lassen sich globale Energie-Impulserhaltungssätze als Sonderfälle formulieren. Somit demonstriert die allgemeine Relativitätstheorie im Gegenteil gerade besonders drastisch, daß sowohl der Energie als auch dem Impuls keineswegs die Eigenschaft einer Basisgröße zukommt.

*keine* Selbstverständlichkeiten und schon gar nicht intuitiv klar. Das gilt im Besonderen für das zweite Axiom und den daraus folgenden Trägheitssatz, was 2000 Jahre Aristotelische Physik belegen. Der Ersatz, der für den Trägheitssatz angeboten wird, nämlich die Aussage, Masse sei die Fähigkeit eines Körpers, Impuls aufzunehmen, ist unzureichend. Insbesondere kann man damit *Trägheitskräften* und ihrer großen Bedeutung in den unterschiedlichsten Gebieten in keiner Weise gerecht werden. Viertens schließlich provoziert das Verwenden des Begriffs der Impulsstromstärke anstelle von Kraft die Gefahr, Impuls und Impulsstromstärke zu verwechseln. Untersuchungen in Unterrichtsversuchen zeigten, daß diese Befürchtung sehr berechtigt ist [400], [490], [491]. Die Folge ist offensichtlich, daß hierdurch bei den Schülerinnen und Schülern ein kompetenter Umgang mit der physikalischen Größe Kraft, um die man, ganz egal wie man sie nennt, im Unterricht nicht herumkommt, deutlich erschwert wird. Kräfte sind als tatsächliche Grundgrößen sehr abstrakt, und die damit verbundenen Schwierigkeiten werden nur verstärkt, wenn man ihnen eine nicht existierende Anschaulichkeit zu verleihen versucht. Anschaulich wird die Sache jedoch, wenn man sich damit begnügt, daß man Kräfte an ihren Wirkungen erkennt.

In ähnlicher Weise werden an anderen Stellen der Druck („Druck ist was der Druckmesser anzeigt ...“) oder das elektrische Potential („Das elektrische Potential ist das, was die Elektrizität antreibt ...“) als nicht weiter hinterfragbare Basisgrößen eingeführt. Alle diese Begriffe sind jedoch als Grundgrößen unbrauchbar. Sie alle können und müssen unter Verwendung tatsächlicher Basisgrößen wie Zeit, Ort, Kraft oder elektrischer Ladung präzise definiert werden. Insbesondere sind Energie und Impuls sinnlos ohne die vorgeordneten Größen Ort, Geschwindigkeit und Kraft. Das gilt auch für diejenigen Vorstellungen, mit denen die Autoren des Karlsruher Physikkurses die Begriffe der Energie und des Impulses als Grundbegriffe plausibel machen wollen. Es liegt folglich ein logischer Zirkel vor, denn um Energie und Impuls als Basisgrößen zu etablieren, benötigt man gerade diejenigen Größen, die dadurch als Basisgrößen abgeschafft werden sollen. Bleibt man bei der bewährten Reihenfolge und betrachtet in der Mechanik zunächst einfache kinematische und sodann einfache dynamische Sachverhalte, wobei die Kräfte als nicht weiter erklärbare Größen beschrieben werden, die man an ihren Wirkungen erkennt, treten solche Schwierigkeiten nicht auf.

### 1.3.2.2 Woher kommt der universelle Antrieb?

Unser nächster Einwand kritisiert die Postulierung eines universellen Strom-Antriebs-Konzepts, das in allen Bereichen der Physik gültig sein soll.

***Argument 6: Der Karlsruher Physikkurs postuliert ein für alle physikalische Disziplinen gültiges allgemeines Antriebsprinzip, das keine Entsprechung in der Realität besitzt.***

Neben dem Energie-Träger-Konzept ist das Strom-Antrieb-Konzept das zweite Standbein der „fluidalen“ Strömungsinterpretation dynamischer Vorgänge durch den Karlsruher Physikkurs [414], [418]. Unter Berufung auf die in Abschnitt 2.6 beschriebene *formale* Übertragbarkeit hydromechanischer Begriffe auf extensive und intensive Größen versucht man dabei, die „Strömungen“ der extensiven Größen durch Differenzen der entsprechenden intensiven Größen zu erklären, die dazu sämtlich als Potentiale interpretiert werden,

in Anlehnung an das elektrische Potential. Eine solche Interpretation ist jedoch in keiner Weise zwingend. Das Strom-Antrieb-Konzept suggeriert einen *kausalen* Zusammenhang zwischen Potentialdifferenz und „strömender“ Größe, von dem überhaupt keine Rede sein kann. Eine solche Einteilung in Ursache und Wirkung ist völlig willkürlich und physikalisch in keiner Weise begründbar. Es handelt sich hier um nichts anderes als die Erfindung einer universellen Antriebskraft der Natur von rein idealistischem Charakter, die stark an die Aristotelische Physik erinnert mit ihrem universellen Drang der Gegenstände, zur Ruhe zu kommen.

Erschwerend kommt hinzu, daß die Deutung der intensiven Größen als Potentiale nicht allgemein möglich ist. Das einfachste Gegenbeispiel ist der Druck bei Wasserströmungen. Wäre dieser ein Potential, so müßte Wasser *jedem* Druckgefälle folgen und damit beispielsweise im Meer von unten nach oben strömen. Ähnliches gilt für die Entropie, wo eine Temperaturdifferenz keineswegs notwendige Voraussetzung für nichtverschwindende Entropiestromstärken ist.

Schülerinnen und Schüler haben keine Möglichkeit, die über einen reinen Modellcharakter hinausgehende Eigenschaft des Strom-Antrieb-Konzepts richtig einzuschätzen, eine an romantische Naturvorstellungen erinnernde Art neue Naturkraft zu postulieren. Sie müssen auch das für bare Münze nehmen und interpretieren es realistisch, das heißt als wirklich existierendes Phänomen. Daher sollten solche Vorstellungen im Unterricht vermieden werden und eine Beschränkung auf die herkömmlichen dynamischen Prinzipien der Physik geübt werden.

### 1.3.2.3 Rückkehr zur Kontinuumsphysik?

Einer der schwerwiegendsten Kritikpunkte bemängelt eine Rückwärtsbewegung zu physikalischen Theorien, die aus dem vorletzten Jahrhundert und noch früheren Epochen stammen und sich schon lange als unhaltbar erwiesen haben.

***Argument 7: Der Karlsruher Physikkurs beinhaltet eine Rückkehr zu Kontinuums- und Substanzkonzepten der idealistischen Physik der frühen Aufklärung, die längst überholt sind und zu fatalen Fehlvorstellungen führen.***

Diese Entwicklung dürfte Unbeteiligte am meisten verblüffen. Die im Karlsruher Physikkurs schwer vermeidbare Fehlinterpretation, Energie, Entropie, Impuls und so weiter als Substanzen und damit als materielle Stoffe aufzufassen, wurde weiter oben ausführlich diskutiert. Damit ist jedoch nur ein Teil der Problematik erfaßt. Denn die Verwendung von Thermo- und Hydrodynamik als Ausgangspunkte führt automatisch auf eine reine Kontinuumsphysik. Das wäre solange in Ordnung, wie man es als Näherung auffaßt. Der Karlsruher Physikkurs erhebt jedoch wie erwähnt einen Universalitätsanspruch und faßt darüberhinaus die formal auftauchenden Kontinua zwar modellmäßig, aber für Schülerinnen und Schüler natürlich sehr konkret als Fluida, also als Flüssigkeiten auf. Damit ist unter anderem faktisch die Rückkehr zu Caloricum beziehungsweise Phlogiston<sup>48</sup> und der

<sup>48</sup>Caloricum oder Phlogiston ist der Name für den Stoff, den der Ansbacher Naturgelehrte Georg Ernst Stahl nach Vorarbeiten des Alchimisten Johann Joachim Becher zu Beginn des 17. Jahrhunderts zur Erklärung beispielsweise von Verbrennungsprozessen postulierte. Nach seiner Theorie entweicht Phlogiston aus Gegenständen bei Verbrennung und dringt bei Erwärmung in sie ein. Phlogiston ist

elektrischen Flüssigkeit<sup>49</sup> vollzogen. Was das für den Bildungsanspruch des Fachs Physik als einer modernen Naturwissenschaft bedeutet, braucht nicht weiter auseinandergesetzt zu werden. Die Sache hat jedoch noch weitergehende Folgen.

Die Interpretation der Elektrizität<sup>50</sup> beziehungsweise des elektrischen Stroms als strömender Flüssigkeit wird im Karlsruher Physikkurs relativ direkt vertreten [414]. Insbesondere kommen keine Ladungsträger vor, womit die Schülerinnen und Schüler, die verbreitet die Vorstellung des elektrischen Stroms als bewegte Elektronen schon mitbringen, wieder umlernen und den Eindruck bekommen müssen, von einem modern erscheinenden Begriff wie dem der Elektronen auf eine zumindest merkwürdige Vorstellung des elektrischen Stroms zurückgedrängt zu werden.

Deutet man Potentialdifferenzen unhinterfragt als Ursache des elektrischen Stroms, so führt das in Verbindung mit der üblichen, aus der Konvention der Richtung der elektrischen Feldlinien folgenden Vorzeichenkonvention des elektrischen Potentials zur fatalen Konsequenz, daß der elektrische Strom *von + nach -* fließt. Das steht in schärfstem Konflikt zu der für Schülerinnen und Schüler leicht nachvollziehbaren tatsächlichen Bewegungsrichtung der Leitungselektronen *von - nach +*, da diese vom Minuspol abgestossen und vom Pluspol angezogen werden. Ursache hierfür ist eine in wissenschaftlichen Publikationen explizit [222], im Unterricht aber wohl weniger deutlich ausgesprochene Umdefinition des Begriffs der Richtung der elektrischen Stromstärke von der anschaulichen Bedeutung der *Driftrichtung der Ladungsträger* zur unanschaulichen Bedeutung der *Richtung des Vektorfelds der Stromdichte*. Diese hängt wie gesagt von der Konvention der Feldlinienrichtung ab. Konsequenterweise sollte man von solchen Begriffen im Unterricht völlig absehen und es dabei belassen, daß die Richtung des elektrischen Stroms durch die Drift der Ladungsträger festzulegen ist. Das führt bei Elektronen in metallischen Leitern auf die Richtung vom Minuspol zum Pluspol, kann bei anderen Ladungsträgern aber auch entgegengesetzt sein. Vor allen Dingen erspart man sich dabei die Diskussion der Frage, wie es sein kann, daß sich bei einem elektrischen Strom die Ladungsträger entgegen der „Stromrichtung“ bewegen.

Der Karlsruher Physikkurs setzt der ganzen Sache nun noch eins drauf, indem er, wie in Abschnitt 1.3.1.4 schon beschrieben, einen neuen Stoff postuliert, nämlich das *Elektronium*. Das soll der „Stoff der Elektrizität“ sein und ist damit nichts anderes als Franklins elektrisches Fluidum. Wir haben schon gesehen, zu welcher extremen Fehlvorstellung das in der Quantenmechanik führt, aber selbst davon abgesehen bewirkt das Elektronium nichts anderes als die weitere Verfestigung der unglücklichen Vorstellung der Elektrizität als einer Substanz.

Vergleichbares geschieht mit allen anderen extensiven Größen. Referenten von Fortbildungen zum Karlsruher Physikkurs pflegen offen einzugestehen, daß sie sich Wärme als „eine Art Substanz“ vorstellen. Damit ist man wieder beim Caloricum oder Phlogiston angelangt. Die Vorstellung von Wärme als Substanz ist seit 224 Jahren überholt. Sie steht in schärfstem Widerspruch zur Erklärung von thermischer Energie als kinetischer

---

damit der hypothetische „Wärmestoff“. Die Phlogiston-Theorie war spätestens 1785 mit Lavoisiers Theorie der Oxidation erledigt.

<sup>49</sup>Diese wurde ab etwa 1767 von Charles Cavendish und Benjamin Franklin zur Erklärung elektrostatischer Phänomene und des elektrischen Stroms verwendet.

<sup>50</sup>Dieser Ausdruck wird in diffuser Weise zur Vermeidung des Begriffs der elektrischen Ladung verwendet.

Energie der ungeordneten Teilchenbewegung. Bei Energie und Entropie wurde alles Nötige dazu oben bereits gesagt, genauso inakzeptabel sind die Interpretationen von Impuls als Substanz, Drehimpuls als Substanz, und so weiter. Der Karlsruher Physikkurs führt dazu, daß sich Schülerinnen und Schüler alle diese abstrakten Größen als reale Stoffe vorstellen.

Eine der ambitioniertesten Vorgaben des Karlsruher Physikkurses ist nach Aussage seiner Autoren, physikalische Altlasten zu entsorgen<sup>51</sup>. Angesichts der geschilderten Situation ist festzustellen, daß dieser Anspruch nicht nur nicht erreicht wird, es werden im Gegenteil längst entsorgte Altlasten reaktiviert.

### 1.3.2.4 Fehlende Begriffsdefinitionen

Ein Kritikpunkt, der weniger den Karlsruher Physikkurs selbst als vielmehr dessen propagierte Umsetzung im Unterricht betrifft, ist der folgende.

***Argument 8: Der Karlsruher Physikkurs führt im Unterricht zu einem Verzicht auf exakte Definitionen der betrachteten physikalischen Begriffe.***

Weiter oben wurde die Strategie des Karlsruher Physikkurses beschrieben, Begriffe wie Impuls, Druck, Energie oder Entropie zu Basisgrößen zu machen. Überaus auffällig ist dabei, daß im Rahmen konkreter Unterrichtsvorschläge<sup>52</sup> solche Größen *nicht definiert werden*. Sie fallen vom Himmel und sollen nach und nach durch ein Ansammeln ihrer Eigenschaften verstanden werden<sup>53</sup>. Das gilt in besonders auffälliger Weise für Impuls, Energie und Entropie, betrifft aber weitere extensive Größen gleichermaßen. Begründet wird das zum einen mit der Aussage, Schülerinnen und Schüler brächten entsprechende intuitive Vorstellungen bereits in den Unterricht mit, und andererseits damit, daß man auch im Alltag Objekte an ihren Eigenschaften und nicht an irgendwelchen Definitionen erkenne. Diese Vorgehensweise ist teilweise durch Prinzipien des Karlsruher Physikkurses selbst angelegt, da dieser einige eigentlich abgeleitete Größen zu undefinierbaren Basisgrößen erklärt, teilweise ist sie aber schwer mit diesen in Einklang zu bringen. Denn die darin auftretende starke Betonung von Analogien in unterschiedlichen physikalischen Bereichen ist nur etablierbar, wenn zuvor sehr genau definiert wird, worum es geht.

Ein solches Vorgehen fördert die beschriebenen Fehlvorstellungen natürlich erst recht. Schwerer wiegt jedoch, daß es in stärkstem Gegensatz zu den erkenntnistheoretischen und naturphilosophischen Eigenschaften der physikalischen Beschreibung unserer Welt steht. Die Physik zeichnet sich als mathematische Naturwissenschaft in besonderem Maße dadurch aus, daß Begriffe exakt definiert und deren formale Strukturen erforscht werden.

<sup>51</sup>„Altlasten der Physik“ ist der Name einer Artikelserie von F. Herrmann und anderen, die seit 1994 zunächst in der Zeitschrift *Physik in der Schule* und später in der Nachfolgezeitschrift *Praxis der Naturwissenschaften – Physik in der Schule* veröffentlicht wird. Inzwischen sind die Aufsätze teilweise auch in Buchform erschienen [233]. Die Anzahl etablierter und vor allen Dingen unentbehrlicher physikalischer Fachbegriffe und Konzepte, die hiernach „entsorgt“ werden sollten, ist beeindruckend; Beschleunigung, Arbeit und Wärme sind nur die prominentesten darunter. das Ganze paßt nach Inhalt und Konzept genau zum hier beschriebenen Karlsruher Physikkurs, soweit es nicht ohnehin schon Bestandteil desselben ist.

<sup>52</sup>Siehe beispielsweise [413] – [418].

<sup>53</sup>Dafür wurde der Begriff „kumulatives Lernen“ geprägt.



Man begnügt sich dabei im allgemeinen gerade nicht mit dem einfachen Zusammentragen charakteristischer Eigenschaften – es sei denn, es handelt sich tatsächlich um Basisgrößen, die nicht auf noch fundamentalere Begriffe zurückführbar sind. Von diesen kommt in der Schulphysik im wesentlichen jedoch nur vier Stück eine wirkliche Bedeutung zu, nämlich *Ort*, *Masse*, *elektrische Ladung* und *Kraft*. Alle anderen Größen lassen sich aus diesen formal präzise ableiten. Aufgabe der Physik ist es insbesondere, quantitative Beziehungen zwischen solchen fundamentalen wie nicht fundamentalen Größen zu finden. Es muß ein wesentliches Bildungsziel jeglichen Physikunterrichts sein, auch in der Schule, diese Aspekte besonders deutlich zu machen. Das bedeutet insbesondere: An irgendeiner Stelle (möglichst früh) *müssen* altersgemäße, aber präzise Definitionen der eingeführten physikalischen Begriffe erfolgen. Das Ansammeln von Eigenschaften derselben, um deren Bedeutung „nach und nach zu verstehen“, *genügt nicht*.

Eine beabsichtigte Begleiterscheinung der Abkehr von Begriffsdefinitionen der propagierten Unterrichtsgänge ist ein starkes Zurückdrängen quantitativer und damit mathematischer Betrachtungen im Physikunterricht. Kurz gesagt, es soll viel weniger und viel später gerechnet werden. Gleichzeitig sollen, ganz in der Tradition Martin Wagenscheins, Phänomene stärker in den Vordergrund gestellt werden. Abgesehen davon, daß man durch ausschließliches Betrachten der Phänomene häufig *gerade nicht* zu den physikalischen Gesetzmäßigkeiten gelangt, begibt man sich damit gefährlich nahe an eine Transformation des Fachs Physik in die traditionelle Naturkunde der früheren Volksschulen. Dabei hilft auch der Verweis auf den angeblich unbedingt notwendigen Alltagsbezug des Physikunterrichts nichts. Dieser besagt genauesehen nicht weniger, als daß man die Schülerinnen und Schüler mit technischer Problemlösekompetenz für den Alltag aussatten müsse<sup>54</sup> und wird teilweise bis zur Forderung gesteigert, Physikunterricht müsse sinnstiftende Wirkung entfalten [376]. Wenn sich das Fach Physik selbst derart unerfüllbaren Ansprüchen aussetzt, kann es nur verlieren. Die Physik muß sich auf ihre Eigenschaft, zu den *Naturwissenschaften* zu gehören, und ihre besondere Stärke besinnen, und diese liegt insbesondere in der erstaunlichen Tatsache der zumindest prinzipiell *mathematischen Beschreibbarkeit der Natur*. Das wenigstens ansatzweise zu vermitteln, sollte eines der wesentlichen Bildungsziele eines jeden Physikunterrichts sein.

### 1.3.3 Didaktische Überlegungen

#### 1.3.3.1 Präkonzepte

Das nächste Argument betrifft unmittelbar die Strategie, mit welcher der Karlsruher Physikkurs Begriffsbildungen im Physikunterricht anstrebt.

---

<sup>54</sup>Im Rahmen des sogenannten *Erlanger Programms* versuchten unter anderem R. Inhetveen [262], P. Janich [271], [272], [273] und P. Lorenzen [349], [350], die Physik generell als reine Technikwissenschaft, genauer gesagt als Disziplin zur Bereitstellung von technischem Verfügungswissen zu beschreiben. Sie bewegten sich dabei auf der Basis der sogenannten *Protophysik*, einer Variante der konstruktivistischen Wissenschaftstheorie, welche die Physik allein aus den an grundlegende Meßverfahren zu stellenden Vorschriften zu rekonstruieren versucht. Ein allerdings etwas unkritischer Überblick dazu ist [270]. Beides, das Erlanger Programm wie auch die Protophysik, dürfen als gescheitert betrachtet werden, was sowohl an physikalischen Unzulänglichkeiten im speziellen als auch an der grundsätzlichen Selbstwidersprüchlichkeit jeder konstruktivistischen Erkenntnistheorie im allgemeinen liegt; siehe dazu exemplarisch einerseits [85] und [130] sowie andererseits [515] und [516].

***Argument 9: Der Karlsruher Physikkurs versucht, Präkonzepte der Schülerinnen und Schüler aufzugreifen, ersetzt diese aber nicht durch etablierte physikalische Konzepte, sondern behält sie weitgehend bei.***

Nicht nur der Karlsruher Physikkurs, sondern inzwischen auch der Bildungsplan für Gymnasien des Landes Baden-Württemberg fordert, bei der Bildung physikalischer Begriffe von den Vorerfahrungen oder Präkonzepten der Schülerinnen und Schüler auszugehen und daraus *schrittweise* durch Zusammentragen der Eigenschaften physikalischer Größen zu deren korrekter Bedeutung und entsprechender wissenschaftlicher Beschreibung zu gelangen. Dazu sind drei Aspekte kritisch anzumerken.

Erstens ist diese Strategie ganz unabhängig vom gewählten Unterrichtsgang nicht generell durchzuhalten. Die Natur tut uns nicht den Gefallen, sich an solche Präkonzepte zu halten, und folglich führen diese häufig selbst bei noch so kleinschrittigem Vorgehen beim besten Willen der Lehrerin oder des Lehrers nicht zum gewünschten Erfolg; die Präkonzepte bleiben bestehen und verwandeln sich so in *Fehlvorstellungen*. Dagegen hilft nur ein planmäßig herbeigeführter Bruch mit den Präkonzepten, etwa in Form provozierter Konflikte, die sich ausschließlich durch einen Wechsel auf tatsächliche physikalische Konzepte lösen lassen. Das muß frühzeitig geschehen, da sich sonst die mitgebrachten falschen Vorstellungen nur noch schwer oder gar nicht mehr ändern lassen. Ein berüchtigtes Beispiel hierfür ist das Bohrsche Atommodell, das, einmal in der Schule unterrichtet, den Schülerinnen und Schülern später meist nicht mehr auszutreiben ist. Aus dieser Perspektive scheint die Vorgehensweise, Präkonzepte stets aufzugreifen und sie immer nur schrittweise zu verbessern, zumindest schon fragwürdig.

Zweitens sollte man zwischen *Präkonzepten* und *Vorwissen* unterscheiden. Das wird exemplarisch am Teilchenmodell der Materie deutlich, das im Rahmen des Karlsruher Physikkurses sehr lange ausgespart wird und erst in der Oberstufe in Erscheinung tritt, weil, so führt man aus, eine diskrete Natur der Materie nicht den Präkonzepten der Schülerinnen und Schüler entspricht. Nun mag es stimmen, daß diese sich Materie unbewußt kontinuierlich vorstellen, aber das tut genaugenommen *jeder*, und außerdem bringen heutzutage schon Kinder in den fünften Klassen Kenntnisse über den Aufbau der Materie aus Atomen und Molekülen mit, oft einschließlich dieser Begriffe selbst. Natürlich können sie sich nicht wirklich etwas darunter vorstellen, aber auch hier gilt wieder, daß das auf *jeden* zutrifft. Daher wäre es höchst albern, die Schülerinnen und Schüler zu zwingen, von diesem Vorwissen wieder abzurücken und zu einem nicht nur aus deren Sicht falschen und vor allem rückständigen Modell zu wechseln. Die Konsequenz daraus muß nicht nur sein, das Teilchenmodell im Physikunterricht so früh wie möglich einzuführen, es sollte sogar schon im Naturphänomene-Unterricht der Unterstufe behandelt werden. Genau das selbe gilt für die Elektrizitätslehre; auch hier muß die Rückkehr von Modellen mit bewegten Elektronen zu einem mysteriösen elektrischen Fluidum, auch wenn man es wohl nicht explizit so nennt, den Schülerinnen und Schülern völlig zu recht als intellektueller Rückschritt erscheinen.

Drittens drängt sich selbst dann, wenn man die Strategie des Aufgreifens und schrittweise Weiterentwickelns von Präkonzepten für einmal hinnimmt, die Frage auf, ob diese Strategie vom Karlsruher Physikkurs eigentlich tatsächlich befolgt wird. Das ist in der Tat zu einem nicht unerheblichen Teil nicht der Fall, wie die Abschnitte 1.3.1.2, 1.3.1.3 und 1.3.2.3 deutlich gezeigt haben dürften. Es werden im Gegenteil verbreitet nicht nur

mitgebrachte Präkonzepte beibehalten und verstärkt, sondern auch neue Fehlvorstellungen provoziert. Darüberhinaus darf durchaus angezweifelt werden, ob alle Vorstellungen, die zu Präkonzepten erklärt werden, auch wirklich solche sind. Auch hierzu wurden Beispiele diskutiert, man denke an den zur vermeintlichen Selbstverständlichkeit werdenden Trägheitssatz, der sicherlich den Vorerfahrungen von fast jedem deutlich widerspricht. Auch Energie ist kein Präkonzept, keine Größe, für die man ein Gefühl mitbringt, sondern ein abstraktes Konzept, daß man durch Berechnungen handhaben muß.

### 1.3.3.2 Analogien

Größten Wert legt der Karlsruher Physikkurs auf die Verwendung von Analogien, was Anlaß zu einem weiteren Gegenargument gibt.

***Argument 10: Der Karlsruher Physikkurs verwendet Analogien in einem Maß, das weit über den für Schülerinnen und Schüler zu bewältigenden Umfang hinausgeht.***

Nahezu jede Darstellung des Karlsruher Physikkurses beginnt mit der oder erwähnt zumindest die Bedeutung der Betrachtung mengenartiger Größen im Unterricht. Das sind Größen, die bilanzierbar sind in dem Sinn, daß für sie eine Kontinuitätsgleichung aufgestellt werden kann; mit anderen Worten, es handelt sich dabei genau um die extensiven Größen in der verallgemeinerten Gibbsschen Fundamentalbeziehung. Diese sollen sämtlich einheitlich behandelt werden und so Anlaß zu flächendeckenden Analogiebetrachtungen geben. Bevor wir auf die Problematik dieses Analogie-Konzepts zu sprechen kommen, sei vorab noch folgende Bemerkung gemacht: Im Zusammenhang mit den mengenartigen Größen wird einerseits zu Recht betont, daß nicht alle davon Erhaltungsgrößen sind, es wird andererseits aber auch behauptet, daß Erhaltungssätze nur für mengenartige Größen sinnvoll sind<sup>55</sup>. Das ist nicht richtig, wie beispielsweise die aus der Translationsinvarianz eines Systems folgende Schwerpunktserhaltung belegt. Der Schwerpunkt eines Systems ist natürlich keine mengenartige Größe.

Die möglichst weitgehende Verwendung von Analogien kann vielleicht als didaktisches Hauptprinzip des Karlsruher Physikkurses angesehen werden. Aus rein physikalischer Sicht ist dagegen auch nichts einzuwenden. Allerdings sollte man diesen Sachverhalt nicht auf tatsächliches Strömungsverhalten der beschriebenen Größen, sondern vielmehr auf die Leistungsfähigkeit des verwendeten mathematischen Formalismus schieben, und genau darin liegt das Problem. Die Verwendung von Analogien auf dieser Ebene ist eine beträchtliche Abstraktionsleistung, die weit über das Abstraktionsvermögen von Schülerinnen und Schülern in der gymnasialen Mittelstufe hinausgeht<sup>56</sup>. Die Folge ist, daß die verwendeten Modelle für bare Münze genommen werden, und das um so mehr, als üblicherweise in der 7. Klasse mit besonders anschaulichen und damit einprägsamen Beispielen begonnen wird. Der Einsatz von Wasserstromkreisen [414] und ähnlichen Systemen mit materiellen Substanzen als Einführung und zum Kennenlernen zentraler Begriffe wie Stromstärke, Antrieb, Potential und so weiter führt zu einer dauerhaften Verfestigung der Vorstellung, daß bei physikalischen Prozessen immer konkrete materielle Substanzen

---

<sup>55</sup>Siehe zum Beispiel [228].

<sup>56</sup>Das gilt mit nur geringen Einschränkungen auch noch für die Oberstufe. Man sollte auch dort im Unterricht immer wieder darauf hinweisen, allfällig auftretende Analogien nicht wörtlich zu nehmen.



strömen. Diese Vorstellung ist in den meisten anderen Fällen völlig falsch und führt zu weiteren, teilweise völlig absurden Fehlvorstellungen. Zum Beispiel ist den Schülerinnen und Schülern als Folge der Analogie Wasserhahn – elektrischer Schalter nicht mehr immer klar, daß in einem elektrischen Stromkreis der elektrische Strom bei geschlossenem und nicht bei offenem Schalter fließt. Dieses haarsträubende Problem kann bei herkömmlichen Unterrichtsformen selbstverständlich gar nicht erst auftreten. Erschwert wird die Sache fast zwangsläufig auch noch durch die schwer zu vermeidende Verwechslung von Strömungsgeschwindigkeit einer strömenden Substanz aus Teilchen und Stromstärke einer physikalischen Größe.

Das Arbeiten mit Analogien ist an und für sich zweifellos eine wertvolle didaktische Vorgehensweise. Sie sollte jedoch sehr dosiert eingesetzt werden, je jünger die unterrichteten Schülerinnen und Schüler sind, desto vorsichtiger. Und sie sollte ein Hilfsinstrument bleiben, die andere Methoden unterstützen kann, anstatt zum Kernstück der Strategie gemacht zu werden.

### 1.3.3.3 Verzicht auf wichtige Begriffe

Besonders ärgerlich ist, daß das Unterrichten nach dem Karlsruher Physikkurs zum Wegfallen einiger eigentlich unverzichtbarer Begriffe führen soll.

***Argument 11: Der Karlsruher Physikkurs verzichtet auf einige nicht nur in der Schulphysik bewährte, sondern aus wissenschaftlicher Sicht unerläßliche Begriffe.***

Ganz im Sinne der selbstgestellten Aufgabe, physikalische Altlasten zu entsorgen, schafft der Karlsruher Physikkurs gleich eine ganze Reihe grundlegender Begriffe ganz oder teilweise ab. Wärme, Arbeit und die unterschiedlichen Energieformen sollen ganz wegfallen, Kraft und Leistung werden umgetauft und tauchen unter ihrem eigenen Namen so gut wie nicht auf, und Teilchenmodelle der Materie werden zwar auch nicht generell abgeschafft, ihre Einführung aber im günstigsten Fall auf einen sehr viel späteren Zeitpunkt verschoben. Der kritische Beobachter sollte daher nicht zögern, sich davon zu überzeugen, ob diese Begriffe tatsächlich völlig entbehrlich sind. Das sind sie in der Hochschulphysik sicher nicht, und wie sich zeigt, ist es in der Schulphysik nicht anders.

Beginnen wir mit der Arbeit. Hierbei handelt es sich um eine Prozeßgröße, und wenn man auf die Betrachtung der Prozesse verzichtet und nur den Zustand des betrachteten Systems davor und danach berücksichtigt, kommt man natürlich auch nur mit dem Begriff der Energie aus; verzichtet man nicht darauf, ist man schon auf Umschreibungen wie Energietransport, Energieumwandlung und dergleichen angewiesen. Das wäre für sich allein kein Problem. Entscheidend ist jedoch, daß die Arbeit operational definiert werden kann und damit erst eine saubere Definition der Energie erlaubt. Ohne eine solche Definition bleibt der Begriff der Energie unklar, wie weiter oben ausführlich beschrieben wurde. Darüberhinaus entspricht der Arbeitsbegriff bei sorgfältiger Behandlung im Unterricht sehr genau den Alltagsvorstellungen, beispielsweise im Sinne der Erfahrung, daß eine Kraft Bewegungen erschwert. Es wäre auch aus dieser Sicht äußerst ungeschickt, auf diese Größe zu verzichten.

Ebenfalls eine Prozeßgröße ist die Wärme. Die Schwierigkeiten, die ihre Verwendung im Unterricht angeblich verursacht, treten höchstens dann auf, wenn sie falsch oder un-

präzise verwendet wird. Ihre genaue Definition ist die folgende: Wärme ist thermische Energie, die durch eine Systemgrenze hindurchtritt. In dieser Form ist sie nicht nur ein präziser Begriff, der für einen sauberen Aufbau der Thermodynamik unverzichtbar ist, sie entspricht auch exakt den Alltagsvorstellungen, wenn man diese nur genau genug betrachtet. Beispielsweise wäre es nicht nur physikalisch falsch, zu behaupten, in einer mit heißem Tee gefüllten Thermoskanne befinde sich Wärme, es würde auch keinerlei Erfahrung entsprechen, denn die Thermoskanne fühlt sich so kühl an wie die Umgebung. Erst wenn die darin enthaltene thermische Energie herauskann und zumindest teilweise in ein Sinnesorgan oder ein Meßgerät eintritt, ist Wärme wahrnehmbar, aber eben genau deshalb, weil diese jetzt durch eine Systemgrenze hindurchtritt. Das läßt sich auf jede Situation übertragen, die mit der Wahrnehmung von Wärme einhergeht. Darüberhinaus ist der Begriff der Wärme unentbehrlich, wenn irgendwelche Sachverhalte, die mit Dissipation oder dem zweiten Hauptsatz der Wärmelehre zusammenhängen, zur Sprache kommen sollen. Wieder haben wir einen Begriff, auf den man nicht verzichten kann.

Über Energieformen, Leistung und Kraft sowie Teilchenmodelle wurde weiter oben bereits ausführliches gesagt, jeweils mit dem Ergebnis, daß ihre fehlende beziehungsweise stiefmütterliche Behandlung im Karlsruher Physikkurs nicht gerechtfertigt ist. Generell darf festgestellt werden, daß die Begriffsauswahl in den herkömmlichen Unterrichtsgängen im Fach Physik vielleicht doch nicht so verkehrt ist, wie bisweilen behauptet wird.

#### 1.3.3.4 Was ist mit der etablierten Fachsprache?

Der Karlsruher Physikkurs ist mehr als nur eine gegenüber bisher üblichen Methoden veränderte Gewichtung und Anordnung physikalischer Begriffe, was uns auf unser abschließendes Gegenargument führt.

***Argument 12: Der Karlsruher Physikkurs berücksichtigt nicht, daß der Maßstab für Begriffsbildungen und -verwendungen im Physikunterricht stets nur die etablierte Hochschulphysik sein kann.***

In diesem Abschnitt verallgemeinern wir die Problematik des vorhergehenden. Dabei steht die Frage im Mittelpunkt, wonach sich die Auswahl und Verwendung von Fachbegriffen im schulischen Physikunterricht zu richten hat, denn der Karlsruher Physikkurs weicht in dieser Hinsicht beträchtlich von der üblichen Linie ab; diese Eigenschaft hat er mit dem Zeigerformalismus-Konzept für die Quantenmechanik gemein, das in Abschnitt 3.1.4 bereits kurz erwähnt wurde und gleich anschließend in Kapitel 2 die Hauptrolle spielt. Tritt man gedanklich einen Schritt zurück und betrachtet die Sache aus etwas mehr Distanz, wundert man sich jedoch sofort, warum sich so eine Frage überhaupt stellen sollte, insbesondere, wenn man sich die selbe Situation auf einen beliebigen anderen Fachbereich übertragen vorstellt, etwa irgendeine Fremdsprache. Es wurde bereits erwähnt, daß sich dort die Frage stellen würde, ob es angesichts der Lernschwierigkeiten, welche die Schülerinnen und Schüler verbreitet mit den offensichtlich schwierigen Vokabeln haben, nicht besser wäre, wenn man stattdessen neue, einfachere Vokabeln verwenden würde.

Doch selbst wenn man diesen wissenschaftstheoretischen Standpunkt ignoriert, wird man kaum an der Feststellung vorbeikommen, daß eine so starke begriffliche und konzeptionelle Veränderung des Unterrichts, wie sie der Karlsruher Physikkurs fordert, nur

zu rechtfertigen ist, wenn sich dadurch die Ergebnisse des Unterrichts signifikant verbessern. Mehrere Erhebungen bei nach dem Karlsruher Physikkurs unterrichteten Klassen erbrachten jedoch keine solche Resultate. Die Schülerinnen und Schüler zeigten überwiegend keine verbesserte Voraussagekompetenz bei physikalischen Problemen [400], [490], [491] insbesondere in den Bereichen Mechanik und Elektrizitätslehre. Stattdessen wurden beispielsweise gravierende Schwierigkeiten mit dem Kraftbegriff festgestellt [400]; die vorgeschlagene Begründung durch Verwechslung von Impuls und Impulsstromstärke und damit von Impuls und Kraft ist gut nachvollziehbar. Lediglich im Bereich der Thermodynamik schienen sich bessere Ergebnisse zu ergeben, was jedoch dadurch relativiert wird, daß diese Disziplin im herkömmlichen Physikunterricht stark vernachlässigt wird und hier eine Vergleichbarkeit von Resultaten bei Erhebungen wohl bezweifelt werden darf. Die überwiegend nicht überzeugenden Ergebnisse der Erhebungen sind um so erstaunlicher, als bei Unterrichtsversuchen schon allein aufgrund der mutmaßlich höheren Motivation aller Beteiligten unabhängig von der Konzeption des untersuchten Unterrichtsganges bessere Ergebnisse zu erwarten sein sollten [400], [498].

Ganz unabhängig von solchen Ergebnissen darf zudem nicht übersehen werden, daß Lernerfolge nicht durch fachlich problematische Unterrichtsgänge erreicht werden sollten. Die Physik ist eine hochentwickelte exakte Naturwissenschaft mit einer präzisen allgemein akzeptierten und verbreiteten Sprache. Die Schulphysik darf nicht versuchen, eigene, in der Wissenschaft unübliche oder sogar unbekannte Begriffe zu etablieren.

Fairerweise sollte ein Aspekt, der im Zusammenhang mit der Fachsprache steht, nicht unterschlagen werden, da hier der Karlsruher Physikkurs eine teilweise vorbildliche, wenn auch nicht neue Sprechweise verwendet. Es handelt sich um den Feldbegriff, der hier sehr viel realistischer gehandhabt wird als in herkömmlichen Darstellungen üblich [218], [223]<sup>57</sup>. Kraftfelder sind reale physikalische Systeme mit ganz konkreten Eigenschaften, die lediglich das Pech haben, unsichtbar zu sein. Diesem Sachverhalt wird üblicherweise nicht genügend Rechnung getragen.

Angesichts der Vehemenz, mit der inzwischen versucht wird, den Karlsruher Physikkurs in Baden-Württembergs Seminaren zu etablieren, und der beschriebenen fachlichen, philosophischen und didaktischen Probleme, die eine Einführung mit sich bringt, erscheint eine etwas konservativere Haltung ratsam. Die etablierten physikalischen Begriffe sind allesamt tragfähig, was sie über lange Zeiträume bewiesen haben, und insbesondere nicht Ursache der Lernschwierigkeiten, die Schülerinnen und Schüler im Fach Physik zweifellos nicht selten haben. Kein Versuch zu deren Überwindung darf zur Abkehr von der bewährten Sprache führen.

Der Karlsruher Physikkurs reklamiert für sich eine einheitliche Darstellung aller physikalischen Disziplinen. Die Einheit der Physik ist jedoch ein Ziel, das derzeit noch nicht einmal als konkretes Forschungsprogramm, bestenfalls als Perspektive vorliegt. Versucht man, dieses Ziel auf der Basis der Gleichgewichtsthermodynamik zu erreichen und noch dazu so, daß das in der Sekundarstufe I unterrichtbar sein soll, nimmt man sich vielleicht doch etwas viel vor.

---

<sup>57</sup>Gleichfalls nicht verschwiegen werden sollte, daß der Karlsruher Physikkurs auch hier gelegentlich über das Ziel hinausschießt, nämlich dann, wenn empfohlen wird, sich Kraftfelder als elastische Gebilde vorzustellen. Hier tritt wieder das generelle Problem auf, daß Schülerinnen und Schüler zur Überzeugung kommen könnten, Kraftfelder seien materielle Gebilde, insbesondere, weil dabei der inakzeptable Ausdruck „Feldstoff“ verwendet wird [229].

Da die Entropie nun einmal derzeit als wesentlicher Bestandteil in den Lehrplänen der Gymnasien auftaucht, kommt man im Unterricht nicht daran vorbei. Folglich ist ein Kompromiß zwischen der erwähnten traditionellen Betrachtungsweise der Wärmelehre und diesem zwar nicht neuen, aber für die Schulphysik in der verlangten Ausführlichkeit neuen Begriff erforderlich. Dieser Kompromiß kann nicht funktionieren, wenn er der Entropie eine wie gesehen nicht vorhandene intuitive Anschaulichkeit andichtet. Wie ein solcher Kompromiß aus unserer Sicht stattdessen aussehen kann, ist in Kapitel 3 dargestellt.

# Kapitel 2

## Bemerkungen zum sogenannten Zeigerformalismus

### 2.1 Quantenmechanik und Schulphysik

Die Behandlung der Quantenmechanik in der gymnasialen Oberstufe Baden-Württembergs ist gekennzeichnet durch ein Umschlagen von einem Extrem in ein anderes. In dem Maß, wie sie im Bildungsplan bis 2003 nur stiefmütterlich behandelt wurde [1], entdeckte man sie in den beiden darauffolgenden Bildungsplänen plötzlich ganz neu [2], [3], was unter anderem dazu führte, daß nun Begriffe wie Verschränkung oder Dekohärenz zu finden sind, die teilweise Gegenstände aktuellster Forschung und in jedem Fall eher als Bestandteile des Hauptstudiums denn der 12. oder 13. Klasse aufzufassen sein dürften; außerdem ist im schriftlichen Abitur in jeder Aufgabe ein Abschnitt zur Quantenmechanik enthalten, sodaß die Schülerinnen und Schüler unabhängig von der Auswahl durch den Fachlehrer nicht um entsprechende Fragestellungen herumkommen. Nun ist gegen eine verstärkte Behandlung der Quantenmechanik im Gymnasium natürlich grundsätzlich überhaupt nichts einzuwenden, ganz im Gegenteil kommt man an der modernen Physik nicht vorbei, wenn man sich dem Humboldtschen Ideal der Allgemeinbildung verpflichtet fühlt. Ein Umdenken bei der Vermittlung der Quantenmechanik in der Oberstufe ist daher zweifellos erforderlich, aber es liegt in der äußerst mathematischen Natur der Sache, daß hier sehr behutsam vorgegangen werden muß.

Die Berücksichtigung dieser Vorgabe lag ganz sicher in der Absicht der Erfinder des Konzeptes, das im Mittelpunkt des vorliegenden Kapitels steht. Daß dabei ähnliche Probleme auftreten wie beim Karlsruher Physikkurs, liegt unter anderem daran, daß hier wie dort das selbe Grundprinzip gymnasialer Physikdidaktik unberücksichtigt blieb: *Begriffliche und fachliche Grundlage jeglicher Schulphysik muß die etablierte Wissenschaft sein.* Es darf weder eine Entwicklung weg von der Hochschulphysik noch der Versuch erfolgen, eigene, von letzterer abweichende Begriffe einzuführen. Gerade die Quantenmechanik ist mehr als alle anderen Teilbereiche der Physik darauf angewiesen, daß sie begrifflich exakt und damit im Sinne der anerkannten Fachsprache unterrichtet wird.

Der Zeigerformalismus, um den es nun gehen soll, macht, wenn er in der Quantenmechanik zum Einsatz kommt, eine Verletzung dieses Prinzips unumgänglich, da zu seiner Rechtfertigung mit der Feynmanschen Pfadintegralformulierung ausgerechnet der schwierigste aller Zugänge zur Quantenmechanik gewählt wird. Dieser ist in den für die

Schulphysik interessierenden Fällen ausgesprochen untauglich<sup>1</sup> und zudem in seiner in der Physik verbreitetsten Erscheinungsform gar nicht definierbar, was aus mathematischer Sicht eine aufwendige Präzisierung beziehungsweise Umformulierung erforderlich macht. Letztere ist zudem nicht einheitlich, da es mehrere unterschiedliche solche Präzisierungen gibt, und vor allem gar nicht allgemein möglich. *Die Quantisierung mit Pfadintegralen kann daher nicht als grundlegendes, konzeptionell mit der kanonischen Quantisierung vergleichbares Prinzip betrachtet werden.*

All das erweist sich aus schulpraktischer Sicht noch als das kleinste Problem. Bei genauerer Betrachtung zeigt sich zwar, daß die in der Oberstufe unterrichtete „Zeigerformalismus-Quantenmechanik“ aus einer Fehlinterpretation der Pfadintegrale entstanden ist, was die grundsätzliche Verwendbarkeit im Unterricht natürlich mehr als in Frage stellt, aber immerhin nicht zu falschen Vorhersagen führt. Schwerer wiegt die trügerische Anschaulichkeit dieses Zugangs, die den Schülerinnen und Schülern die irreführende Möglichkeit vorgaukelt, sich quantenmechanische Vorgänge in klassischen Bildern vorstellen zu können. Das wäre auch dann der Fall, wenn Pfadintegrale in der Oberstufe angemessen behandelt werden könnten, wovon natürlich keine Rede sein kann. Als Mittel zum Unterrichten von Quantenmechanik sind sie höchstens dann geeignet, *wenn bereits fundierte Vorstellungen derselben vorhanden sind.*

Um sich aus didaktischer Sicht ein Urteil über die Verwendung des Zeigerformalismus zur Behandlung der Quantenmechanik in der gymnasialen Oberstufe machen zu können, sind Kenntnisse über Pfadintegrale nicht unbedingt erforderlich, so daß man die nun folgende elementare Einführung derselben auch überlesen kann. Dennoch kann es sicher nicht schaden, ein wenig darüber zu wissen. Der mathematische Aufwand einer präzisen Diskussion der Pfadintegrale ist beträchtlich, weswegen wir die Sache in zwei Schritten bewältigen. Die folgenden Abschnitte liefern zunächst eine Darstellung der physikalischen, formal jedoch unsaubereren Betrachtungsweise, die beim Leser lediglich voraussetzt, daß er mit der elementaren Quantenmechanik in ihrer Standard-Hilbertraum-Version vertraut ist; erst anschließend widmen wir uns einer mathematisch exakten Betrachtung des Gegenstands aus mehreren unterschiedlichen Blickwinkeln, wofür jeweils etwas weitergehende Kenntnisse der Funktionalanalysis sowie einige Grundbegriffe aus der Maßtheorie und der Theorie der stochastischen Prozesse erforderlich sind. Mit konkreten Berechnungen von Pfadintegralen werden wir uns hier nicht beschäftigen<sup>2</sup>. Wir setzen in diesem Kapitel durchweg  $\hbar = 1$  und in Hamilton-Operatoren zusätzlich auch  $\hbar = 2m$ , sofern nicht ausdrücklich etwas anderes gesagt wird.

<sup>1</sup>Die vollständige Lösung des quantenmechanischen nichtrelativistischen Kepler-Problems mit Pfadintegralen beispielsweise wurde erst 1979 von Duru und Kleinert gefunden [132], [133], [310], [311], [312]; vergleiche auch [251] und [264]. Möglich wurde das erst durch Einsatz raffinierter Tricks und einem schon einige Jahre zuvor entwickelten Verfahren aus der Himmelsmechanik, der sogenannten Kustaanheimo-Stiefel-Transformation [339], die das dreidimensionale Kepler-Problem auf das Problem vierdimensionaler harmonischer Oszillatoren transformiert. Etwas später gelang die Lösung mit Pfadintegralen auch für das zweidimensionale [133], [263] und das allgemeine  $d$ -dimensionale Kepler-Problem [98], [251], [493] sowie auch für die Dirac-Gleichung mit Coulomb-Potential und damit für das relativistische Kepler-Problem [306]. Generell sind Berechnungen mit Pfadintegralen im Rahmen der nichtrelativistischen Einteilchen-Quantenmechanik selbst im Vergleich zu den Standardverfahren überaus aufwendig. Praktische Bedeutung erlangen Pfadintegrale hauptsächlich erst in der relativistischen Quantenmechanik, der Quantenfeldtheorie und der statistischen Mechanik.

<sup>2</sup>Ausführliche Beispiele zur Berechnung von Pfadintegralen liefern beispielsweise [203], [313] und [441].



## 2.2 Pfadintegrale in der Quantenmechanik – heuristischer Zugang

Die Feynmansche Formulierung der Quantenmechanik stellt den Versuch dar, die Quantisierung eines physikalischen Systems im Gegensatz zum herkömmlichen Zugang nicht über die Ersetzung klassischer Größen durch lineare Operatoren, sondern durch die ausschließliche Verwendung der klassischen Wirkungsfunktion des Systems zu bewerkstelligen. Diese Idee war gleichzeitig überhaupt die erste, wenn auch zunächst noch nicht zur vollen praktischen Einsatzfähigkeit entwickelte Formulierung der neueren Quantenmechanik, denn die wesentlichen Prinzipien finden sich bereits 1924 in einer Arbeit von Gregor Wentzel [517], die noch vor Heisenbergs Matrizenmechanik (1925) und Schrödingers Wellenmechanik (1926) veröffentlicht wurde<sup>3</sup>. Wentzels Pionierleistung war danach lange Zeit vergessen; erst 1996 wurde sie von Antoci und Liebscher als eine solche wiederentdeckt [15]. Vergleichbare Ansätze tauchen eher beiläufig etwas später auch bei Dirac auf [121], [122], [123]; in beiden Fällen wies die formale Präzision natürlich das für die Pionier-Quantenmechanik der Prä-von Neumann-Ära typische eher bescheidene Ausmaß auf. Feynman war, darauf aufbauend, dann der erste, der die Pfadintegral-Formulierung der Quantenmechanik zur Praxisreife und damit auch in das Bewußtsein der physikalischen Welt brachte [165], [168]. Allerdings führte auch er seinen Zugang auf heuristische, mathematisch in keiner Weise sauber definierten Form ein. Dem Mangel an Präzision steht jedoch der Vorteil dieses Zugangs gegenüber, ein sehr anschauliches, intuitives Verständnis seiner physikalischen Interpretation zu ermöglichen – wenn man bereits ein solches für herkömmliche Zugänge der Quantenmechanik entwickelt hat. Dieses Vorverständnis ist unter anderem deshalb wichtig, weil ohne ein solches eine fehlerhafte, vermeintlich klassische Interpretation der Pfadintegrale nahezu unvermeidlich ist. Wir betrachten zunächst diese heuristische Herleitung, unter ständiger Beachtung der soeben notierten Bemerkungen.

### 2.2.1 Der Schrödinger-Propagator

In der Quantenmechanik wird die Dynamik eines Systems bekanntlich durch einen zeitabhängigen Zeitentwicklungsoperator  $\hat{U}(t_2, t_1)$  beschrieben. Er wird bestimmt durch die Operator-Differentialgleichung

$$i \frac{\partial \hat{U}(t_2, t_1)}{\partial t} = \hat{H} \hat{U}(t_2, t_1), \quad (2.1)$$

wobei  $\hat{H}$  der Hamilton-Operator des Systems ist. Die zeitliche Entwicklung des Zustandsvektors ist damit

$$|\psi(t)\rangle = \hat{U}(t, t_0) |\psi(t_0)\rangle, \quad (2.2)$$

und Ableiten nach  $t$  liefert unter Verwendung von (2.1)

$$i \frac{\partial |\psi(t)\rangle}{\partial t} = \hat{H} |\psi(t)\rangle,$$

---

<sup>3</sup>Wentzel ist insbesondere durch seine zusammen mit Kramers und Brillouin entwickelte halbklassische Näherungsmethode bekannt, die in der Literatur üblicherweise als WKB-Methode bezeichnet wird.

also die zeitabhängige Schrödingergleichung.

Die grundlegende Idee der Pfadintegralformulierung der Quantenmechanik besteht nun darin, eine Integraldarstellung von (2.2) zu finden. Durch Einschleiben der dualen Darstellung des 1-Operators,

$$\mathbf{1} = \int |x\rangle \langle x| dx$$

bei der Ortsdarstellung von  $\psi$  ergibt sich zunächst<sup>4</sup>

$$\begin{aligned} \psi(x_2, t_2) &= \langle x_2 | \psi(t_2) \rangle = \int \langle x_2 | \hat{U}(t_2, t_1) | x_1 \rangle \langle x_1 | \psi(t_1) \rangle dx_1 \\ &= \int \langle x_2 | \hat{U}(t_2, t_1) | x_1 \rangle \psi(x_1, t_1) dx_1. \end{aligned}$$

Mit Hilfe der Heavisideschen Stufenfunktion

$$\theta(t) = \begin{cases} 0 & \text{für } t < 0 \\ 1 & \text{für } t \geq 0 \end{cases}$$

definieren wir den *retardierten Propagator*

$$K(x_2, t_2; x_1, t_1) = \langle x_2 | \hat{U}(t_2, t_1) | x_1 \rangle \theta(t_2 - t_1)$$

und erhalten wie gewünscht

$$\psi(x_2, t_2) = \int K(x_2, t_2; x_1, t_1) \psi(x_1, t_1) dx_1. \quad (2.3)$$

Man beachte: Bei  $dx$  und bei  $\psi$  erscheint  $x$  jeweils mit demselben Index.  $\psi(x_2, t_2)$  entsteht also aus einer Überlagerung aller Zustände, die sich aus den  $\psi(x_1, t_1)$  mit beliebigem  $x_1$  entwickeln. Das entspricht genau dem Huygenschen Prinzip der Wellenoptik: Den Wert der Wellenfunktion  $\psi$  am Ort  $x_2$  zum Zeitpunkt  $t_2$  erhält man als gewichtete Summe der Werte der Wellenfunktion an allen Orten  $x_1$  zum Zeitpunkt  $t_1 < t_2$ .

Wegen  $\langle x_2 | x_1 \rangle = \delta(x_1 - x_2)$  gilt insbesondere

$$K(x_2, t_0; x_1, t_0) = \delta(x_1 - x_2).$$

Wir betrachten den speziellen Fall  $|\psi(t_1)\rangle = |x_1\rangle$ , das heißt ein zum Zeitpunkt  $t_1$  bei  $x_1$  lokalisiertes System. Der Zustand  $|\psi\rangle = |x_1\rangle$  liegt unmittelbar nach einer Ortsmessung mit dem Ergebnis  $x_1$  vor. Zum Zeitpunkt  $t_2 > t_1$  gilt dann

$$|\psi(t_2)\rangle = \hat{U}(t_2, t_1)|x_1\rangle,$$

---

<sup>4</sup>Die Eigenvektoren  $|x\rangle$  des Ortsoperators sind keine quadratintegrablen Funktionen und damit eigentlich keine Elemente des Hilbertraums der Zustände des betrachteten Systems. Diese scheinbare Schwierigkeit läßt sich jedoch mathematisch präzise bewältigen. Erweitert man den Hilbertraum zu einem sogenannten *Gelfandschen Raumtripel*, so lassen sich solche wilden Zustände als *Eigendistributionen* der zugehörigen Operatoren auffassen. Man kann zeigen, daß Faltungen von Zustandsvektoren mit solchen Eigendistributionen wieder auf Zustandsvektoren führen, womit gleichzeitig der Begriff der *Wellenpakete* eine exakte Definition findet. Genauer dazu steht beispielsweise in [378]. Gelfandsche Raumtripel werden uns in einem ganz anderen Zusammenhang in Abschnitt 2.6 erneut begegnen.



also in Ortsdarstellung

$$\psi(x_2, t_2) = \langle x_2 | \psi(t_2) \rangle = \langle x_2 | \hat{U}(t_2, t_1) | x_1 \rangle = K(x_2, t_2; x_1, t_1).$$

Der Propagator  $K(x_2, t_2; x_1, t_1)$  ist somit die Ortsdarstellung des Zustandes, der sich unter dem Einfluß des Zeitentwicklungsoperators aus einem bei  $x_1$  zum Zeitpunkt  $t_1$  lokalisierten Zustand zum Zeitpunkt  $t_2 > t_1$  entwickelt hat. Das bedeutet: Führt man zuerst zum Zeitpunkt  $t_1$  eine Ortsmessung mit dem Ergebnis  $x_1$  durch, dann zum Zeitpunkt  $t_2$  eine weitere Ortsmessung mit dem Ergebnis  $x_2$ , so ist die Zustandsfunktion unmittelbar danach genau  $K(x_2, t_2; x_1, t_1)$ , denn es gilt

$$\langle x_2 | \psi(t_1) \rangle = \langle x_2 | x_1 \rangle = \delta(x_1 - x_2).$$

### 2.2.2 Pfadintegral-Darstellung des Schrödinger-Propagators

Um den Propagator  $K(x_b, t_b; x_a, t_a)$  etwas genauer zu betrachten, notieren wir zunächst, daß für den Zeitentwicklungsoperator die Beziehung

$$\hat{U}(t_a, t_b) = \hat{U}(t_a, t) \hat{U}(t, t_a), \quad t_b > t > t_a.$$

gilt, und zerlegen sodann das Zeitintervall  $[t_a, t_b]$  in  $n$  äquidistante Abschnitte der Länge  $\frac{t_b - t_a}{n} = t_{i+1} - t_i \equiv \varepsilon$ . Anschließend betrachten wir den Grenzfall  $n \rightarrow \infty$ , das heißt  $\varepsilon \rightarrow 0$ . Der Zeitentwicklungsoperator läßt sich damit als

$$\hat{U}(t_a, t_b) = \prod_{i=0}^{n-1} \hat{U}(t_{i+1}, t_i), \quad t_0 = t_a, \quad t_n = t_b$$

schreiben. Zwischen die  $\hat{U}$ 's wird nun jedesmal ein **1**-Operator eingeschoben. Für nur eine Zwischenstelle bedeutet dies

$$K(x_b, t_b; x_a, t_a) = \int \langle x_b | \hat{U}(t_b, t) | x \rangle \langle x | \hat{U}(t, t_a) | x_a \rangle dx,$$

also

$$K(x_b, t_b; x_a, t_a) = \int K(x_b, t_b; x, t) K(x, t; x_a, t_a) dx. \quad (2.4)$$

Es wird also über alle Wege integriert, die von  $x_a$  zum Zeitpunkt  $t_a$  über irgendeinen Punkt  $x$  zum Zeitpunkt  $t$  nach  $x_b$  zum Zeitpunkt  $t_b$  führen. (2.4) nennt man *Chapman-Kolmogoroff-Gleichung*<sup>5</sup>.

Nun lassen wir die Anzahl der Zwischenstellen gegen Unendlich gehen:

$$K(x_b, t_b; x_a, t_a) = \lim_{n \rightarrow \infty} \iiint \cdots \int \prod_{i=0}^{n-1} K(x_{i+1}, t_{i+1}; x_i, t_i) dx_1 dx_2 dx_3 \cdots dx_{n-1}. \quad (2.5)$$

<sup>5</sup>Diese Gleichung ist in der Theorie der stochastischen Prozesse von großer Bedeutung; sie deutet gleichzeitig darauf hin, daß die Propagatoren  $K(x, t; x_a, t_a)$  eine Halbgruppe bilden, was in Abschnitt 2.3 von Interesse sein wird.

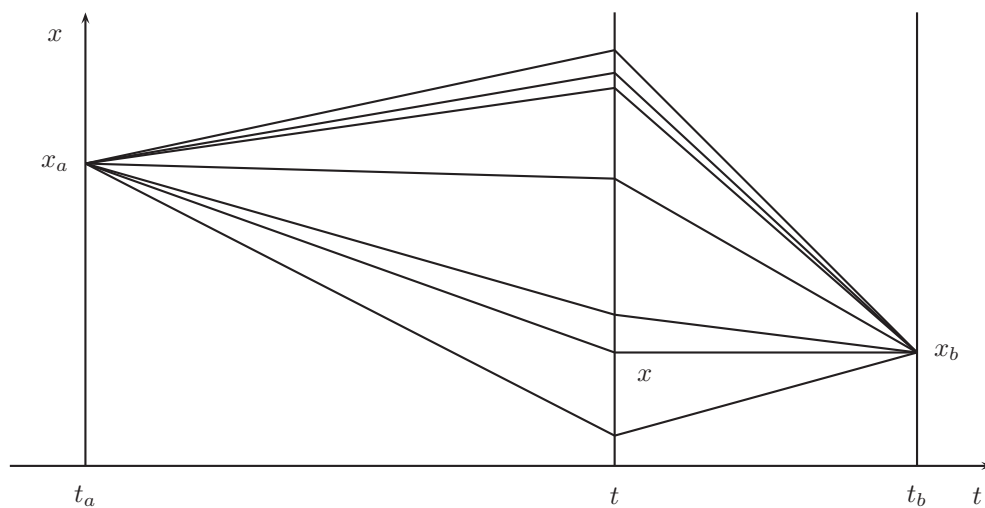


Abbildung 2.1: Summation über alle Wege bei einer Zwischenstelle

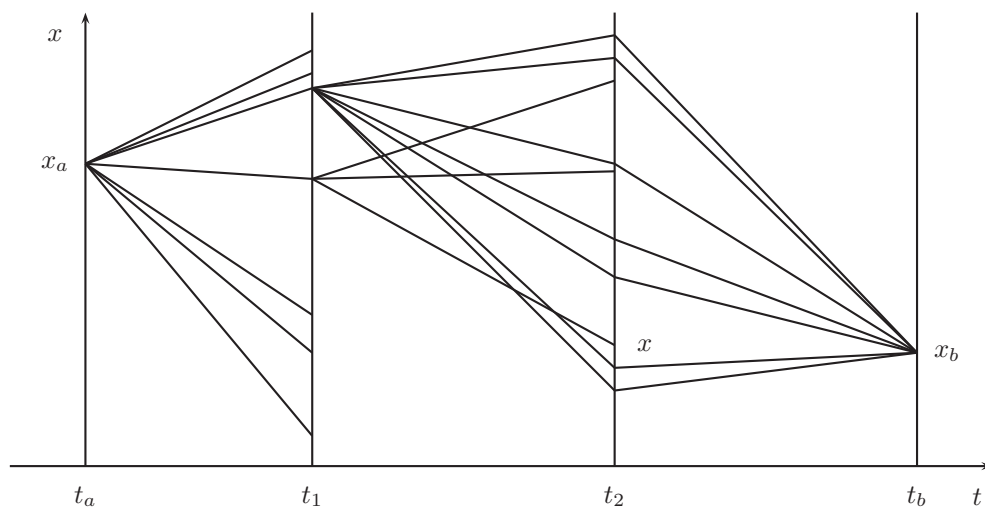


Abbildung 2.2: Summation über alle Wege bei zwei Zwischenstellen

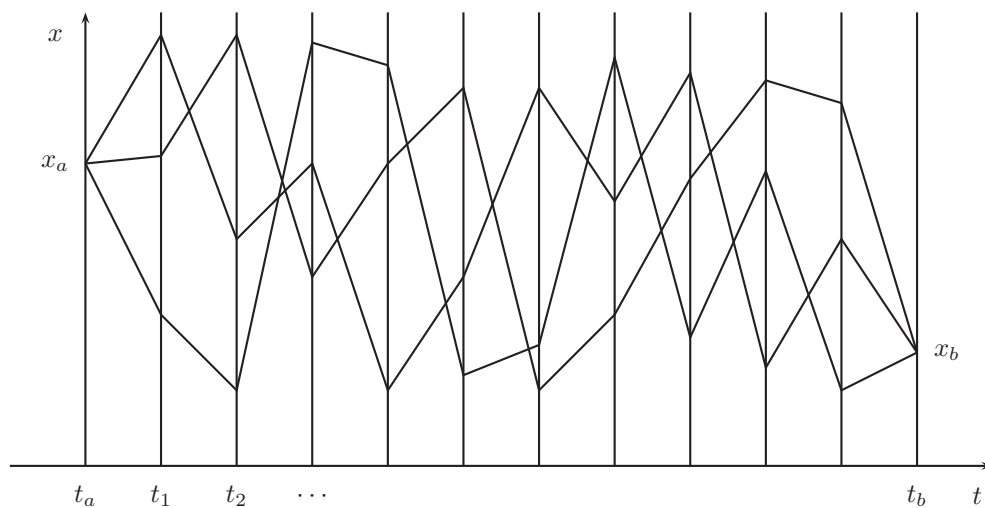
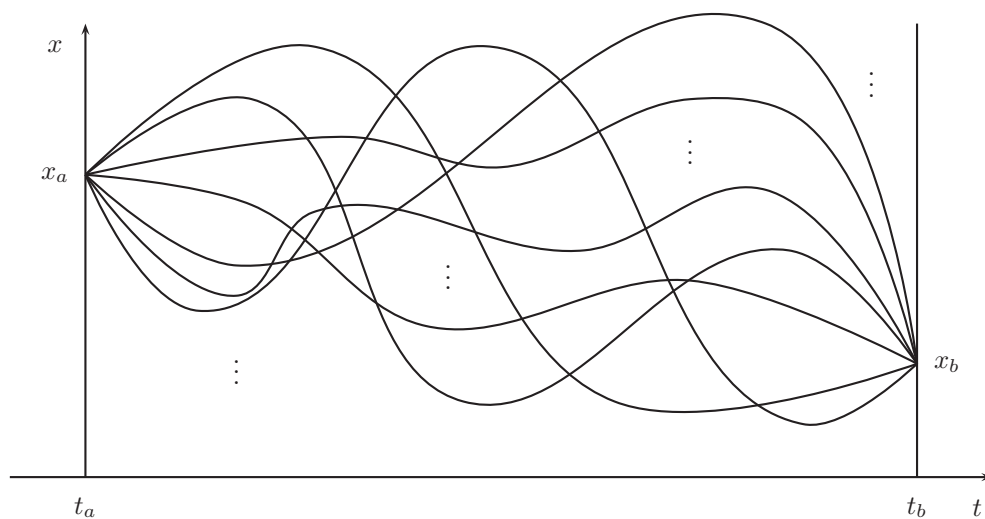


Abbildung 2.3: Summation über alle Wege bei zehn Zwischenstellen

Abbildung 2.4: Summation über alle Wege von  $x_a$  nach  $x_b$  im kontinuierlichen Grenzfall

Die einzelnen Raum-Zeit-Punkte auf den Achsen der Zwischenstellen decken damit den Bereich der Raum-Zeit zwischen  $x_a$  und  $x_b$  vollständig ab. Es wird folglich über alle stetigen Raum-Zeit-Wege von  $x_a$  nach  $x_b$ , also über die Menge  $\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}$  aller stetigen Funktionen auf  $[t_a, t_b]$  mit  $x(t_a) = x_a$  und  $x(t_b) = x_b$  integriert<sup>6</sup>.

Wir berechnen nun den Propagator unter folgenden speziellen Voraussetzungen: Das betrachtete System bestehe aus  $N$  Teilchen mit den Massen  $m_i$ ,  $i = 1, 2, \dots, N$ , für seinen Hamilton-Operator gelte  $\partial\hat{H}/\partial t = 0$  und außerdem  $\hat{H} = \hat{T} + \hat{V}$ . Dabei ist

$$\hat{T} = \sum_{i=1}^N \frac{\hat{P}_i^2}{2m_i}$$

der Operator der kinetischen Energie<sup>7</sup> und  $\hat{V} = \hat{V}(x)$  der Operator der potentiellen Energie, der nur von  $x = (x_1, x_2, \dots, x_{3N})$  und nicht von  $\dot{x}$  abhängig sein soll<sup>8</sup>. Daraus folgt zunächst nach der nunmehr möglichen direkten Integration von (2.1)

$$K(x_{i+1}, t_{i+1}; x_i, t_i) = \langle x_{i+1} | e^{-i\hat{H}\varepsilon} | x_i \rangle$$

mit  $\varepsilon = t_{i+1} - t_i = \Delta t$ . Für den Exponentialausdruck gilt einerseits bis zur zweiten Ordnung in  $\varepsilon$

$$e^{(\hat{A} + \hat{B})\varepsilon} = 1 + (\hat{A} + \hat{B})\varepsilon + O(\varepsilon^2);$$

andererseits findet man ebenfalls bis zur zweiten Ordnung

$$e^{\hat{A}\varepsilon} e^{\hat{B}\varepsilon} = [1 + \hat{A}\varepsilon + O(\varepsilon^2)][1 + \hat{B}\varepsilon + O(\varepsilon^2)] = 1 + \hat{A}\varepsilon + \hat{B}\varepsilon + O(\varepsilon^2) = e^{\hat{B}\varepsilon} e^{\hat{A}\varepsilon}$$

Damit folgt in erster Näherung

$$\langle x_{i+1} | e^{-i(\hat{T} + \hat{V})\varepsilon} | x_i \rangle = \langle x_{i+1} | e^{-i\hat{T}\varepsilon} e^{-i\hat{V}\varepsilon} | x_i \rangle = \langle x_{i+1} | e^{-i\hat{T}\varepsilon} | x_i \rangle e^{-iV(x_i)\varepsilon}. \quad (2.6)$$

Für das Matrix-Element gilt (wieder nach Einschieben eines **1**-Operators)

$$\begin{aligned} \langle x_{i+1} | e^{-i\hat{T}\varepsilon} | x_i \rangle &= \int_{\mathbb{R}^{3N}} \langle x_{i+1} | e^{-i\hat{T}\varepsilon} | p \rangle \langle p | x_i \rangle dp \\ &= \int_{\mathbb{R}^{3N}} \langle x_{i+1} | p \rangle \langle p | x_i \rangle \exp\left(-i\varepsilon \sum_{k=1}^N \frac{p_k^2}{2m_k}\right) dp, \end{aligned}$$

<sup>6</sup>Auch wenn im vorliegenden Abschnitt mathematische Details weitgehend ausgeblendet werden, ist an dieser Stelle doch eine Bemerkung angebracht. Es wird hier stillschweigend vorausgesetzt, daß die Zick-Zack-Wege beim Grenzübergang  $n \rightarrow \infty$  in *alle möglichen stetigen Wege* übergehen; das ist auch richtig, aber nicht trivial. Der Grund für diesen Sachverhalt liegt darin, daß die Menge aller Polygonzüge von  $x_a$  nach  $x_b$  dicht in  $\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}$  ist, wenn man dabei die durch die Maximumsnorm induzierte Topologie zugrunde legt.

<sup>7</sup>Natürlich setzen wir hier stillschweigend kartesische Koordinaten voraus. Entsprechende Ausdrücke in verallgemeinerten Koordinaten erhält man durch geeignete Koordinatentransformationen; allerdings ist es dabei wesentlich, erst die komplette kanonische Quantisierungsprozedur durchlaufen zu lassen und die Transformationen erst anschließend an den resultierenden Operatoren vorzunehmen, da man andernfalls im allgemeinen einen unvollständigen Hamilton-Operator erhält.

<sup>8</sup>Nelson und Sheeks diskutieren in [387] eine Verallgemeinerung auf geschwindigkeitsabhängige Potentiale.

und wegen

$$\langle x_{i+1}|p\rangle = \frac{1}{(2\pi)^{3N/2}} e^{ipx_{i+1}}$$

weiter

$$\begin{aligned} \langle x_{i+1}|e^{-i\hat{T}\varepsilon}|x_i\rangle &= \frac{1}{(2\pi)^{3N/2}} \int_{\mathbb{R}^{3N}} \exp[ip(x_{i+1} - x_i)] \exp\left(-i\varepsilon \sum_{k=1}^N \frac{p_k^2}{2m_k}\right) dp \\ &= \frac{1}{(2\pi)^{3N/2}} \int_{\mathbb{R}^{3N}} \exp\left\{i \sum_{k=1}^N \left[-\frac{\varepsilon p_k^2}{2m_k} + p_k(x_{k,i+1} - x_{k,i})\right]\right\} dp \\ &= \frac{1}{(2\pi)^{3N/2}} \prod_{k=1}^N \int_{\mathbb{R}^3} \exp\left\{i \left[-\frac{\varepsilon p_k^2}{2m_k} + p_k(x_{k,i+1} - x_{k,i})\right]\right\} dp_k \\ &= \frac{1}{(2\pi)^{3N/2}} \prod_{l=1}^{3N} \int_{-\infty}^{\infty} \exp\left\{i \left[-\frac{\varepsilon p_l^2}{2m_l} + p_l(x_{l,i+1} - x_{l,i})\right]\right\} dp_l, \end{aligned} \quad (2.7)$$

wobei der Index  $l$  in der vierten Zeile über die skalaren Komponenten von  $p$ ,  $x_i$  und  $x_{i+1}$  läuft und für die Massen folglich  $m_k = m_{3l-2} = m_{3l-1} = m_{3l}$  gilt; der Vorfaktor kommt von der Normierung. Die hier auftretenden Integrale sind Gauß-Integrale der Form

$$I = \int_{-\infty}^{\infty} e^{i(\alpha x^2 - \beta x)} dx \quad (2.8)$$

und existieren im Lebesgueschen Sinne natürlich nicht. Man kann sie jedoch als uneigentliche Integrale auffassen und mit einigen Tricks berechnen.

Zunächst bringen wir (2.8) durch quadratische Ergänzung und anschließende Substitution  $t = \alpha x - \beta/2\alpha$  auf die Form

$$\int_{-\infty}^{\infty} e^{i(\alpha x - \beta/2\alpha)^2 - i\beta^2/4\alpha} dx = e^{-i\beta^2/4\alpha} \int_{-\infty}^{\infty} e^{i(\alpha x - \beta/2\alpha)^2} dx = \frac{e^{-i\beta^2/4\alpha}}{\alpha} \int_{-\infty}^{\infty} e^{it^2} dt.$$

Zur Berechnung des hierbei resultierenden Integrals betrachten wir stattdessen die nach dem Cauchyschen Integralsatz geltende Relation

$$\oint_{\Gamma} e^{iz^2} dz = 0$$

und wählen dabei als Integrationsweg einen auf der reellen Achse beginnenden Achtelkreis mit Radius  $R$  und dem Ursprung als Mittelpunkt (Abbildung 2.5). Wir werten zunächst dieses Integral aus und erhalten das gewünschte Integral anschließend mit Hilfe des Grenzübergangs  $R \rightarrow \infty$ . Zerlegt man  $\Gamma$  in die drei Teilwege  $OA$ ,  $AB$  und  $BO$ , so schreibt sich (2.9)

$$\int_{OA} e^{iz^2} dz + \int_{AB} e^{iz^2} dz + \int_{BO} e^{iz^2} dz = 0. \quad (2.9)$$

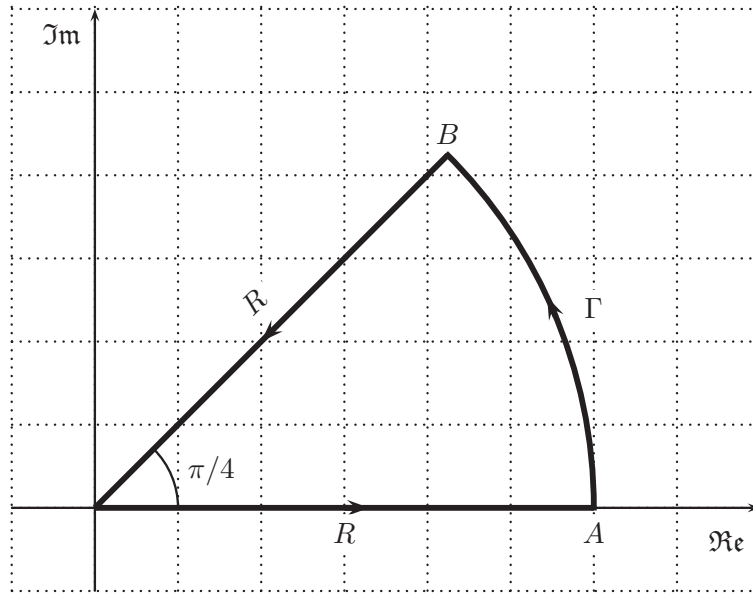


Abbildung 2.5: Integrationsweg für das Kurvenintegral (2.9)

Diese Teilwege lassen sich leicht parametrisieren; man erhält

$$\begin{aligned} \text{Für } OA : z &= x && \text{mit } x \in [0, R], \\ \text{Für } AB : z &= Re^{i\theta} && \text{mit } \theta \in [0, \pi/4], \\ \text{Für } BO : z &= -re^{i\pi/4} && \text{mit } r \in [0, R] \end{aligned}$$

und damit

$$\int_0^R e^{ix^2} dx + \int_0^{\pi/4} e^{iR^2 e^{2i\theta}} iRe^{i\theta} d\theta - \int_0^R e^{-r^2} e^{i\pi/4} dr = 0.$$

Das zweite Integral schreiben wir in der Form

$$\int_0^{\pi/4} e^{iR^2 e^{2i\theta}} iRe^{i\theta} d\theta = \int_0^{\pi/4} e^{iR^2 \cos 2\theta - R^2 \sin 2\theta} iRe^{i\theta} d\theta$$

und schätzen es anschließend zunächst ab zu

$$\left| \int_0^{\pi/4} e^{iR^2 \cos 2\theta - R^2 \sin 2\theta} iRe^{i\theta} d\theta \right| \leq \int_0^{\pi/4} e^{-R^2 \sin 2\theta} R d\theta = \frac{R}{2} \int_0^{\pi/2} e^{-R^2 \sin \phi} d\phi.$$

Wie man sich leicht überzeugt, gilt für  $0 \leq \phi \leq \pi/2$  die Ungleichung  $\sin \phi \geq 2\phi/\pi$ , und wir finden weiter

$$\frac{R}{2} \int_0^{\pi/2} e^{-R^2 \sin \phi} d\phi \leq \frac{R}{2} \int_0^{\pi/2} e^{-2R^2 \phi/\pi} d\phi = \frac{\pi}{4R} (1 - e^{-R^2}).$$

Damit folgt

$$\lim_{R \rightarrow \infty} \int_0^{\pi/4} e^{iR^2 e^{2i\theta}} iR e^{i\theta} d\theta = 0. \quad (2.10)$$

Das dritte Integral können wir für  $R \rightarrow \infty$  direkt auswerten. Wir betrachten dazu das Doppelintegral

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)} dx dy = \int_{-\infty}^{\infty} e^{-x^2} dx \int_{-\infty}^{\infty} e^{-y^2} dy,$$

das wir durch Transformation auf Polarkoordinaten ausrechnen können:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)} dx dy = \int_0^{\infty} \int_0^{2\pi} e^{-r^2} r dr d\phi = 2\pi \left[ -\frac{1}{2} e^{-r^2} \right]_0^{\infty} = \pi.$$

Es ergibt sich also

$$\int_{-\infty}^{\infty} e^{-x^2} dx \int_{-\infty}^{\infty} e^{-y^2} dy = \pi,$$

damit weiter

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$$

und aufgrund der Symmetrie des Integranden schließlich

$$\int_0^{\infty} e^{-x^2} dx = \frac{\sqrt{\pi}}{2}.$$

Für das dritte Integral gilt folglich

$$e^{i\pi/4} \int_0^{\infty} e^{-r^2} dr = \frac{\sqrt{\pi}}{2} e^{i\pi/4}. \quad (2.11)$$

Aus (2.9), (2.10) und (2.11) erhält man nun

$$\int_0^{\infty} e^{ix^2} dx = \frac{\sqrt{i\pi}}{2}$$

und für (2.8) folglich

$$\int_{-\infty}^{\infty} e^{i(\alpha x^2 - \beta x)} dx = e^{-i\beta^2/4\alpha} \frac{\sqrt{i\pi}}{2\alpha}$$

Damit findet man nach Einsetzen in (2.7) mit  $\alpha = -\varepsilon/2m_l$  und  $\beta = x_{l,i} - x_{l,i+1}$

$$\langle x_{i+1} | e^{-i\hat{T}\varepsilon} | x_i \rangle = \frac{1}{(2i\varepsilon\pi)^{3N/2}} \prod_{k=1}^N m_k^{3/2} \exp \left[ \frac{i}{2\varepsilon} \sum_{k=1}^N m_k (x_{k,i+1} - x_{k,i})^2 \right]$$



und erhält für den Propagator

$$K(x_{i+1}, t_{i+1}; x_i, t_i) = \frac{1}{(2i\varepsilon\pi)^{3N/2}} \prod_{k=1}^N m_k^{3/2} \exp \left\{ \frac{\varepsilon i}{2} \left[ \sum_{k=1}^N m_k \left( \frac{x_{k,i+1} - x_{k,i}}{\varepsilon} \right)^2 - V(x_i) \right] \right\}.$$

Wegen  $\varepsilon = \Delta t$  kann man die Abkürzung

$$\frac{x_{i+1} - x_i}{\varepsilon} = \frac{\Delta x}{\Delta t}$$

einführen; für den Fall unendlich vieler Unterteilungen ( $n \rightarrow \infty$ ) gilt daher

$$\lim_{\Delta t \rightarrow 0} \frac{\Delta x}{\Delta t} = v = \dot{x}$$

und damit auch

$$\sum_{k=1}^N \frac{m_k}{2} \left( \frac{x_{k,i+1} - x_{k,i}}{\varepsilon} \right)^2 - V(x_i) = \frac{1}{2} \sum_{k=1}^N m_k \dot{x}_k^2 - V(x_i) = T - V.$$

Man startet also in der Rechnung mit  $\hat{T} + \hat{V} = \hat{H}$  und landet nach Auswertung des Integrals bei  $T - V = L$ , also der *klassischen* Lagrange-Funktion. Der Übergang von der unquantisierten Hamilton-Funktion  $H$  zum quantisierten Hamilton-Operator  $\hat{H}$  liefert nach Durchführung obiger Rechnung die klassische Lagrange-Funktion im Exponenten. Für den Propagator gilt damit

$$K(x_{i+1}, t_{i+1}; x_i, t_i) = \frac{1}{(2i\varepsilon\pi)^{3N/2}} \prod_{k=1}^N m_k^{3/2} e^{iL(x, \dot{x})\Delta t}.$$

Setzt man dies in  $K(x_b, t_b; x_a, t_a)$  ein, so findet man

$$K(x_b, t_b; x_a, t_a) = \lim_{\varepsilon \rightarrow 0} \frac{1}{A^n} \iint \cdots \int \exp \left[ i \sum_{i=0}^{n-1} L_{\text{kl}}(x_{i+1}, t_{i+1}; x_i, t_i) \right] \times \\ \times \varepsilon dx_1 dx_2 \cdots dx_{n-1},$$

mit

$$A := \frac{(2i\varepsilon\pi)^{3N/2}}{\prod_{k=1}^N m_k^{3/2}}. \quad (2.12)$$

Führen wir die klassische Wirkung

$$S_{\text{kl}} = \int L_{\text{kl}}(x(t), \dot{x}(t)) dt,$$

ein und betrachten den Grenzfall  $n \rightarrow \infty$ , erhalten wir für den Propagator

$$K(x_b, t_b; x_a, t_a) = \lim_{n \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} \frac{1}{A^n} \iint \cdots \int \exp \left[ i \int L_{\text{kl}}(x, \dot{x}) \right] dx_1 dx_2 \cdots dx_{n-1} dt$$

$$= \lim_{n \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} \frac{1}{A^n} \iint \cdots \int e^{iS_{kl}} dx_1 dx_2 \cdots dx_{n-1}. \quad (2.13)$$

Das Intervall  $[x_a, t_a; x_b, t_b]$  wird hierbei unendlich oft unterteilt, und an jeder Teilungsstelle wird jeweils über  $x_i$  integriert. Dies gibt Anlaß zu der abkürzenden Schreibweise

$$K(x_b, t_b; x_a, t_a) = \int_{\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}} e^{iS_{kl}(x(t))} \mathcal{D}x(t). \quad (2.14)$$

Dabei ist

$$\mathcal{D}x(t) = \lim_{\varepsilon \rightarrow 0} \prod_{i=1}^{\infty} \frac{dx_i}{A}$$

ein Maß auf der Menge aller Wege von  $(x_a, t_a)$  nach  $(x_b, t_b)$ , das alle Elemente dieser Menge gleich gewichtet und uns gleich noch weiter beschäftigen wird. Folglich bedeutet das Symbol  $\int_{\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}} \cdots \mathcal{D}x(t)$  Integration über alle Wege von  $(x_a, t_a)$  nach  $(x_b, t_b)$ . Da man

diese Wege als Graphen von Funktionen auffassen kann, handelt sich um eine Funktionalintegration, hier speziell um die Integration des Funktionals  $e^{iS_{kl}(x(t))}$ , mit den Funktionen  $x(t)$  (den Wegen) als Argumenten. Ein solches Integral heißt *Feynmansches Pfadintegral*. Für die Wellenfunktion des betrachteten Systems gilt damit die Zeitentwicklung

$$\psi(x_2, t_2) = \int \int_{\mathcal{C}_{x_1, x_2}^{[t_1, t_2]}} e^{iS_{kl}(x(t))} \psi(x_1, t_1) \mathcal{D}x(t) d^{3N} x_1. \quad (2.15)$$

Die Lagrange-Funktion wird bei einer solchen Integration nicht nur entlang der klassischen Wege ausgewertet, auf denen gemäß dem Hamiltonschen Prinzip die Wirkung extremalisiert wird, sondern längs aller mathematisch möglicher Wege von  $(x_a, t_a)$  nach  $(x_b, t_b)$ . Die verschiedenen Wege liefern dabei im Integranden gleiche Beiträge, aber verschiedene Phasen; diese sind die klassischen Wirkungen längs der Wege. Durch Superposition aller Wege und destruktive und konstruktive Interferenz erhält man schließlich den Propagator  $K$ . Da es sich bei der Integration über alle Wege um Raum-Zeit-Wege handelt, spricht man auch von „sum over histories“. Gleichzeitig läßt sich die bisher nur angedeutete Analogie zum Huygensschen Prinzip präzisieren: Die Wellenfunktion  $\psi(x_0 + \delta x, t + \delta t)$  eines Systems erhält man als Superposition von Wellenfunktionen dieses Systems an allen Raumpunkten  $x$  zum Zeitpunkt  $t$ , wobei diese Wellenfunktionen alle den selben Betrag, aber unterschiedliche Phasen aufweisen. Die Phasenänderung ist dabei jeweils proportional zur Wirkung, die das System aufsammeln würde, wenn es sich gemäß dem Hamiltonschen Prinzip der klassischen Mechanik von  $(x_0, t)$  nach  $(x_0 + \delta x, t)$  bewegen würde.

Wir erhalten damit eine alternative Axiomatik der Quantenmechanik. Sie erlangt ihre eigentliche Bedeutung erst in der statistischen Mechanik<sup>9</sup> und insbesondere in der Quan-

<sup>9</sup>Siehe beispielsweise [313].

tenfeldtheorie<sup>10</sup>. Für die nichtrelativistische Einteilchen-Quantenmechanik ist sie ziemlich unbrauchbar. Das soll keineswegs heißen, daß hier keine Resultate zu erzielen wären; für einen Großteil der quantenmechanischen Systeme, deren zugehörige Schrödingergleichungen lösbar sind, existieren auch Lösungen über Pfadintegrale<sup>11</sup>. Die herkömmlichen Wege der elementaren Hilbertraum-Quantenmechanik, algebraisch wie analytisch, sind für nichtrelativistische Systeme jedoch stets erheblich einfacher.

### 2.2.3 Feynman-Integrale und Standard-Quantenmechanik

Die haushohe technische Überlegenheit der Standard-Hilbertraum-Quantenmechanik gegenüber den Pfadintegralen bei nichtrelativistischen Einteilchen-Anwendungen legt zwar für konkrete Anwendungen den Einsatz ersterer nahe; das macht den alternativen Zugang natürlich aus philosophischer Sicht nicht weniger interessant. Vielmehr stellt sich dadurch automatisch die Frage nach der Relation zwischen der Pfadintegral- und der Hilbertraum-Formulierung. Daher ist es sehr lehrreich, sich von der Äquivalenz der beiden Zugänge zu überzeugen. Feynman führte das in seiner ersten Arbeit über den Gegenstand explizit vor [165]; wir beschreiben im folgenden einen etwas allgemeineren Fall.

Zum Nachweis der Äquivalenz der beiden Formulierungen betrachten wir eine typische quantenmechanische Situation: Ein System wird zunächst in einem gewissen Zustand präpariert, und später wird eine bestimmte Größe daran gemessen; die Präparation sei zum Zeitpunkt  $t_1$  abgeschlossen, die Messung beginne zum Zeitpunkt  $t_2$  und irgendwo dazwischen sei die Gegenwart  $t$ . Die Standardfrage ist dann diejenige nach der Wahrscheinlichkeit, das System nach der Messung in bestimmten Eigenzuständen der gemessenen Größe vorzufinden. In die Sprache der Pfadintegrale übersetzt entspricht das einer Aufteilung des Raum-Zeit-Gebietes  $\mathcal{G}$ , in dem sich das System aufhalten kann, in ein Gebiet  $\mathcal{G}_1$ , in dem sich das System zunächst für Zeitpunkte  $\tau < t_1$  befindet, ein Gebiet  $\mathcal{G}_2$ , in dem es sich für  $\tau > t_2$  befindet, und den Rest. Dabei sind  $\mathcal{G}_1$  und  $\mathcal{G}_2$  durch die Präparation beziehungsweise die Messung auf spezifische Weise festgelegt, während der Rest die gesamte Raum-Zeit zwischen  $t_1$  und  $t_2$  ist. Die oben erwähnte Wahrscheinlichkeit ist dann die bedingte Wahrscheinlichkeit dafür, das System für  $\tau > t_2$  in  $\mathcal{G}_2$  zu finden, wenn es für  $\tau < t_1$  in  $\mathcal{G}_1$  war.

In diesem Fall können wir den Propagator (2.13) mit  $t \equiv t_j$  und

$$S(x_{i+1}, x_i) = \int_{t_i}^{t_{i+1}} L(x(t), \dot{x}(t)) dt$$

in der Gestalt

$$K(x_2, t_2; x_1, t_1) = \lim_{\varepsilon \rightarrow 0} \cdots \iint \cdots \exp \left[ i \sum_i S(x_{i+1}, x_i) \right] \cdots \frac{dx_i}{A} \frac{dx_{i+1}}{A} \cdots$$

<sup>10</sup>Die Pionierarbeit hierzu stammt von Feynman selbst und stellt gleichzeitig einen der bahnbrechenden ersten Beiträge zur Quantenelektrodynamik dar [166]. Die Berechnung von Streuquerschnitten und die Quantisierung nichtabelscher Eichtheorien sind vielleicht die wichtigsten, aber längst nicht die einzigen Einsatzgebiete quantenfeldtheoretischer Pfadintegrale. Beispiele für Anwendungen in der Hochenergiephysik findet man unter anderem in [221], [269], [294] und [540].

<sup>11</sup>Einen Überblick über analytische Brerechnungsmethoden für Pfadintegrale findet man in [203] und in [313].

$$= \lim_{\varepsilon \rightarrow 0} \cdots \iint \cdots \exp \left[ i \sum_{i=j}^{\infty} S(x_{i+1}, x_i) \right] \exp \left[ i \sum_{i=-\infty}^{j-1} S(x_{i+1}, x_i) \right] \cdots \frac{dx_i}{A} \frac{dx_{i+1}}{A} \cdots$$

schreiben<sup>12</sup>. Nach Ausführen der Integrationen erhalten wir folglich einen Ausdruck der Form

$$K(\mathcal{G}_1, \mathcal{G}_2) = \int \varphi^*(x, t) \psi(x, t) dt,$$

mit

$$\psi(x_j, t) = \lim_{\varepsilon \rightarrow 0} \iint_{\mathcal{G}_1} \cdots \exp \left[ i \sum_{i=-\infty}^{j-1} S(x_{i+1}, x_i) \right] \frac{dx_{j-1}}{A} \frac{dx_{j-2}}{A} \cdots$$

und

$$\varphi(x_j, t) = \lim_{\varepsilon \rightarrow 0} \iint_{\mathcal{G}_2} \cdots \exp \left[ -i \sum_{i=j}^{\infty} S(x_{i+1}, x_i) \right] \frac{1}{A} \frac{dx_{j+1}}{A} \frac{dx_{j+2}}{A} \cdots$$

Die Bedeutung dieser beiden Größen liegt auf der Hand.  $\psi$  beschreibt vollständig den Zustand des Systems im Gebiet  $\mathcal{G}_1$  und ist damit die *Wellenfunktion* des Systems nach der Präparation,  $\varphi$  beschreibt dessen Zustand nach der Messung der erwähnten physikalischen Größe und ist folglich eine der Eigenfunktionen des diese Größe repräsentierenden selbstadjungierten Operators.  $K(\mathcal{G}_1, \mathcal{G}_2)$  ist die Übergangsamplitude vom Zustand  $\psi$  in den Zustand  $\varphi$  und entsprechend ist  $|K(\mathcal{G}_1, \mathcal{G}_2)|^2$  die Wahrscheinlichkeit, bei Vorliegen des Zustands  $\psi$  den zum Zustand  $\varphi$  gehörenden Eigenwert zu messen. Wir haben genau die Vorschrift der gewöhnlichen Quantenmechanik zur Berechnung solcher Wahrscheinlichkeiten wiedergefunden.

Die Interpretation von  $\psi$  als Wellenfunktion des betrachteten Systems ist natürlich nur gerechtfertigt, wenn  $\psi$  die Schrödingergleichung des Systems erfüllt. Um das zu sehen, lassen wir die Zeit in  $\psi$  um ein infinitesimales Stück  $\varepsilon$  weiterlaufen. Da das wegen  $t_j = t_{j-1} + \varepsilon$  gleichbedeutend mit der zusätzlichen Berücksichtigung des nächsten  $x_i$  in der Summe im Integranden von  $\psi$  ist, erhalten wir

$$\psi(x_j, t + \varepsilon) = \iiint_{\mathcal{G}_1} \cdots \exp \left[ i \sum_{i=-\infty}^j S(x_{i+1}, x_i) \right] \frac{dx_j}{A} \frac{dx_{j-1}}{A} \frac{dx_{j-2}}{A} \cdots$$

und nach Ausführen aller Integrationen von  $i = -\infty$  bis  $i = j - 1$  weiter

$$\psi(x_j, t + \varepsilon) = \int_{\mathbb{R}^{3N}} \exp [i S(x_{i+1}, x_i)] \psi(x_j, t) \frac{dx_j}{A}.$$

Das erinnert schon verdächtig an einen Zeitentwicklungsoperator und wird sich gleich als formal integrierte Schrödingergleichung erweisen.

Setzt man wieder die spezielle Form

$$\hat{H} = \sum_{i=1}^N \frac{\hat{P}_i^2}{2m_i} + V(x)$$

<sup>12</sup>Dieses Resultat ist natürlich darauf angewiesen, daß die Lagrang-Funktionen nur von den verallgemeinerten Koordinaten und Geschwindigkeiten, aber keinen höheren Ableitungen der Koordinaten abhängt.

des Hamilton-Operators voraus, so folgt

$$S(x_{i+1}, x_i) = \frac{\varepsilon}{2} \sum_{k=1}^N m_k \left( \frac{x_{k,i+1} - x_{k,i}}{\varepsilon} \right)^2 - \varepsilon V(x_i)$$

und damit

$$\psi(x_{j+1}, t + \varepsilon) = \int_{\mathbb{R}^{3N}} \exp \left\{ \frac{i\varepsilon}{2} \left[ \sum_{k=1}^N m_k \left( \frac{x_{k,j+1} - x_{k,j}}{\varepsilon} \right)^2 - V(x_{j+1}) \right] \right\} \psi(x_j, t) \frac{dx_j}{A}.$$

Die Substitution  $x_{j+1} - x_j = \xi$  zusammen mit der Abkürzung  $x_{j+1} = x$  macht daraus

$$\psi(x, t + \varepsilon) = \exp(-i\varepsilon V(x)) \int_{\mathbb{R}^{3N}} \exp \left( \frac{i}{2\varepsilon} \sum_{i=1}^N m_i \xi_i^2 \right) \psi(x - \xi, t) \frac{d\xi}{A}.$$

$\varepsilon$  ist eine infinitesimale Größe, folglich kann  $\psi(x - \xi, t)$  in eine Taylorreihe entwickelt werden; das liefert

$$\begin{aligned} \psi(x, t + \varepsilon) &= \exp(-i\varepsilon V(x)) \int_{\mathbb{R}^{3N}} \exp \left( \frac{i}{2\varepsilon} \sum_{i=1}^N m_i \xi_i^2 \right) \times \\ &\quad \times \left[ \psi(x, t) - \sum_{i=1}^N \xi_i \frac{\partial \psi(x, t)}{\partial x_i} + \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \xi_i \xi_j \frac{\partial^2 \psi(x, t)}{\partial x_k \partial x_l} \mp \dots \right] \frac{d\xi}{A}, \end{aligned}$$

und die Aufgabe besteht nun darin, die Integrale auf der rechten Seite der Reihe nach auszuwerten, wobei es genügt, bis zur dritten Ordnung in  $\xi$  zu gehen, denn die Potenzen ab  $\xi^4$  aufwärts liefern Beiträge, die alle zweiter Ordnung oder höher in  $\varepsilon$  sind und vernachlässigt werden können. Speziell geht es um die Ausdrücke

$$\begin{aligned} I_1 &= \int_{-\infty}^{\infty} \exp \left( \frac{im_i \xi_j^2}{2\varepsilon} \right) d\xi_j, & I_2 &= \int_{-\infty}^{\infty} \exp \left( \frac{im_i \xi_j^2}{2\varepsilon} \right) \xi_j d\xi_j \\ I_3 &= \int_{-\infty}^{\infty} \exp \left( \frac{im_i \xi_j^2}{2\varepsilon} \right) \xi_j^2 d\xi_j, & I_4 &= \int_{-\infty}^{\infty} \exp \left( \frac{im_i \xi_j^2}{2\varepsilon} \right) \xi_j^3 d\xi_j. \end{aligned}$$

Diese vier Integrale existieren allesamt nicht als Lebesgue-Integrale, sie können jedoch als uneigentliche Integrale berechnet werden.  $I_2$  und  $I_4$  verschwinden, weil die Integranden ungerade sind. Das erste haben wir im vorigen Abschnitt berechnet; mit dem dort gewonnenen Resultat erhalten wir hier

$$\int_{-\infty}^{\infty} \exp \left( \frac{im_i \xi_j^2}{2\varepsilon} \right) d\xi_j = \sqrt{\frac{2\pi i \varepsilon}{m_i}}.$$

Das dritte berechnen wir wie in Abschnitt 2.2.2; analog zur Vorgehensweise bei der Auswertung von (2.8) finden wir zunächst

$$\int_{-\infty}^{\infty} e^{it^2} t^2 dt = \sqrt{i\pi}$$

und mit  $t = im_i \xi_j^2 / 2\varepsilon$

$$\int_{-\infty}^{\infty} \exp\left(\frac{im_i \xi_j^2}{2\varepsilon}\right) \xi_j^2 d\xi_j = \frac{i\varepsilon}{m_i} \sqrt{\frac{2\pi i\varepsilon}{m_i}}.$$

Aus der Wellenfunktion wird damit

$$\psi(x, t + \varepsilon) = \exp(-i\varepsilon V(x)) \frac{(2i\varepsilon\pi)^{3N/2}}{A \prod_{i=1}^N m_i^{3/2}} \left[ \psi(x, t) - \frac{i\varepsilon}{2} \sum_{i=1}^N \frac{1}{m_i} \frac{\partial^2 \psi(x, t)}{\partial x_i^2} \pm \dots \right],$$

Entwickelt man nun auch  $\psi(x, t + \varepsilon)$  und die Exponentialfunktion auf der rechten Seite in Taylorreihen, erhält man

$$\begin{aligned} & \psi(x, t) + \varepsilon \frac{\partial \psi(x, t)}{\partial t} + \frac{\varepsilon^2}{2} \frac{\partial^2 \psi(x, t)}{\partial t^2} + \dots \\ &= \frac{(2i\varepsilon\pi)^{3N/2}}{A \prod_{i=1}^N m_i^{3/2}} \left[ 1 - i\varepsilon V(x) - \frac{1}{2} (-i\varepsilon V(x))^2 \pm \dots \right] \times \\ & \quad \times \left[ \psi(x, t) - \frac{i\varepsilon}{2} \sum_{i=1}^N \frac{1}{m_i} \frac{\partial^2 \psi(x, t)}{\partial x_i^2} + \dots \right] \end{aligned}$$

und nach ausmultiplizieren bis zur ersten Ordnung in  $\varepsilon$

$$\begin{aligned} & \psi(x, t) + \varepsilon \frac{\partial \psi(x, t)}{\partial t} + \dots \\ &= \frac{(2i\varepsilon\pi)^{3N/2}}{A \prod_{i=1}^N m_i^{3/2}} \left\{ \psi(x, t) - \varepsilon \left[ \frac{i}{2} \sum_{i=1}^N \frac{1}{m_i} \frac{\partial^2 \psi(x, t)}{\partial x_i^2} + i V(x) \psi(x, t) \right] \pm \dots \right\} \end{aligned}$$

Vergleich der nullten Ordnungen in  $\varepsilon$  liefert zunächst für den Normierungsfaktor in Übereinstimmung mit (2.12)

$$A = \frac{(2i\varepsilon\pi)^{3N/2}}{\prod_{i=1}^N m_i^{3/2}},$$

und Vergleich der ersten Ordnungen in  $\varepsilon$  ergibt

$$i \frac{\partial \psi(x, t)}{\partial t} = \frac{1}{2} \sum_{i=1}^N \frac{1}{m_i} \frac{\partial^2 \psi(x, t)}{\partial x_i^2} + V(x) \psi(x, t),$$

also genau die Schrödingergleichung für die Wellenfunktion  $\psi(x, t)$ . Führt man eine analoge Prozedur für  $\varphi^*$  durch, erhält man dieselbe Gleichung, allerdings mit dem Unterschied, daß die Zeit gerade das umgekehrte Vorzeichen erhält. Folglich erfüllt die komplex konjugierte Funktion  $\varphi$  wieder genau die Schrödingergleichung, sodaß deren oben erwähnte Interpretation als spezielle Wellenfunktion, nämlich als einer Eigenfunktion der gemessenen Observablen, gerechtfertigt ist. Damit ist die Äquivalenz des Pfadintegralformalismus zur Standardformulierung für die wesentlichen Fälle der nichtrelativistischen Quantenmechanik nachgewiesen – zumindest auf derjenigen heuristischen Ebene, auf welcher sich der gesamte Pfadintegralformalismus, wie er bis hierher diskutiert wurde, auch selbst bewegt. Wir werden weiter unten sehen, wie auch dieser Äquivalenzbeweis in eine mathematisch exakte Form gebracht werden kann.

#### 2.2.4 Der klassische Limes und die Methode der stationären Phase

Da Pfadintegrale wesentlich Gebrauch von klassischen Begriffen machen, liefern sie als interessantes Nebenprodukt eine sehr direkte Veranschaulichung des Übergangs von der Quantenmechanik zur klassischen Mechanik im Sinne des Grenzübergangs  $\hbar \rightarrow 0$ ; einer heuristischen Darstellung dieses Sachverhalts widmet sich der vorliegende Abschnitt; sie geht im wesentlichen auf Feynman selbst zurück [165].

Wie gesehen erfolgt die Berechnung von Übergangsamplituden im Pfadintegralformalismus durch die Summation über alle möglichen Wege vom Start- zum Zielzustand. Man kann sich nun leicht davon überzeugen, daß sich im klassischen Grenzfall die allermeisten dieser Wege durch destruktive Interferenz herausheben und gleichzeitig durch konstruktive Interferenz nur der klassische Weg übrigbleibt. Entscheidend ist dabei, daß sich die Wege jeweils in einer Phase proportional zur Wirkung unterscheiden. Abgesehen davon sind alle Wege einschließlich desjenigen Weges  $x_{\text{kl}}$ , der klassisch herauskommen würde, gleichberechtigt. Das besondere an diesem klassischen Weg ist jedoch, daß er gemäß dem Hamiltonschen Prinzip die klassische Wirkung  $S$  extremalisiert.  $S$  ist folglich bei  $x_{\text{kl}}$  stationär, das heißt, die Wirkung ändert sich bei kleinen Variationen des Weges in der Umgebung von  $x_{\text{kl}}$  kaum, bei kleinen Änderungen in der Umgebung eines beliebigen anderen Weges  $x$  jedoch im Allgemeinen sehr stark. Sind  $x_1$  und  $x_2$  zwei solche Wege in der Umgebung von  $x$ , so gilt folglich  $S(x_1) \neq S(x_2)$ , und aufgrund der Kleinheit der Planckschen Konstanten liegen die Zahlenwerte der beiden Phasen  $S(x_1)/\hbar$  und  $S(x_2)/\hbar$  extrem weit auseinander. Die Phasenfaktoren  $\exp[iS(x_1)/\hbar]$  und  $\exp[iS(x_2)/\hbar]$  entsprechen dadurch auf der komplexen Einheitskugel Punkten, die ebenfalls weit auseinander liegen können. Sind  $x_3$  und  $x_4$  dagegen zwei Wege in der Umgebung von  $x_{\text{kl}}$ , so liegen sowohl die Phasen  $S(x_3)/\hbar$  und  $S(x_4)/\hbar$  als auch die den Phasenfaktoren  $\exp[iS(x_1)/\hbar]$  und  $\exp[iS(x_2)/\hbar]$  zugeordneten Punkte auf der komplexen Einheitskugel nahe beieinander. Läßt man  $\hbar$  formal noch kleiner werden, gilt das alles in gesteigertem Maß. Im Grenzfall  $\hbar \rightarrow 0$  führt das dazu, daß sich in der Umgebung eines beliebigen Weges  $x$



bei der Aufsummierung aller Wege stets zwei Wege finden lassen, die sich aufgrund ihrer Phasendifferenz gerade gegenseitig wegheben, während sich die Beiträge der Wege in der Umgebung von  $x_{kl}$  gerade zu  $x_{kl}$  verstärken. Man erhält in diesem Grenzfall also genau das klassische Verhalten des Systems. Das hier angedeutete Verfahren nennt man aus naheliegenden Gründen die *Methode der stationären Phase*.

Es sei betont, daß dieses Argument rein heuristischer Natur ist und erheblich der Präzisierung bedarf; dies hat das Verfahren natürlich mit dem Feynmanschen Pfadintegralformalismus gemein. Spezielle Varianten der mathematisch rigorosen Rekonstruktionen von letzterem, auf die wir gleich ausführlicher zurückkommen, gestatten in der Tat genau eine solche Präzisierung und damit eine formal strenge Beschreibung des Übergangs von der Quantenmechanik zur klassischen Mechanik.

Wir weisen außerdem darauf hin, daß eine solche Darstellung des klassischen Grenzfalls der Quantenmechanik, ganz egal ob heuristisch oder mathematisch exakt, in keiner Weise erklären kann, *warum* wir unsere Umwelt im allgemeinen klassisch wahrnehmen und quantenmechanische Effekte erst mit aufwendigen experimentellen Mitteln nachweisbar sind. Das ist deshalb so erstaunlich, weil die Welt nach allem, was wir wissen, vollständig aus quantenmechanischen Objekten aufgebaut ist und die Quantenmechanik beziehungsweise ihre Verallgemeinerung in Gestalt der Quantenfeldtheorie als die fundamentale Theorie der Natur zu betrachten ist. Der Lösung dieses sogenannten *Lokalitätsproblems* ist man inzwischen mit der Theorie der *Dekohärenz* womöglich ein erhebliches Stück näher gekommen, wenngleich das ganze nach wie vor als Forschungsprojekt und keinesfalls als abgeschlossene Angelegenheit aufzufassen ist. Kurz gesagt erklärt die Dekohärenz die Beobachtung vom klassischen Verhalten physikalischer Systeme durch die (quantenmechanische!) Wechselwirkung derselben mit ihrer Umgebung, wodurch die sehr wohl vorhandenen nichtklassischen Superpositionen aufgrund der riesigen Anzahl von Freiheitsgraden des Gesamtsystems unbeobachtbar werden. Das Auftreten einer klassischen Welt ist danach selbst ein rein quantenmechanisches Phänomen. Einen Überblick über dieses aktuelle Gebiet der Quantenmechanik liefern [287] und [449].

### 2.2.5 Pfadintegrale im Phasenraum

Bis hierher haben wir uns ausschließlich im Konfigurationsraum des betrachteten Systems aufgehalten. Das verallgemeinern wir jetzt auf den Phasenraum des Systems, also auf den von den  $3N$  Koordinaten und  $3N$  Impulsen aufgespannten  $6N$ -dimensionalen Raum. Klassisch ist jeder Zeitentwicklung des Systems mit Randwerten  $(x_a, p_a)$  und  $(x_b, p_b)$  eine Phasenraum-Trajektorie von  $(x_a, p_a)$  nach  $(x_b, p_b)$  zugeordnet, also ein Element des Raums  $\mathcal{C}[(x_a, p_a), (x_b, p_b)] \equiv \mathcal{C}(a, b)$  der stetigen Kurven von  $(x_a, p_a)$  nach  $(x_b, p_b)$  im Phasenraum.

Wir greifen dazu die Gleichungen (2.5), (2.6) und (2.7) aus Abschnitt 2.2.2 wieder auf; bauen wir sie zusammen, erhalten wir

$$K(x_b, t_b; x_a, t_a) = \lim_{n \rightarrow \infty} \int_{\mathbb{R}^{3N}} \int_{\mathbb{R}^{3N}} \cdots \int_{\mathbb{R}^{3N}} \exp \left\{ i \sum_{j=0}^{n-1} \left[ \sum_{k=1}^N p_{j,k} (x_{j+1,k} - x_{j,k}) - H(p_j, x_j) \Delta t \right] \right\} \prod_{k=1}^{n-1} dx_k \prod_{l=1}^n \frac{dp_l}{2\pi}. \quad (2.16)$$

Schreiben wir für die Summanden in (2.16)

$$\sum_{k=1}^N p_{j,k} (x_{j+1,k} - x_{j,k}) - H(p_j, x_j) \Delta t = \left[ \sum_{k=1}^N p_{j,k} \frac{x_{j+1,k} - x_{j,k}}{\Delta t} - H(p_j, x_j) \right] \Delta t,$$

erhalten wir im Grenzwert  $n \rightarrow \infty$

$$K(x_b, t_b; x_a, t_a) = \lim_{n \rightarrow \infty} \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} \cdots \int_{\mathbb{R}^N} \exp \left\{ i \int_{t_a}^{t_b} \left[ \sum_{k=1}^N p_{j,k} \dot{x}_{j,k} - H(p_j, x_j) \right] dt \right\} \times \\ \times \prod_{k=1}^{n-1} dx_k \prod_{l=1}^n \frac{dp_l}{2\pi}.$$

Das Integral im Exponenten ist wieder die klassische Wirkung, und damit nimmt der Propagator die Form

$$K(x_b, t_b; x_a, t_a) = \lim_{n \rightarrow \infty} \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} \cdots \int_{\mathbb{R}^N} e^{iS_{\text{kl}}[p_i, x_i]} \prod_{k=1}^{n-1} dx_k \prod_{l=1}^n \frac{dp_l}{2\pi} \quad (2.17)$$

an. Man schreibt dafür wieder symbolisch

$$K(x_b, t_b; x_a, t_a) = \iint_{\mathcal{C}(a,b)} e^{iS_{\text{kl}}[p_i, x_i]} \mathcal{D}x \mathcal{D}p.$$

Hierbei handelt es sich um eine Integration über alle stetigen Wege von  $(x_a, p_a)$  nach  $(x_b, p_b)$  im Phasenraum. In der bisherigen, spezielleren Formulierung werden wegen der Integration über alle möglichen Raum-Zeit-Wege auch alle möglichen Steigungen solcher Wege berücksichtigt. Dies sind gerade alle möglichen verschiedenen  $p$ 's. In der allgemeineren Formulierung werden die  $p$ 's bei der Integration über alle möglichen Wege im Phasenraum explizit als Koordinaten mitberücksichtigt. Für die Zeitentwicklung von Wellenfunktionen erhalten wir entsprechend

$$\psi(q_2, t_2) = \int_{\mathbb{R}^{3N}} \iint_{\mathcal{C}(a,b)} e^{iS_{\text{kl}}[p_i, q_i]} \psi(q_1, t_1) \mathcal{D}q \mathcal{D}p d^{3N} q_1. \quad (2.18)$$

Die Resultate (2.15) und (2.18) wurden unter der Voraussetzung des Vorliegens eines sehr speziellen Hamilton-Operators hergeleitet. Das können wir verallgemeinern; wir betrachten dazu ein System, das durch verallgemeinerte Koordinaten  $q_1, q_2, \dots, q_N$  und  $p_1, p_2, \dots, p_N$  dazu kanonisch konjugierten verallgemeinerte Impulse sowie der allgemeinen Lagrange-Funktion

$$L(p, q) = \sum_{i=1}^N p_i \dot{q}_i - H(p_1, p_2, \dots, p_N, q_1, q_2, \dots, q_N) \equiv \sum_{i=1}^N p_i \dot{q}_i - H(p_i, q_i, t)$$

beschrieben werden soll; diese führt zum Wirkungsfunktional

$$S_{\text{kl}}[p_i, q_i] := \int_{t_a}^{t_b} \left[ \sum_{i=1}^N p_i(t) \dot{q}_i(t) - H(p_i(t), q_i(t), t) \right] dt.$$

Die Hamilton-Funktion ist hier keinerlei Beschränkungen unterworfen. Aus (2.18) wird so das allgemeine Phasenraum-Pfadintegral

$$\psi(q_2, t_2) = \int_{\mathbb{R}^{3N}} \iint_{\mathcal{C}(a,b)} e^{iS_{\text{kl}}[p_i, q_i]} \psi(q_1, t_1) \mathcal{D}q \mathcal{D}p d^{3N} q_1.$$

Der Vorteil der Formulierung im Phasenraum liegt in deren fundamentalen Natur: Während die Formulierbarkeit von Pfadintegralen im Konfigurationsraum vom Hamilton-Operator abhängt und nicht immer möglich ist, können Pfadintegrale im Phasenraum auf dem hier beschriebenen informellen Niveau in völliger Allgemeinheit für beliebige Hamilton-Operatoren aufgestellt werden<sup>13</sup>.

Damit beschließen wir die etwas hemdsärmelige Vorbetrachtung, die in erster Linie den Sinn hatte, eine möglichst leicht zugängliche Vorstellung der Pfadintegrale zu liefern und gleichwohl in der Quantenmechanik und insbesondere der Quantenfeldtheorie weite Anwendungen erfährt. Um zu sehen, daß sich daraus zumindest für die Quantenmechanik *keine* formal korrekte Betrachtung machen läßt, ist zunächst eine Präzisierung der vereinfachten Darstellung erforderlich, da ansonsten überhaupt nicht wirklich klar ist, wovon eigentlich die Rede ist; diese Präzisierung sowie die anschließenden mathematisch strengen Umformulierungen werden uns nun weit in die Funktionalanalysis und Stochastik führen.

## 2.3 Halbgruppen, Produktformeln und der Satz von Cameron

Bevor wir uns nun also daranmachen, anstelle der bisherigen heuristischen Herleitungen mit viel Aufwand mathematisch exakte Definitionen für Pfadintegrale anzugeben, werden wir das wie angekündigt ausführlich begründen. Es wurde bereits angedeutet, daß die Feynmansche Formulierung der Quantenmechanik exakte Definitionen ihrer zentralen Begriffe nicht vorweisen kann. Das liegt an einigen mathematischen Unzulänglichkeiten, die dazu führen, daß das Feynmansche Integral in seiner üblichen Form mathematisch gar nicht existiert. Um das zu sehen, müssen wir jedoch zuerst einmal die heuristische Version der Pfadintegrale mathematisch präziser formulieren, da sonst keine Existenzaussagen möglich sind; das wird Gegenstand des folgenden Teilabschnittes sein. Anschließend überzeugen wir uns davon, daß diese vermeintlich exakte Formulierung mathematisch nicht durchführbar ist. Die Ersetzung der heuristischen durch mathematisch existierende und damit überhaupt erst definierbare Versionen ist ein anspruchsvolles Stück mathematische Physik, wie wir im nächsten Abschnitt sehen werden. Der vermeintlich so schön anschauliche Zugang erweist sich so plötzlich als überaus kompliziert.

Nicht in Abrede gestellt werden soll damit natürlich die außergewöhnliche praktische Bedeutung, welche dieses heuristische Feynman-Integral in denjenigen Einsatzgebieten hat, wo es auch anwendbar ist, also insbesondere in der statistischen Mechanik, der Quantenfeldtheorie und in Versuchen, vereinheitlichte Wechselwirkungs- und Quantengravitationstheorien aufzustellen. Das beseitigt jedoch nicht die Notwendigkeit einer sauberen mathematischen Formulierung, die zwar zweifellos nicht bei jeder Anwendung zum

<sup>13</sup>Details findet man beispielsweise in [308] und [369].

Einsatz kommen, so aber doch gewissermaßen im Hintergrund als Absicherung vorhanden sein muß. Umso bedeutsamer ist dann auch die Tatsache, daß solche mathematischen Präzisierungen tatsächlich möglich sind. Die Situation ist vergleichbar mit derjenigen der Distributionentheorie; Distributionen werden an allen Ecken und Enden in der Quantenmechanik physikalisch robust und mathematisch bedenkenlos verwendet, was sich jedoch rechtfertigen läßt, da es eine präzise mathematische Theorie der Distributionen gibt.

## 2.3.1 Integraldarstellung des Zeitentwicklungsoperators

### 2.3.1.1 Die Trotter-Produktformel

Wir beginnen mit einem Hilfssatz, der sich nicht nur in diesem, sondern auch im folgenden Abschnitt als nützlich erweisen wird und der Tatsache geschuldet ist, daß operatorwertige Exponentialausdrücke wie etwa  $\exp(\hat{A} + \hat{B})$  oder  $\exp(\hat{A}\hat{B})$  nur dann unproblematisch sind, wenn die Operatoren im Exponenten kommutieren; tun sie dies nicht, wird die Sache kompliziert. Das erkennt man schnell, wenn man versucht, so etwas auszurechnen und zum Beispiel bei der *Baker-Campbell-Hausdorff-Formel*

$$e^{\hat{A}} e^{\hat{B}} = \exp \left\{ \sum_{n=1}^{\infty} \sum_{\substack{p_j+q_j>0 \\ 1 \leq j \leq n}} \frac{(-1)^{n-1}}{n \prod_{j=1}^n p_j! q_j! \sum_{j=1}^n (p_j + q_j)} \times \right. \\ \left. \times \underbrace{[\hat{A}, [\hat{A}, \dots [\hat{A}, \hat{B}, [\hat{B}, \dots [\hat{B}, \dots [\hat{A}, [\hat{A}, \dots [\hat{A}, [\hat{B}, [\hat{B}, \dots [\hat{B}, \dots]]]]]]]]]}_{p_1 \text{ mal}} \underbrace{]}_{q_1 \text{ mal}} \dots \underbrace{]}_{p_n \text{ mal}} \underbrace{]}_{q_n \text{ mal}} \dots \right\}$$

landet; ausgeschrieben lautet das

$$e^{\hat{A}} e^{\hat{B}} = \exp \left\{ \hat{A} + \hat{B} + \frac{1}{2} [\hat{A}, \hat{B}] + \frac{1}{12} \left( [\hat{A}, [\hat{A}, \hat{B}]] - [\hat{B}, [\hat{A}, \hat{B}]] \right) - \frac{1}{24} [\hat{B}, [\hat{A}, [\hat{B}, \hat{A}]]] \right. \\ + \frac{1}{120} \left( [\hat{B}, [\hat{A}, [\hat{B}, [\hat{A}, \hat{B}]]]] - [\hat{A}, [\hat{A}, [\hat{B}, [\hat{A}, \hat{B}]]]] \right) \\ + \frac{1}{360} \left( [\hat{B}, [\hat{A}, [\hat{A}, [\hat{A}, \hat{A}]]]] - [\hat{A}, [\hat{B}, [\hat{B}, [\hat{A}, \hat{B}]]]] \right) \\ + \frac{1}{720} \left( [\hat{B}, [\hat{B}, [\hat{B}, [\hat{A}, \hat{B}]]]] - [\hat{A}, [\hat{A}, [\hat{A}, [\hat{A}, \hat{B}]]]] \right) \\ + \frac{1}{160} [\hat{A}, [\hat{B}, [\hat{A}, [\hat{B}, [\hat{A}, \hat{B}]]]]] - \frac{1}{180} [\hat{B}, [\hat{A}, [\hat{B}, [\hat{B}, [\hat{A}, \hat{B}]]]]] \\ + \frac{1}{288} [\hat{A}, [\hat{B}, [\hat{B}, [\hat{B}, [\hat{A}, \hat{B}]]]]] - \frac{1}{360} [\hat{A}, [\hat{B}, [\hat{A}, [\hat{A}, [\hat{A}, \hat{B}]]]]] \\ + \frac{1}{360} [\hat{B}, [\hat{B}, [\hat{A}, [\hat{B}, [\hat{A}, \hat{B}]]]]] + \frac{1}{480} [\hat{B}, [\hat{A}, [\hat{A}, [\hat{A}, [\hat{A}, \hat{B}]]]]] \\ - \frac{1}{480} [\hat{A}, [\hat{A}, [\hat{B}, [\hat{B}, [\hat{A}, \hat{B}]]]]] - \frac{1}{480} [\hat{B}, [\hat{A}, [\hat{A}, [\hat{B}, [\hat{A}, \hat{B}]]]]] \\ \left. + \frac{1}{720} [\hat{A}, [\hat{A}, [\hat{A}, [\hat{B}, [\hat{A}, \hat{B}]]]]] + \frac{1}{1440} [\hat{B}, [\hat{B}, [\hat{A}, [\hat{A}, [\hat{A}, \hat{B}]]]]] + \dots \right\}.$$

Auch der dazu duale Fall ist nicht einfacher, er führt auf die *Zassenhaus-Formel*

$$\begin{aligned}
e^{\hat{A}+\hat{B}} &= \exp(\hat{A}) \exp(\hat{B}) \exp\left(\frac{1}{2}[\hat{B}, \hat{A}]\right) \exp\left(\frac{1}{3}[[\hat{B}, \hat{A}], \hat{B}] + \frac{1}{6}[[\hat{B}, \hat{A}], \hat{A}]\right) \times \\
&\times \exp\left\{\frac{1}{8}\left([[[\hat{B}, \hat{A}], \hat{B}], \hat{B}] + [[[\hat{B}, \hat{A}], \hat{A}], \hat{B}]\right) + \frac{1}{24}[[[\hat{B}, \hat{A}], \hat{A}], \hat{A}]\right\} \times \\
&\times \exp\left\{\frac{1}{30}\left([[[[\hat{B}, \hat{A}], \hat{B}], \hat{B}], \hat{B}] + [[[[\hat{B}, \hat{A}], \hat{A}], \hat{A}], \hat{B}]\right) \right. \\
&\quad + \frac{1}{20}\left([[[[\hat{B}, \hat{A}], \hat{A}], \hat{B}], \hat{B}] + [[[\hat{B}, \hat{A}], \hat{A}], [\hat{B}, \hat{A}]]\right) \\
&\quad \left. + \frac{1}{10}[[[\hat{B}, \hat{A}], \hat{B}], [\hat{B}, \hat{A}]] + \frac{1}{120}[[[[\hat{B}, \hat{A}], \hat{A}], \hat{A}], \hat{A}]\right\} \\
&\times \exp\left\{\frac{1}{144}\left([[[[[\hat{B}, \hat{A}], \hat{B}], \hat{B}], \hat{B}], \hat{B}] + [[[[[\hat{B}, \hat{A}], \hat{A}], \hat{A}], \hat{A}], \hat{B}]\right) \right. \\
&\quad + \frac{1}{72}\left([[[[[\hat{B}, \hat{A}], \hat{A}], \hat{B}], \hat{B}], \hat{B}] + [[[[[\hat{B}, \hat{A}], \hat{A}], \hat{A}], \hat{B}], \hat{B}] \right. \\
&\quad \left. + [[[[[\hat{B}, \hat{A}], \hat{A}], \hat{A}], [\hat{B}, \hat{A}]]]\right) \\
&\quad + \frac{1}{24}\left([[[[\hat{B}, \hat{A}], \hat{B}], \hat{B}], [\hat{B}, \hat{A}]] + [[[[\hat{B}, \hat{A}], \hat{A}], \hat{B}], [\hat{B}, \hat{A}]]\right) \\
&\quad \left. + \frac{1}{720}[[[[[\hat{B}, \hat{A}], \hat{A}], \hat{A}], \hat{A}], \hat{A}]\right\} + \dots
\end{aligned}$$

Es liegt auf der Hand, daß das sehr unhandlich wird, sobald man mehr als die ersten paar Terme mitzunehmen gedenkt<sup>14</sup>.

Wenn die Operatoren  $\hat{A}$  und  $\hat{B}$  gewisse Zusatzbedingungen erfüllen, gibt es für den hier interessierenden Fall einer Operatorsumme im Exponenten etwas besseres, nämlich die sogenannte *Trotter-Produktformel*. Sie wurde von H. F. Trotter in sehr allgemeiner Form für Halbgruppen auf Banachräumen bewiesen [508], [509], die Originalfassung des Beweises ist allerdings etwas aufwendig und recht technisch. Greift man jedoch zu Hilfsmitteln, die später dazukamen, wird die Sache sehr viel überschaubarer; P. R. Chernoff fand eine verallgemeinerte Produktformel [96], [97], aus welcher man diejenige von Trotter schnell ableiten kann [144]. Wir werden das im folgenden detailliert nachvollziehen.

Wesentlich für die Handhabbarkeit operatorwertiger Exponentialausdrücke ist die Betrachtung von Scharen solcher Objekte anstelle von nur einzelnen. Das geschieht durch Hinzufügen eines kontinuierlichen Parameters im Exponenten, wodurch die Exponentialausdrücke zwar nicht generell Gruppen-, zumindest aber Halbgruppeneigenschaften erlangen. Daher erinnern wir zunächst an einige grundlegende Begriffe aus der Theorie der Halbgruppen<sup>15</sup>.

<sup>14</sup>Details sowie Herleitungen zu diesen Formeln finden sich beispielsweise in [134] und [528].

<sup>15</sup>Der klassische Text über Halbgruppen ist [247]; er stammt allerdings aus der Zeit vor der Entdeckung der hier interessierenden Resultate. Einen aktuellen Überblick liefern [144] und [145].

**Definition:** Eine Familie  $\{\hat{T}(t)\}_{t>0}$  beschränkter Operatoren auf einem Banachraum  $\mathcal{B}$  heißt *Halbgruppe* auf  $\mathcal{B}$ , wenn  $\hat{T}(t+s) = \hat{T}(t)\hat{T}(s)$  gilt für alle  $s, t \in \mathbb{R}_0^+$ ; sie heißt *stark stetige Halbgruppe* oder kurz  $(C_0)$ -Halbgruppe, wenn  $\lim_{t \rightarrow 0} \hat{T}(t)f = f$  gilt für alle  $f \in \mathcal{B}$ .

Der Operator

$$\hat{\Omega} = \lim_{t \rightarrow 0} \left\{ \frac{1}{t} [\hat{T}(t) - \mathbb{1}] \right\}$$

heißt *infinitesimaler Generator* oder kurz Generator von  $\hat{T}(t)$ .

Wegen  $\hat{T}(0) = \mathbb{1}$  gilt für einen solchen Generator

$$\lim_{t \rightarrow 0} \frac{\hat{T}(t)f - \hat{T}(0)f}{t} = \hat{\Omega}f$$

oder

$$\frac{d}{dt} \hat{T}(t)f = \hat{\Omega}f$$

beziehungsweise, in ausintegrierter Form,

$$\hat{T}(t)f = f + \int_0^t \hat{T}(s)\hat{\Omega}f ds,$$

weswegen man symbolisch  $\hat{T}(t) = e^{\Omega t}$  schreibt<sup>16</sup>. Wir werden die schlampige, aber verbreitete Konvention verwenden, mit  $\hat{T}(t)$  je nach Kontext ein einzelnes Element oder die gesamte Halbgruppe zu meinen.

Für den Beweis der Trotter-Produktformel betrachten wir nun zunächst die bereits erwähnte Verallgemeinerung derselben.

**Satz:** (*Chernoff-Produktformel*)  $\{\hat{F}(t)\}_{t \geq 0}$  sei eine Familie kontrahierender linearer Abbildungen auf dem Banachraum  $\mathcal{B}$  mit  $\hat{F}(0) = \mathbb{1}$  und  $\|[\hat{F}(t)]^m\| \leq M$  für alle  $t \geq 0$  und  $m \in \mathbb{N}$  und einer Konstanten  $M \geq 1$ . Außerdem sei  $\mathcal{D}$  ein dichter Teilraum von  $\mathcal{B}$ , so daß

$$\hat{A}f := \lim_{h \searrow 0} \frac{\hat{F}(h)f - f}{h}$$

für alle  $f \in \mathcal{D}$  existiert und  $(\lambda_0 - \hat{A})\mathcal{D}$  ebenfalls ein dichter Teilraum von  $\mathcal{B}$  ist. Dann ist der Abschluß  $\hat{\bar{A}}$  von  $\hat{A}$  Generator einer beschränkten  $(C_0)$ -Halbgruppe  $\{\hat{T}(t)\}_{t \geq 0}$ , für die

$$\hat{T}(t) = \lim_{n \rightarrow \infty} [\hat{F}(t/n)]^n \quad (2.19)$$

gilt. Die Konvergenz ist gleichmäßig auf kompakten Teilmengen von  $[0, \infty)$ .

<sup>16</sup>Das ist im hier diskutierten allgemeinen Fall tatsächlich nur eine symbolische Schreibweise, die man nicht zu wörtlich nehmen darf, da sie zu Fehlschlüssen führen kann, weil Halbgruppen nicht alle Eigenschaften der Exponentialfunktion aufweisen. Eine wirkliche mathematische Relation wird daraus erst, wenn die Operatoren im Exponenten selbstadjungiert sind.

BEWEIS: Für  $\varepsilon > 0$  betrachte man den beschränkten Operator

$$\hat{A}_\varepsilon := \frac{\hat{F}(\varepsilon) - \mathbf{1}}{\varepsilon};$$

dafür gilt

$$\lim_{\varepsilon \rightarrow 0} \hat{A}_\varepsilon f = \hat{A} f$$

für alle  $f \in \mathcal{D}$ . Die  $\hat{A}_\varepsilon$  generieren Halbgruppen  $\{e^{t\hat{A}_\varepsilon}\}_{t \geq 0}$ , die nach Voraussetzung jeweils für alle  $t \geq 0$  die Abschätzung

$$\|e^{t\hat{A}_\varepsilon}\| \leq e^{-t/\varepsilon} \|e^{t\hat{F}(\varepsilon)/\varepsilon}\| \leq e^{-t/\varepsilon} \sum_{n=0}^{\infty} \frac{t^n \|[\hat{F}(\varepsilon)]^n\|}{\varepsilon^n n!} \leq M$$

erfüllen. Nach dem Trotter-Kato-Theorem [508], [300] ist damit der Abschluß  $\overline{\hat{A}}$  von  $\hat{A}$  Generator einer  $(C_0)$ -Halbgruppe  $\{\hat{T}(t)\}_{t \geq 0}$  mit

$$\lim_{\varepsilon \rightarrow 0} \|\hat{T}(t) f - e^{t\hat{A}_\varepsilon} f\| = 0,$$

und damit auch

$$\lim_{j \rightarrow \infty} \|\hat{T}(t) f - e^{j\hat{A}_{t/j}} f\| = 0,$$

jeweils für alle  $f \in \mathcal{B}$  und gleichmäßig in  $[0, t_0]$  für beliebiges  $t_0 > 0$ , wodurch eine Abschätzung für die linke Seite von (2.19) vorliegt. Um nun auch eine entsprechende Abschätzung für die rechte Seite zu erhalten, schreibt man zunächst für festes  $j$

$$\begin{aligned} e^{j[\hat{F}(t/j) - \mathbf{1}]} - [\hat{F}(t/j)]^j &= e^{-j} \{e^{j\hat{F}(t/j)} - e^j [\hat{F}(t/j)]^j\} \\ &= e^{-j} \sum_{k=0}^{\infty} \frac{j^k}{k!} \{[\hat{F}(t/j)]^k - [\hat{F}(t/j)]^j\}. \end{aligned}$$

Die zweite geschweifte Klammer läßt sich für  $k > j$  als

$$[\hat{F}(t/j)]^k - [\hat{F}(t/j)]^j = \sum_{l=j}^{k-1} \{[\hat{F}(t/j)]^{l+1} - [\hat{F}(t/j)]^l\} = \sum_{l=j}^{k-1} [\hat{F}(t/j)]^l \{[\hat{F}(t/j)] - \mathbf{1}\}$$

und für  $k < j$  als

$$[\hat{F}(t/j)]^k - [\hat{F}(t/j)]^j = \sum_{l=k}^{j-1} \{[\hat{F}(t/j)]^{l+1} - [\hat{F}(t/j)]^l\} = \sum_{l=k}^{j-1} [\hat{F}(t/j)]^l \{[\hat{F}(t/j)] - \mathbf{1}\}$$

schreiben. Zusammen mit  $\|[\hat{F}(t)]^m\| \leq M$  ergibt das für alle  $k \in \mathbb{N}$  und alle  $f \in \mathcal{B}$

$$\left\| [\hat{F}(t/j)]^k f - [\hat{F}(t/j)]^j f \right\| \leq |j - k| M \|\hat{F}(t/j) f - f\|.$$

Daraus wiederum folgt mit der Cauchy-Schwarzschen Ungleichung

$$\left\| e^{j\hat{A}_{t/j}} f - [\hat{F}(t/j)]^j f \right\| \leq e^{-j} \leq M \|\hat{F}(t/j) f - f\| \sum_{k=0}^{\infty} \frac{j^k}{k!} |j - k|$$



$$\begin{aligned}
&\leq e^{-j} \leq M \|\hat{F}(t/j) f - f\| \left( \sum_{k=0}^{\infty} \frac{j^k}{k!} \right)^{1/2} \left( \sum_{k=0}^{\infty} \frac{j^k}{k!} (j-k)^2 \right)^{1/2} \\
&= e^{-j} M \|\hat{F}(t/j) f - f\| e^{j/2} j^{1/2} e^{j/2} \\
&= \sqrt{j} M \|\hat{F}(t/j) f - f\| = \frac{tM}{\sqrt{j}} \|\hat{A}_{t/j} f\|,
\end{aligned}$$

also für alle  $f \in \mathcal{D}$

$$\lim_{j \rightarrow \infty} \left\| e^{j\hat{A}_{t/j}} f - [\hat{F}(t/j)]^j f \right\| = 0$$

gleichmäßig in  $]0, t_0]$ . Nach Voraussetzung gilt außerdem

$$\left\| e^{t\hat{A}_{t/j}} - [\hat{F}(t/j)]^j \right\| \leq 2M,$$

damit insgesamt

$$\left\| \hat{T}(t) f - e^{t\hat{A}_{t/j}} f \right\| - \left\| e^{j\hat{A}_{t/j}} f - [\hat{F}(t/j)]^j f \right\| \geq \left\| \hat{T}(t) f - [\hat{F}(t/j)]^j f \right\|$$

und weiter

$$\lim_{j \rightarrow \infty} \left\| \hat{T}(t) f - [\hat{F}(t/j)]^j f \right\| = 0.$$

Daraus folgt (2.19). □

Im folgenden sei  $\mathcal{D}(\hat{A})$  stets die Definitionsmenge des Operators  $\hat{A}$ . Eine Konsequenz aus der Chernoff-Produktformel ist nun das für den vorliegenden Abschnitt namensgebende Resultat.

**Satz:** (Trotter-Produktformel)  $\{\hat{T}(t)\}_{t>0}$  und  $\{\hat{T}'(t)\}_{t>0}$  seien  $(C_0)$ -Halbgruppen auf einem Banachraum  $\mathcal{B}$ , welche für alle  $t > 0$  die Bedingungen

$$\|\hat{T}(t)\| \leq e^{\omega t}, \quad \|\hat{T}'(t)\| \leq e^{\omega' t} \quad (2.20)$$

erfüllen, mit Konstanten  $\omega, \omega' \in \mathbb{R}$ .  $\hat{\Omega}$  beziehungsweise  $\hat{\Omega}'$  seien die Generatoren der beiden Halbgruppen.  $\mathcal{D}(\hat{\Omega} + a\hat{\Omega}') = \mathcal{D}(\hat{\Omega}) \cap \mathcal{D}(\hat{\Omega}')$  und  $(\lambda - \hat{\Omega} - a\hat{\Omega}') \mathcal{D}$  seien dicht in  $\mathcal{B}$  für ein  $\lambda > \omega + a\omega'$  und beliebige Zahlen  $a \in \mathbb{R}_0^+$ . Dann ist der Abschluß von  $\hat{\Omega} + a\hat{\Omega}'$  Generator einer  $(C_0)$ -Halbgruppe  $\{\hat{S}_a(t)\}_{t>0}$ , und diese ist von der Form

$$S_a(t) = \lim_{h \rightarrow 0} [\hat{T}(h) \hat{T}'(ah)]^{[t/h]} \quad (2.21)$$

mit  $[t/h] = \min \{n \in \mathbb{N} \mid n \leq t/h\}$ .

BEWEIS: Für die beiden Halbgruppen  $\{\hat{T}(t)\}_{t>0}$  und  $\{\hat{T}'(t)\}_{t>0}$  gilt nach Voraussetzung

$$\|[\hat{T}(h) \hat{T}'(ah)]^k\| \leq e^{(\omega + a\omega')t}$$

für alle  $h$  und alle  $k$ . Nun setzen wir

$$\hat{F}(t) := \hat{T}(t) \hat{T}'(at), \quad t > 0;$$

damit erhalten wir zunächst

$$\lim_{h \searrow 0} \frac{\hat{T}(t) \hat{T}'(at) f - f}{h} = \lim_{h \searrow 0} \frac{\hat{T}'(ah) f - f}{h} + \lim_{h \searrow 0} \frac{\hat{T}(h) f - f}{h} = \hat{\Omega} f + a \hat{\Omega}' f$$

für alle  $f \in \mathcal{D}(\hat{\Omega} + a \hat{\Omega}')$  und nach Anwendung der Chernoff-Produktformel mit  $n = [t/h]$  die Behauptung.  $\square$

Formel (2.21) findet man häufig auch in der symbolischen Form<sup>17</sup>

$$e^{-t(\overline{\hat{A} + \hat{B}})} = \lim_{n \rightarrow \infty} (e^{-t\hat{A}/n} e^{-t\hat{B}/n})^n.$$

Diese Schreibweise wird unten gleich zum Einsatz kommen.

Wir erwähnen ausdrücklich, daß nicht jede beliebige Summe der Generatoren zweier stark stetiger Halbgruppen, die selbst wieder der Generator einer stark stetigen Halbgruppe ist, automatisch durch die Trotter-Produktformel darstellbar ist; der dadurch gegebene Ausdruck ist nicht immer eine konvergente Folge. Entscheidend sind dabei die Bedingungen (2.20), die man auch *Stabilitätsbedingungen* für Halbgruppen nennt. Beispiele für Produkte aus Halbgruppen, für welche die Trotter-Produktformel nicht gilt, findet man in [329] und [338]. Die oben beschriebene Version der Trotterschen Produktformel gilt gleichwohl unter sehr allgemeinen Voraussetzungen<sup>18</sup>. In vielen Fällen genügt bereits folgende schwächere Form:

**Korollar:** (*Trotter-Produktformel für unitäre Gruppen*)  $\mathcal{H}$  sei ein Hilbertraum,  $\hat{A}$  und  $\hat{B}$  seien selbstadjungierte Operatoren auf  $\mathcal{H}$ , und  $\hat{A} + \hat{B}$  sei wesentlich selbstadjungiert auf  $\mathcal{D}(\hat{A}) \cap \mathcal{D}(\hat{B}) \subset \mathcal{H}$ . Dann gilt

$$e^{it(\overline{\hat{A} + \hat{B}})} = \lim_{n \rightarrow \infty} (e^{it\hat{A}/n} e^{it\hat{B}/n})^n, \quad (2.22)$$

wobei die Konvergenz gleichmäßig in  $t$  ist für alle beschränkte Teilmengen von  $\mathbb{R}$ . Sind  $\hat{A}$  und  $\hat{B}$  außerdem nach unten beschränkt, dann gilt auch

$$e^{-t(\overline{\hat{A} + \hat{B}})} = \lim_{n \rightarrow \infty} (e^{-t\hat{A}/n} e^{-t\hat{B}/n})^n \quad (2.23)$$

mit ebenfalls in  $t$  gleichmäßiger Konvergenz für alle beschränkten Teilmengen von  $\mathbb{R}$ .

Diese spezielle Version der Trotterschen Formel findet sich üblicherweise in der Literatur, meist unter derselben Bezeichnung wie die oben beschriebene allgemeinere; im weiteren Verlauf des vorliegenden Buches wird die schwächere Variante ausreichen.

### 2.3.1.2 Anwendung auf den Zeitentwicklungsoperator

Die Trotter-Produktformel in der speziellen Form für unitäre Gruppen läßt sich nun unmittelbar auf den Standard-Hamiltonoperator

$$\hat{H} = -\nabla^2 + V = \hat{H}_0 + V$$

<sup>17</sup>Man beachte aber auch hier die in Anmerkung 17 erwähnten Einschränkungen.

<sup>18</sup>Weitere Verallgemeinerungen liefern [303] und [305].

der nichtrelativistischen Quantenmechanik anwenden, denn dieser ist für eine umfangreiche Klasse von Potentialen wesentlich selbstadjungiert<sup>19</sup>. E. Nelson führte das 1964 erstmals durch [386], mit folgendem Resultat:

**Satz:** Sei  $V \in \mathcal{L}^2(\mathbb{R}^d) + \mathcal{L}^\infty(\mathbb{R}^d)$ . Dann gilt

$$\begin{aligned} \left( e^{-it(\hat{H}_0+V)} f \right) (x_0) &= \lim_{n \rightarrow \infty} \left( \frac{4\pi it}{n} \right)^{-3n/2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \dots \int_{\mathbb{R}^d} e^{iS_n(x_0, x_1, \dots, x_n, t)} \times \\ &\quad \times f(x_n) dx_1 dx_2 \dots dx_n \end{aligned} \quad (2.24)$$

mit

$$S_n(x_0, x_1, \dots, x_n, t) = \sum_{i=1}^n \frac{t}{n} \left[ \frac{1}{4} \left( \frac{n|x_i - x_{i-1}|}{t} \right)^2 - V(x_i) \right].$$

**BEWEIS:** Wir notieren zunächst, daß der freie Hamilton-Operator unter der Fourier-Transformation  $\mathfrak{F}$  die Relation

$$\hat{H}_0 = \mathfrak{F}^{-1} \lambda^2 \mathfrak{F}$$

erfüllt; daraus folgt für beliebige beschränkte meßbare Funktionen

$$\varphi(\hat{H}_0) = \mathfrak{F}^{-1} \varphi(\lambda^2) \mathfrak{F},$$

also insbesondere für  $\Im t \leq 0$

$$e^{-i\hat{H}_0 t} = \mathfrak{F}^{-1} e^{-i\lambda^2 t} \mathfrak{F}.$$

Sei nun  $\alpha \in \mathbb{R}^+$  mit  $\Re \alpha > 0$ . Dann gehört die Funktion  $\varphi(\lambda) = e^{-\alpha \lambda^2}$  zu  $\mathcal{L}^2(\mathbb{R}^d) \cap \mathcal{L}^\infty(\mathbb{R}^d)$  und hat die Fouriertransformierte  $\bar{\varphi}(x) = (2\alpha)^{-d/2} e^{-x^2/4\alpha}$ . Daraus folgt für beliebige Funktionen  $f \in \mathcal{L}^2(\mathbb{R}^d)$

$$\left( e^{-\alpha \hat{H}_0} f \right) (x) = \frac{1}{(4\pi\alpha)^{d/2}} \int_{\mathbb{R}^d} e^{-\|x-y\|^2/4\alpha} f(y) dy. \quad (2.25)$$

Schreiben wir  $\alpha = \varepsilon + it$  mit  $\varepsilon > 0$ , so erhalten wir für  $f \in \mathcal{L}^2(\mathbb{R}^d) \cap \mathcal{L}^1(\mathbb{R}^d)$

$$\text{l.i.m.}_{\varepsilon \rightarrow 0} e^{-i(t-i\varepsilon)\hat{H}_0} f = e^{-it\hat{H}_0} f,$$

wobei mit l.i.m. der Grenzwert in der  $\mathcal{L}^2$ -Norm gemeint ist. Damit gilt punktweise

$$(e^{-it\hat{H}_0} f)(x) = \lim_{\varepsilon \rightarrow 0} (e^{-i(t-i\varepsilon)\hat{H}_0} f)(x)$$

und mit (2.25) weiter

$$(e^{-it\hat{H}_0} f)(x) = \lim_{\varepsilon \rightarrow 0} \frac{1}{[4\pi i(t-i\varepsilon)]^{d/2}} \int_{\mathbb{R}^d} e^{-\|x-y\|^2/4i(t-i\varepsilon)} f(y) dy$$

<sup>19</sup>Das wurde für die wichtigsten Fälle von Kato bewiesen, siehe dazu [299] - [304].

$$= \frac{1}{(4\pi it)^{d/2}} \int_{\mathbb{R}^d} e^{i\|x-y\|^2/4t} f(y) dy.$$

Um das jetzt auf ganz  $\mathcal{L}^2(\mathbb{R}^d)$  auszudehnen, betrachten wir eine abgeschlossene Kugel im  $\mathbb{R}^d$  mit dem Ursprung als Mittelpunkt und Radius  $R$  sowie deren charakteristische Funktion

$$\chi_R(x) = \begin{cases} 1 & \text{für } \|x\| \leq R \\ 0 & \text{für } \|x\| > R \end{cases}$$

Für beliebige Funktionen  $f \in \mathcal{L}^2(\mathbb{R}^d)$  gilt damit  $\chi_R(x) f \in \mathcal{L}^1(\mathbb{R}^d)$ , und nach dem Theorem von Plancherel folgt für die zugehörigen Fouriertransformierten

$$\lim_{R \rightarrow \infty} \overline{\chi_R(x) f} = \bar{f}$$

und damit weiter

$$(e^{-it\hat{H}_0} f)(x) = \lim_{R \rightarrow \infty} \frac{1}{(4\pi it)^{d/2}} \int_{\|y\| \leq R} e^{i\|x-y\|^2/4t} f(y) dy. \quad (2.26)$$

Nach Voraussetzung ist nun  $V \in \mathcal{L}^2(\mathbb{R}^d) + \mathcal{L}^\infty(\mathbb{R}^d)$  und  $\hat{H}_0 + V$  folglich auf  $\mathcal{D}(\hat{H}_0)$  wesentlich selbstadjungiert. Damit wird die Trotter-Produktformel anwendbar, und (2.22) liefert

$$e^{-it(\hat{H}_0+V)} = \lim_{n \rightarrow \infty} (e^{-it\hat{H}_0/n} e^{-itV/n})^n.$$

Durch Einsetzen von (2.26) erhalten wir die Behauptung.  $\square$

Auf analogem Weg läßt sich eine entsprechende Relation für Integrale im Phasenraum beweisen. – Nelsons Anwendung der Trotter-Produktformel auf den Standard-Hamiltonoperator paßt auf den ersten Blick genau auf Feynmans heuristische Betrachtung, wie sie in den Abschnitten 2.2.2 und 2.2.5 beschrieben wurde und in Ausdrücken wie (2.14), (2.15) oder (2.18) gipfelte, denn

$$\hat{U}(t, t_0) = e^{-i(t-t_0)(\hat{H}_0+V)}$$

ist gerade der zum Standard-Hamilton-Operator gehörende Zeitentwicklungsoperator. Damit ist

$$\psi(x, t) = \lim_{n \rightarrow \infty} \left( \frac{4\pi it}{n} \right)^{-dn/2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \dots \int_{\mathbb{R}^d} e^{iS_n(x, x_1, \dots, x_n, t)} f(x_n) dx_1 dx_2 \dots dx_n \quad (2.27)$$

Lösung der zeitabhängigen Schrödingergleichung

$$i \frac{\partial \psi}{\partial t} = (-\nabla^2 + V) \psi$$

mit Anfangsbedingung  $\psi(x, 0) = f(x)$ . Das entspricht gerade (2.13) beziehungsweise (2.17). Die Interpretation von (2.27) als *Pfadintegral*

$$\psi(x, t) = \int_{\Gamma_x} e^{iS(\gamma(x))} f(\gamma(0)) \mathcal{D}\gamma(x) \quad (2.28)$$

mit

$$\mathcal{D}\gamma(x) = \lim_{n \rightarrow \infty} \left( \frac{4\pi i t}{n} \right)^{-3n/2} \prod_{j=1}^n dx_j$$

und

$$S(x(t)) = \lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{t}{n} \left[ \frac{1}{4} \left( \frac{n|x_i - x_{i-1}|}{t} \right)^2 - V(x_i) \right] = \int_0^t \left( \frac{1}{4} v^2 - V \right) dt' = S_{\text{kl}}$$

über eine geeignete Menge von Wegen  $\{\Gamma_x\}$ , die alle zum Zeitpunkt  $t$  am Ort  $x$  enden, scheint nun nahezuliegen. Das wäre so etwas wie eine mathematisch korrekte Fassung von (2.15) aus Abschnitt 2.2.2; eine solche Vorgehensweise funktioniert jedoch nicht, wie wir gleich sehen werden.

## 2.3.2 Mathematische Unzulänglichkeiten der Feynmanschen Pfadintegrale

### 2.3.2.1 Unendlichdimensionale Maße

Das Haupthindernis, das dem Ziel der exakten Definition eines Integrals wie (2.14), (2.15) oder (2.27) entgegensteht, ist die Tatsache, daß es nicht möglich ist, den dabei verwendeten Maßen  $\mathcal{D}x$  beziehungsweise  $\mathcal{D}p \mathcal{D}q$  eine konkrete mathematische Bedeutung zuzuordnen. Erst wenn man von der unendlichfach iterierten Integration etwa in (2.27) zu einem einzelnen Integral übergeht, das durch ein Maß auf einem unendlichdimensionalen Raum definiert ist, kann man mit Berechtigung von einem Pfadintegral sprechen. Der nächstliegende Versuch, dies zu tun, ist die Interpretation von  $\mathcal{D}x$  als unendlichdimensionales Lebesgue-Maß und dem daraus gebildeten Integral als Lebesgue-Integral auf einem unendlichdimensionalen Raum. Es sei daran erinnert, daß das Lebesgue-Maß  $\lambda$  im  $\mathbb{F}^d$  durch folgende Eigenschaften eindeutig bestimmt ist:

- (i)  $\lambda$  ist *lokal endlich*: Für jede abgeschlossene Borelmenge  $B \subset \mathbb{F}^d$  gilt  $\mu(B) < \infty$ ,
- (ii)  $\lambda$  ist *strikt positiv*: Für jede nichtleere Teilmenge  $A$  von  $\mathbb{F}^d$  gilt  $\lambda(A) > 0$ ,
- (iii)  $\lambda$  ist *translationsinvariant*: Ist  $A$  eine beliebige Lebesgue-meßbare Teilmenge von  $\mathbb{F}^d$  und  $x \in \mathbb{F}^d$  beliebig, dann gilt  $\lambda(A) = \lambda(A + x)$ .

Die Konstruktion eines Maßes mit diesen Eigenschaften für unendlichdimensionale Räume ist jedoch *nicht möglich*. Das sieht man sehr schnell, wenn man einen unendlichdimensionalen Banachraum  $\mathcal{E}$  betrachtet und annimmt,  $\lambda$  sei ein nichttriviales Borelmaß auf  $\mathcal{E}$  mit den Eigenschaften (i), (ii) und (iii). Für  $\delta > 0$  sei  $K_\delta$  die offene Kugel mit dem Ursprung als Mittelpunkt und Radius  $\delta$ . Da  $\mathcal{E}$  unendlichdimensional ist, gibt es eine unendliche Folge  $\{K_{n,\delta/4}\}_{n \in \mathbb{N}}$  von disjunkten offenen Kugeln mit Radius  $\delta/4$  und  $\bigcup_{n=0}^{\infty} K_{n,\delta/4} \subsetneq K_\delta$ . Jede offene Kugel in  $\mathbb{F}^d$  gehört zu  $\mathcal{B}(\mathbb{F}^d)$ . Wegen (ii) gilt

$$0 < \lambda(K_{0,\delta/4}) = \lambda(K_{1,\delta/4}) = \lambda(K_{2,\delta/4}) = \dots < \infty,$$

woraus mit (iii)

$$\lambda(K_\delta) \geq \sum_{n=0}^{\infty} \lambda(K_{n,\delta/4}) = \infty,$$

folgt – ein Widerspruch zu (i). Mit anderen Worten: *Es gibt kein unendlichdimensionales Lebesgue-Maß.* Die einfachste Möglichkeit einer Präzisierung der Feynmanschen Pfadintegrale scheidet damit schon einmal von vorne herein aus.

Natürlich müssen es gar nicht unbedingt Lebesgue-Integrale sein. Der Begriff „Pfadintegral“ erinnert nachdrücklich an die in Abschnitt 2.2 beschriebene gewünschte und im vorigen Abschnitt ebenfalls erwähnte anschauliche Deutung im Sinne einer Integration über eine geeignete Menge von Kurven, also einer Funktionalintegration. Die allgemeinsten Grundmengen, die eine solche Interpretation zulassen, sind Räume stetiger Funktionen, was auch zu der anschaulichen Deutung paßt. Wie sich gleich zeigen wird, sieht die Sache hier jedoch zunächst auch nicht besser aus.

### 2.3.2.2 Maße auf Räumen stetiger Funktionen

Es gilt also nun, anstelle von gewöhnlichen Lebesgue-Maßen geeignete Maße für Teilmengen einer Menge von Kurven zur Definition des gewünschten Integrals zu verwenden. Wir betrachten dazu wieder den Raum  $\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}$  der stetigen Kurven  $\gamma : \mathbb{R} \rightarrow \mathbb{R}^d$  mit  $\gamma(t_a) = x_a$  und  $\gamma(t_b) = x_b$ . Gesucht ist ein Maß  $\mu$  auf diesem Raum, so daß das daraus gebildete Integral die Eigenschaften des heuristischen Feynman-Integrals aufweist. Doch auch hier gilt: *Ein solches Maß auf  $\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}$  gibt es nicht.* Ein besonders einfacher Beweis dieses Sachverhalts stammt von Davey [113]. Die Idee dabei ist, wesentliche, physikalisch zwingend zu fordernde Eigenschaften von  $\mu$  nachzuprüfen. Im Sinne der beschriebenen Interpretation sollte  $\mu$  ein Wahrscheinlichkeitsmaß sein, wofür

$$\mu(\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}) < \infty$$

gelten müßte; zusätzlich sollte auf jeden Fall die Chapman-Kolmogorov-Gleichung (2.4) erfüllt sein. Außerdem müssen natürlich die etablierten Resultate der Standard-Quantenmechanik reproduziert werden. Das einfachste quantenmechanische System ist ein freies Teilchen, also eines mit Hamilton-Operator  $\hat{H} = -\nabla^2$ . Hierfür hat der Propagator bekanntlich<sup>20</sup> die Gestalt

$$K(x_a, x_b, t_a, t_b) = \frac{1}{\sqrt{4\pi i (t_b - t_a)}} \exp \left[ \frac{i (x_b - x_a)^2}{4 (t_b - t_a)} \right],$$

das heißt, es müßte

$$\int_{\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}} \exp \left[ \frac{i}{4} \int_{t_a}^{t_b} \dot{x}(t) dt \right] d\mu(x(t)) = \frac{1}{\sqrt{4\pi i (t_b - t_a)}} \exp \left[ \frac{i (x_b - x_a)^2}{4 (t_b - t_a)} \right]$$

gelten. Die zweite und dritte dieser Forderungen führen jedoch auf

$$\mu(\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}) = \int_{\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}} d\mu(x) = \int_{\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}} |e^{iS(x)}| d\mu(x)$$

<sup>20</sup>Siehe beispielsweise [464].

$$\begin{aligned}
&\geq \left| \int_{\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}} e^{iS(x)} d\mu(x) \right| = \int_0^\infty \left| \int_{\mathcal{C}_{x_a, x_c}^{[t_a, t_c]}} e^{iS(x)} d\mu(x) \int_{\mathcal{C}_{x_c, x_b}^{[t_c, t_b]}} e^{iS(x')} d\mu(x') \right| dx_c \\
&= \int_0^\infty \frac{1}{4\pi \sqrt{(t_c - t_a)(t_b - t_c)}} dx_c = \infty,
\end{aligned}$$

im Widerspruch zur ersten. Damit erweist es sich als unmöglich, ein für die hier verfolgten Zwecke geeignetes Wahrscheinlichkeitsmaß auf  $\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}$  zu finden. Folglich lassen sich Ausdrücke der Form (2.14), (2.15) oder (2.27) auch auf diesem Weg nicht sinnvoll definieren.

### 2.3.2.3 Komplexe Maße

Da wie gesehen kein Weg auf gewöhnliche maßtheoretische Weise zu einer vernünftigen Definition Feynmanscher Integrale führt, bleibt offensichtlich nur die Möglichkeit, dies über irgendwelche verallgemeinerte Maße zu versuchen. Charakteristisch für das heuristische Feynman-Integral ist die Eigenschaft, alle Pfade aus  $\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}$  gleich zu gewichten und sie lediglich mit einer jeweils unterschiedlichen Phase  $e^{iS}$  zu belegen; so kommt man auf die Idee, ein Maß  $\tilde{\mu}$  zu suchen, das für geeignete integrierbare Funktionen

$$\int_{\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}} f(x) d\tilde{\mu}(x) = \int_{\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}} e^{iS(x)} f(x) d\mu(x)$$

liefert.  $\tilde{\mu}$  muß folglich ein *komplexes Maß* sein, also eine  $\sigma$ -additive Funktion von der Borel-Algebra  $\mathcal{B}(\mathcal{C}_{x_a, x_b}^{[t_a, t_b]})$  nach  $\mathfrak{U}^1$ . Daraus erhält man dann in der üblichen Weise ein Integral, indem man ein solches zunächst für in geeignetem Sinne einfache Funktionen definiert und es sodann auf die gewünschte Menge integrierbarer Funktionen erweitert. Als einfache Funktionen kann man hier sogenannte *zylindrischen Funktionen* wählen; diese sind gegeben durch folgende

**Definition:** Eine komplexe Funktion  $f$  auf dem unendlichdimensionalen separablen Hilbertraum  $\mathcal{H}$  heißt *zylindrische Funktion*, wenn es eine Projektion  $\hat{P}$  auf einen endlichdimensionalen Unterraum  $\mathcal{A}$  von  $\mathcal{H}$  gibt, so daß  $f(x) = f(\hat{P}x)$  gilt für alle  $x \in \mathcal{H}$ .  $\mathcal{A}$  heißt dann *Basis* von  $f$ . Die Menge aller Zylinderfunktionen auf  $\mathcal{H}$  heißt  $\mathcal{Z}(\mathcal{H})$ .

Nun sei  $\mathcal{H} = \mathcal{C}_{x_a, x_b}^{[t_a, t_b]}$  und  $\mathcal{F}(\mathcal{C}_{x_a, x_b}^{[t_a, t_b]})$  die Menge aller Funktionen auf  $\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}$ ; diese können als Funktionen mit unendlich vielen Variablen aufgefaßt werden. Mit  $\mathcal{Z}(\mathcal{C}_{x_a, x_b}^{[t_a, t_b]})$  wird hieraus die Menge derjenigen Funktionen ausgewählt, die tatsächlich nur von endlich vielen Variablen abhängen. Ist  $f$  eine beliebige zylindrische Funktion auf  $\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}$  mit Basis  $\mathcal{A} = \{\gamma_0, \gamma_1, \dots, \gamma_n\} \subset \mathcal{C}_{x_a, x_b}^{[t_a, t_b]}$  und  $\gamma_0 = x_a, \gamma_n = x_b$ , außerdem  $z \in \mathfrak{U}^1$  beliebig und



$t_0, t_1, \dots, t_n \in [0, t]$ , so läßt sich dafür ein  $\tilde{\mu}$ -Integral definieren gemäß

$$\begin{aligned} \int_{\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}} f(\gamma) d\tilde{\mu}(\gamma) &= \int_{\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}} \exp \left[ z \int_0^t |\dot{\gamma}(t')|^2 dt' \right] f(\gamma(t_0), \gamma(t_1), \dots, \gamma(t_n)) d\mu \\ &= \prod_{j=0}^{n-1} \left[ -\frac{c}{\pi} (t_{j+1} - t_j) \right]^{d/2} \int_{\mathbb{Z}^d} \int_{\mathbb{Z}^d} \dots \int_{\mathbb{Z}^d} \exp \left( z \sum_{j=0}^{n-1} \frac{|\gamma_{j+1} - \gamma_j|^2}{t_{j+1} - t_j} \right) \times \\ &\quad \times f(\gamma_0, \gamma_1, \dots, \gamma_n) d\gamma_0 d\gamma_1 \dots d\gamma_n. \end{aligned} \quad (2.29)$$

Anschließend erweitert man diesen Integralbegriff auf eine größere Teilmenge von  $\mathcal{F}(\mathcal{C}_{x_a, x_b}^{[t_a, t_b]})$ , bestehend aus Funktionen, die sich beliebig genau durch zylindrische Funktionen approximieren lassen. Diese Funktionen sind dann  $\tilde{\mu}$ -integrierbar. Man überzeugt sich leicht, daß unter der Voraussetzung der Existenz eines solchen Maßes die Wahl  $z = i/4$  genau das heuristische Feynman-Integral reproduzieren würde. Cameron konnte jedoch 1960 zeigen, daß auch dieser Versuch nicht funktioniert; genauer gesagt bewies er das folgende Resultat [89].

**Satz:** (Cameron) *Es gibt kein endliches komplexes Maß  $\tilde{\mu}$ , so daß für  $\Im z \neq 0$  für jede Funktion  $f \in \mathcal{Z}(\mathcal{C}_{x_a, x_b}^{[t_a, t_b]})$  das Integral  $\int_{\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}} f(\gamma) d\tilde{\mu}(\gamma)$  durch (2.29) gegeben ist.*

**BEWEIS:** Wir nehmen an, es gäbe ein solches Maß und zeigen, daß diese Annahme zu einem Widerspruch führt. Für  $0 < t_0 < t_1 < \dots < t_n \in [0, t]$  und  $a \in \mathfrak{L}^1$  mit  $\Re a \geq 0$  definieren wir die Funktion  $g : \mathcal{C}_{x_a, x_b}^{[t_a, t_b]} \rightarrow \mathfrak{L}^1$  durch

$$g(\gamma) := \exp \left[ -a \sum_{j=0}^{n-1} \frac{|\gamma_{j+1} - \gamma_j|^2}{t_{j+1} - t_j} \right].$$

Hierfür erhält man mit Hilfe der Substitution  $\chi_j = \frac{\gamma_{j+1} - \gamma_j}{\sqrt{t_{j+1} - t_j}}$  für  $j = 0, 1, \dots, n-1$

sowie der weiter oben auch schon verwendeten Relation  $\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$

$$\begin{aligned} &\prod_{j=0}^{n-1} \left[ -\frac{z}{\pi} (t_{j+1} - t_j) \right]^{d/2} \int_{\mathbb{Z}^d} \int_{\mathbb{Z}^d} \dots \int_{\mathbb{Z}^d} \exp \left( z \sum_{j=0}^{n-1} \frac{|\gamma_{j+1} - \gamma_j|^2}{t_{j+1} - t_j} \right) \times \\ &\quad \times g(\gamma_0, \gamma_1, \dots, \gamma_n) d\gamma_0 d\gamma_1 \dots d\gamma_n \\ &= \left( -\frac{c}{\pi} \right)^{nd/2} \left\{ \int_{-\infty}^{\infty} \exp [(z - a) \chi^2] d\chi \right\}^{nd} = \left( 1 - \frac{a}{2z} \right)^{-nd/2}, \end{aligned}$$

und folglich auch, sofern das Maß  $\tilde{\mu}$  existiert,

$$\int_{\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}} g(\gamma) d\tilde{\mu}(\gamma) = \left(1 - \frac{a}{2z}\right)^{-nd/2}.$$

Nun weiß man, daß  $\sigma$ -additive komplexwertige Maße stets von endlicher totaler Variation sind<sup>21</sup>, das heißt  $|\tilde{\mu}|$  muß ein endliches positives Maß auf  $\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}$  sein. Das kann mit Hilfe der aus  $|g(\gamma)| \leq 1$  für alle  $\gamma \in \mathcal{C}_{x_a, x_b}^{[t_a, t_b]}$  folgenden Abschätzung

$$|\tilde{\mu}|(\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}) = \left| \int_{\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}} d\tilde{\mu}(\gamma) \right| \geq \int_{\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}} |g(\gamma)| d\tilde{\mu}(\gamma) \geq \left| \int_{\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}} g(\gamma) d\tilde{\mu}(\gamma) \right| = \left|1 - \frac{a}{2z}\right|^{-nd/2}$$

überprüft werden. Wählt man nun speziell  $a = z - \varepsilon$  mit  $\varepsilon > 0$ , so findet man

$$\left|1 - \frac{a}{2z}\right| = \left|1 - \frac{a\bar{z}}{|z|^2}\right| = \frac{\varepsilon}{|z|}$$

und damit

$$\left|1 - \frac{a}{2z}\right| < 1$$

für alle genügend kleine  $\varepsilon$ . Daraus folgt

$$\lim_{n \rightarrow \infty} \left|1 - \frac{a}{2z}\right|^{-nd/2} = \infty,$$

ein Widerspruch. □

Insbesondere ist damit auch der gerade interessierende Fall  $z = i/4$  ausgeschlossen. Es gibt also kein  $\sigma$ -additives komplexwertiges Maß, welches die Anforderungen erfüllt, die erforderlich sind, um Feynman-Integralen auf diesem Weg eine mathematische Bedeutung zu verleihen.

### 2.3.2.4 Weitere Schwierigkeiten

Die Nichtexistenz des in Feynman-Integralen auftretenden Maßes ist nicht das einzige Problem. Selbst wenn es irgendwie möglich wäre, ein geeignetes Maß zu beschaffen, sähe die Sachlage nicht besser aus. Der Raum  $\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}$  ist nicht kompakt, folglich würde auch die Existenz eines reellen  $\sigma$ -additiven Maßes darauf nichts nützen, denn ein solches würde, angewandt auf den gesamten Raum  $\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}$ , ebenfalls einen unendlichen Wert liefern. Daher wäre der Integrand in (2.14) nicht absolut integrierbar und das Integral wieder nicht definiert.

Es kommt noch ein weiterer Aspekt hinzu: Die Wirkungsfunktion

$$S_{\text{kl}} = \int L_{\text{kl}}(x(t), \dot{x}(t)) dt$$

<sup>21</sup>Siehe beispielsweise [107].

hängt im Allgemeinen nicht nur vom Weg des Systems, sondern auch von dessen zeitlicher Ableitung ab. Damit ist sie auf den meisten Elementen von  $\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}$  ein extrem singuläres Objekt, da von den Elementen dieses Raums nur Stetigkeit gefordert wird. Außerdem hat der Integrand in (2.15) oder (2.18) für reellwertige Potentiale den konstanten Betrag 1. Beides führt schon jeweils alleine dazu, daß solche Integrale im Sinne absolut konvergenter Lebesgue-Integrale nicht existieren<sup>22</sup>.

Noch schwieriger ist die Situation bei Feynman-Integralen im Phasenraum. Hier laufen die Integrale über Pfade, die von den Orts- *und* den Impulskordinaten abhängen. Wie wir gesehen haben und noch genauer sehen werden, sind nichtdifferenzierbare Pfade im Konfigurationsraum ausdrücklich zugelassen, und in der Tat sind fast alle möglichen Pfade nicht differenzierbar. Folglich sind fast alle verallgemeinerten Impulse nicht einmal stetig, denn diese sind proportional zu den zeitlichen Ableitungen der verallgemeinerten Koordinaten. Damit ist man im Phasenraum noch wesentlich weiter als im Konfigurationsraum von der Möglichkeit entfernt, Feynman-Integrale rigoros definieren zu können.

Es ist zwar in der theoretischen Physik üblich, all das zu ignorieren und mit Feynman-Integralen zu hantieren, als ob sie existierten; bekanntlich geschieht das aus praktischer Sicht mit großem Erfolg, sowohl im Hinblick auf den heuristischen Wert als auch auf den immensen Nutzen für konkrete Berechnungen in der relativistischen Quantenmechanik und der Quantenfeldtheorie<sup>23</sup>. Das macht jedoch die Notwendigkeit einer präzisen mathematischen Formulierung um so deutlicher. Darüber hinaus führt die Tatsache, daß es eine solche für Feynman-Integrale in ihrer Originalversion nicht gibt, wenig überraschend auch zu ganz direkten Schwierigkeiten. Versucht man beispielsweise, bei Feynman-Integralen gewisse, von gewöhnlichen Integrationstheorien wie derjenigen von Lebesgue, Lebesgue-Stieltjes oder dergleichen vertraute Tricks wie Differenzieren unter dem Integral, dem Satz von Fubini oder Koordinatentransformationsformeln anzuwenden, kann das zu falschen Resultaten führen.

Es erscheint sehr merkwürdig, daß Feynman-Integrale vermeintlich eine Formulierung der Quantenmechanik erlauben, die ausschließlich von klassischen Größen Gebrauch macht; vermutlich ist diese Fehleinschätzung einer der wesentlichen Auslöser für die merkwürdigen Konzepte, die in der Schulphysik zur Einführung der Quantenmechanik auftreten und Gegenstand einiger der folgenden Abschnitte des vorliegenden Buches sind. In der Tat ist der Feynmansche Zugang zur Quantenmechanik genau so vollständig nicht-klassisch wie alle anderen; das wird gerade auch an den eben geschilderten mathematischen Unzulänglichkeiten unübersehbar deutlich. Diese führen dazu, daß eine klassische

---

<sup>22</sup>Man könnte mutmaßen, daß die hier geübte Beschränkung auf die nichtrelativistische Quantenmechanik zumindest Mitschuld an diesem Mangel tragen könnte, doch leider sieht die Sache für relativistische Systeme nicht wirklich besser aus. Ichinose konnte zeigen, daß Ausdrücke der Form  $\int_C f(x) e^{iS(x)} \mathcal{D}x$  im obigen Sinn für durch eine Dirac-Gleichung beschriebene Systeme in zweidimensionalen Raum-Zeiten existieren [256] - [261], Zastawniak bewies jedoch, daß sie für solche Systeme in vierdimensionalen Raum-Zeiten nicht sinnvoll definierbar sind [537], [538]. Auch hierfür gibt es mathematisch präzise Rekonstruktionen; siehe etwa [179], [180], [181] oder [421]. Wie bereits angedeutet ist wie nicht selten anderswo auch hier die Quantenfeldtheorie die Lösung der Probleme; der Grund dafür ist im wesentlichen die mathematische Äquivalenz derselben zu speziellen Gebieten der statistischen Mechanik. Abschnitt 2.4.3.3 kommt kurz darauf zurück, eine weitergehende Diskussion quantenfeldtheoretischer Pfadintegrale würde jedoch den Rahmen des vorliegenden Buches sprengen.

<sup>23</sup>Es sei erneut nachdrücklich bemerkt, daß das in der nichtrelativistischen Quantenmechanik *nicht der Fall ist*.

Definition der grundlegenden quantenmechanischen Größen *gerade nicht möglich ist*.

Trotz allem kann die überzeugende praktische Tauglichkeit des heuristischen Feynman-Integrals in vielen Bereichen eigentlich kein Zufall sein. Und ähnlich wie bei den Distributionen ist es auch hier möglich, die Sache auf mathematisch solide Fundamente zu stellen, allerdings längst nicht in vergleichbar einheitlicher und lückenloser Form. Denn es gibt im Fall der Pfadintegrale gleich mehrere Möglichkeiten der Präzisierung, die alle etwas miteinander zu tun haben, sich im Detail jedoch teilweise erheblich unterscheiden. Gemeinsam haben sie alle die Eigenschaften, jeweils nicht unmittelbar in voller Allgemeinheit auf quantenmechanische Systeme anwendbar zu sein und außerdem allesamt einen gewissen nicht unerheblichen mathematischen Aufwand erforderlich zu machen. Insbesondere erlauben sie keine klassische Interpretationen. Drei Beispiele betrachten wir nun exemplarisch mehr oder weniger genau<sup>24</sup>.

## 2.4 Die Feynman-Kac-Formel

Die Schwierigkeiten, die beim Versuch auftreten, ein Maß im Sinne des oben beschriebenen heuristischen Feynman-Integrals zu konstruieren, sind wesentlich durch den oszillatorischen Charakter des Phasenfaktors  $\exp(iS)$  verursacht; mit anderen Worten, das  $i$  ist schuld. Dieses  $i$  tritt in den Pfadintegralen auf, weil sie als Integralkerne des Zeitentwicklungsoperators  $\hat{U} = \exp(it\hat{H})$  konstruiert werden. Daher bietet es sich an, zur Vermeidung der maßtheoretisch störenden Oszillationen stattdessen Integralkerne des Operators  $\hat{A} = \exp(t\hat{H})$  zu betrachten. Diese Strategie wurde aufbauend auf Vorarbeiten von Norbert Wiener wesentlich von Mark Kac ausgearbeitet [290] – [292], der sich zum Ziel gesetzt hatte, mit Feynmans Pfadintegralen aus mathematischer Sicht klar zu kommen. Was dabei heraus kam, ist als *Feynman-Kac-Formel* bekannt geworden und leistet genau das gewünschte, nämlich eine Integraldarstellung des erwähnten Exponentialoperators zu liefern.

### 2.4.1 Wiener-Prozeß, Wiener-Maß und Wiener-Integrale

Läßt man das Ziel, (2.14) oder (2.15) eine präzise Bedeutung zu verleihen, für einmal links liegen, bleibt die Feststellung, daß es durchaus ein mathematisch wohldefiniertes Maß auf dem Raum der stetigen Kurven gibt, nämlich das sogenannte *Wiener-Maß*. Entsprechend sind Pfadintegrale genau besehen bereits seit den Zwanziger-Jahren des vorigen Jahrhunderts bekannt, wenn auch anfangs nicht unter ebendiesem Namen. Betrachtungen solcher Art stammen ursprünglich aus der Theorie der stochastischen Prozesse und wurden erstmals von Norbert Wiener im Zusammenhang mit der mathematisch exakten Modellierung der Brownschen Bewegung beschrieben [524], [525], [526]<sup>25</sup>. Wir

<sup>24</sup>Einen relativ aktuellen, ausführlichen Überblick über Pfadintegrale aus der Sicht der mathematischen Physik liefert [274].

<sup>25</sup>Der physikalische Prozeß der Brownschen Bewegung wurde erstmals 1785 von dem niederländischen Arzt und Botaniker Jan Ingenhousz beobachtet und 1827 von dem schottischen Botaniker Robert Brown wiederentdeckt [82]. Die mathematische Definition und Beschreibung solcher Prozesse in der heutigen Form geht ursprünglich und unabhängig voneinander auf Albert Einstein [137], [138] und Marian Smoluchowski [484] zurück.

werden daher nun einen kurzen Ausflug in die Theorie der Zufallsprozesse machen, wobei wir uns auf diejenigen Begriffe der stochastischen Analysis beschränken<sup>26</sup>, die für die Quantenmechanik von Bedeutung sind<sup>27</sup>.

### 2.4.1.1 Grundbegriffe aus der Stochastik

Wir sammeln zunächst einige stochastische Grundbegriffe, die für das folgende gebraucht werden. Eine stetige Zufallsvariable  $X$  mit der Wahrscheinlichkeitsdichte

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left[ -\frac{1}{2} \left( \frac{x - \mu}{\sigma} \right)^2 \right]$$

heißt  $\mu$ - $\sigma$ -normalverteilt mit Erwartungswert  $\mu$  und Standardabweichung  $\sigma$ . Entsprechend sagen wir,  $n$  stochastisch unabhängige Zufallsvariable  $X_1, X_2, \dots, X_n$  seien normalverteilt, wenn ihre Wahrscheinlichkeitsdichte

$$f(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^n \prod_{i=1}^n \sigma_i} \exp \left[ -\frac{1}{2} \sum_{i=1}^n \left( \frac{x_i - \mu_i}{\sigma_i} \right)^2 \right]$$

lautet. Es seien nun  $(\Omega, \mathcal{F}, \mathfrak{P})$  ein Wahrscheinlichkeitsraum und  $(\mathcal{E}, \mathcal{E})$  ein meßbarer Raum, das heißt,  $\Omega$  sei eine beliebige Menge,  $\mathcal{F}$  eine  $\sigma$ -Algebra darüber und  $\mathfrak{P}$  ein Wahrscheinlichkeitsmaß auf  $\mathcal{F}$ , außerdem sei  $\mathcal{E}$  eine weitere Menge und  $\mathcal{E}$  wiederum eine  $\sigma$ -Algebra über  $\mathcal{E}$ . Eine Familie  $\{X_t\}_{t \geq 0}$  von Zufallsvariablen  $X_t : (\Omega, \mathcal{F}, \mathfrak{P}) \rightarrow (\mathcal{E}, \mathcal{E})$  mit  $t \in \mathbb{R}_0^+$  heißt *stochastischer Prozeß*. Für stochastische Prozesse definieren wir in üblicher Weise die Mittelwertfunktion

$$m(t) := \mathbf{E}(X_t)$$

als Erwartungswert und die Kovarianzfunktion

$$\mathbf{C}(s, t) := \text{cov}(X_s, X_t) = \mathbf{E}((X_s - m(s))(X_t - m(t))).$$

Im folgenden sei  $\mathcal{E} = \mathbb{R}$  und  $\mathcal{F} = \mathcal{B}(\mathbb{R})$  die Borel- $\sigma$ -Algebra über  $\mathbb{R}$ . Ein reellwertiger stochastischer Prozeß heißt *Gauß-Prozeß*, falls für alle  $t_1 < t_2 < \dots < t_n$  mit  $n \in \mathbb{N}$  die Verteilung des Vektors  $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$  eine  $n$ -dimensionale Gauß-Verteilung ist. Das läßt sich leicht auf  $d$  Dimensionen verallgemeinern. Dazu sei jetzt  $\mathcal{E} = \mathbb{R}^d$  und  $\mathcal{E} = \mathcal{B}(\mathbb{R}^d)$ . Ein  $d$ -dimensionaler Gauß-Prozeß ist ein Vektor  $\{X_{1,t}, X_{2,t}, \dots, X_{d,t}\}_{t \geq 0}$  aus  $d$  Gaußschen Prozessen. Wir betrachten nun eine spezielle Klasse Gaußscher Prozesse.

**Definition:** (*Wiener-Prozesse*)

- (i) Ein stochastischer Prozeß  $\{W_t\}_{t \geq 0}$  heißt *Wiener-Prozeß*, falls  $\{W_t\}_{t \geq 0}$  ein Gauß-Prozeß ist mit  $m(t) = 0$  und  $\mathbf{C}(s, t) = \min\{s, t\}$ .

<sup>26</sup>Ausführliches dazu findet man zum Beispiel in [268], [298] oder [336].

<sup>27</sup>Der hier beschrittene ursprüngliche Zugang zum Wiener-Maß ist zwar der natürlichste und anschaulichste, es geht aber auch direkter ohne Verwendung stochastischer Prozesse, wie Nelson gezeigt hat [386]; vergleiche auch [427] und [504].

- (ii) Ein  $d$ -dimensionaler stochastischer Prozeß  $\{(W_{1,t}, W_{2,t}, \dots, W_{d,t})\}_{t \geq 0}$  heißt  $d$ -dimensionaler Wiener-Prozeß, falls die Koordinaten  $W_{i,t}$  jeweils unabhängige Wiener-Prozesse sind.

Im folgenden schreiben wir jeweils abkürzend  $W_t$  für den Vektor  $(W_{1,t}, W_{2,t}, \dots, W_{d,t})$ .  $d$ -dimensionale Wiener-Prozesse stellen die mathematische Modellierung der  $d$ -dimensionalen *Brownschen Bewegung* dar. Wir erwähnen ohne Beweis einige weitere mathematische Eigenschaften<sup>28</sup>. Wiener-Prozesse besitzen stets unabhängige, stationäre normalverteilte Zuwächse  $W_{t_{i+1}} - W_{t_i}$ ,  $i = 1, 2, \dots, n$ . Es gilt  $\mathfrak{P}(W_0 = 0) = 1$ . Sind  $A_1, A_2, \dots, A_n$  irgendwelche meßbaren Teilmengen des  $\mathbb{R}^d$ , erhält man

$$\begin{aligned} & \mathfrak{P}(W_{t_n} \in A_n, \dots, W_{t_2} \in A_2, W_{t_1} \in A_1) \\ &= \int_{A_n} \dots \int_{A_2} \int_{A_1} K(x_n - x_{n-1}, t_n - t_{n-1}) \dots K(x_2 - x_1, t_2 - t_1) K(x_1, t_1) dx_1 dx_2 \dots dx_n, \end{aligned}$$

mit der Dichte

$$K(x, t) = \frac{1}{(2\pi t)^{d/2}} e^{-x^2/2t}.$$

Anschaulich bedeutet das, daß der Weg, den ein Brownsches Teilchen bis zum Zeitpunkt  $t_1$  zurückgelegt hat, völlig unerheblich für dessen weitere Bewegung ist. Mit anderen Worten: Der Prozeß hat kein Gedächtnis. Eine Brownsche Bewegung startet zu jedem Zeitpunkt neu, das heißt, für einen Wiener-Prozeß  $W_t$  und eine beliebig vorgegebene Zeit  $t'$  ist  $\tilde{W}_t = W_{t+t'} - W_t$  ebenfalls ein Wiener-Prozeß mit den selben Verteilungen, der unabhängig von  $W_t$  ist<sup>29</sup>. Prozesse mit dieser Eigenschaft heißen *Markov-Prozesse*. Für beliebige Dichtefunktionen  $K(x, t)$  sind Markov-Prozesse vollständig festgelegt durch die Angabe der Anfangsverteilung  $\mathfrak{P}(W_0 \in A)$  und der *Übergangsfunktion*

$$\mathfrak{P}(W_{t'} \in A \mid W_t = x) = \int_A K(x' - x, t' - t) dx', \quad t' > t;$$

hierbei handelt es sich um eine bedingte Wahrscheinlichkeit, nämlich die Wahrscheinlichkeit, daß sich das Teilchen zum Zeitpunkt  $t'$  am Ort  $x'$  befindet, wenn es zuvor zum Zeitpunkt  $t$  am Ort  $x$  war.

Die hochinteressanten Eigenschaften der Brownschen Bewegung dürften hinlänglich bekannt sein. Die  $d$ -dimensionale Brownschen Bewegung trifft jede  $d + 1$ -dimensionale Kugel um ihren Startpunkt unendlich oft und windet sich um jeden Punkt mit Ausnahme des Startpunktes ebenfalls unendlich oft. Das bedeutet, daß die Menge der Punkte im  $\mathbb{R}^d$ , die der Pfad einer  $d$ -dimensionalen Brownschen Bewegung unendlich oft trifft, selbst eine  $d + 1$ -dimensionale Menge ist; ein solcher Pfad ist eine raumfüllende Kurve. Außerdem ist die Brownsche Bewegung selbstähnlich, das heißt, auf jeder Längenskala zeigt sie das gleiche Verhalten; betrachtet man den Pfad eines Brownschen Teilchens

<sup>28</sup>Details findet man beispielsweise in [268] oder [298].

<sup>29</sup>Inzwischen wurden bei realen Brownschen Teilchen Abweichungen davon entdeckt; genauer gesagt wurden bei der Bewegung von in wässrigen Lösungen suspendierten Plastik- und Quarzkugeln mit Radien im Mikrometerbereich Gedächtniseffekte in Zeitskalen von Millisekunden nachgewiesen [354].

unter beliebiger Verkleinerung oder Vergrößerung, sieht man stets wieder eine Brownsche Bewegung<sup>30</sup>. Die Pfade der Brownschen Bewegung sind unendlich lang und zwar überall stetig, aber im allgemeinen nirgends differenzierbar; genauer gesagt hat die echte Teilmenge der nicht differenzierbaren Brownschen Pfade in der Menge aller Brownschen Pfade das Wahrscheinlichkeitsmaß 1. Die so seltsam anmutenden Zick-Zack-Kurven sind somit im Vergleich zu schön glatten Kurven nicht die Ausnahme, sondern die Regel. Das ist ein Spezialfall von folgendem

**Satz:** *Fast alle stetigen Kurven sind nirgends differenzierbar.*

**BEWEIS:** Es genügt zu zeigen, daß für einen beliebigen Brownschen Pfad die Wahrscheinlichkeit, nirgends differenzierbar zu sein, 1 ist. Sei also  $W_t(\omega)$  ein Brownscher Pfad, und es werde angenommen,  $W_t(\omega)$  sei in  $t_0$  differenzierbar. Dann gibt es einen Vektor  $W'_{t_0}(\omega)$  und dazu für alle  $\varepsilon > 0$  ein  $\delta > 0$ , so daß

$$\left| \frac{W_{t_0+\Delta t}(\omega) - W_{t_0}(\omega)}{\Delta t} - W'_{t_0}(\omega) \right| \leq \varepsilon$$

gilt für  $|\Delta t| \leq \delta$ . Nun sei ohne Beschränkung der Allgemeinheit  $\Delta t > 0$ . Wiener-Prozesse sind Gauß-Prozesse mit den oben erwähnten Eigenschaften, folglich ist die Größe in den Betragsstrichen eine Gaußsche Zufallsverteilung mit Wahrscheinlichkeitsdichte

$$f(x) = \frac{1}{\sqrt{2\pi} W'_{t_0}(\omega)} \exp \left\{ -\frac{1}{2} \left[ \frac{x - W'_{t_0}(\omega)}{\Delta t} \right]^2 \right\}.$$

Die Wahrscheinlichkeit, daß diese Zufallsvariable jedes beliebige  $\varepsilon > 0$  übersteigt, wird beliebig groß, wenn  $\Delta t$  nur genügend klein wird, somit existiert der Grenzwert

$$W'_{t_0}(\omega) = \lim_{\Delta t \rightarrow 0} \frac{W_{t_0+\Delta t}(\omega) - W_{t_0}(\omega)}{\Delta t}$$

nicht, und  $W_t(\omega)$  ist mit der Wahrscheinlichkeit 1 nirgends differenzierbar.  $\square$

#### 2.4.1.2 Integration über Brownsche Pfade

Nun betrachten wir den Raum  $\mathcal{C}_{[0,\infty)}^0(\mathbb{R}^d) \equiv \mathcal{C}$  der stetigen Kurven  $\gamma : \mathbb{R}_0^+ \rightarrow \mathbb{R}^d$  mit  $\gamma(0) = 0$  und darauf die  $d$ -dimensionalen Zufallsvariablen  $\Gamma_t : \mathcal{C} \rightarrow \mathbb{R}^d$  mit  $\Gamma_t(\gamma) = \gamma(t)$ . Außerdem definieren wir  $\mathcal{E} := \{ \{ \Gamma_t \in \mathcal{B} \} \mid \mathcal{B} \in \mathcal{B}(\mathbb{R}^d), t \geq 0 \}$  und konstruieren mit

$$\mathcal{C} = \bigcap \{ \mathcal{S} \subset \mathcal{E} \mid \mathcal{S} \text{ ist eine } \sigma\text{-Algebra} \}$$

die kleinste in  $\mathcal{E}$  enthaltene  $\sigma$ -Algebra. Jeder Wiener-Prozeß  $W = \{W_t\}_{t \geq 0}$  auf  $(\Omega, \mathcal{F}, \mathfrak{P})$  definiert dann eine Funktion von  $\Omega$  nach  $\mathcal{C}$  mit  $W(\omega) = \{W_t(\omega)\}_{t \geq 0}$ . Für  $\mathcal{A} := \{ \Gamma_t \in \mathcal{B} \}$  mit  $\mathcal{B} \in \mathcal{B}$ ,  $t \geq 0$ , gilt

$$W^{-1}(\mathcal{A}) = \{ \omega \mid W(\omega) \in \mathcal{A} \} = \{ \omega \mid W_t(\omega) \in \mathcal{B} \} \in \mathcal{F},$$

<sup>30</sup>Man spricht hier auch von *Skaleninvarianz* und von einer *fraktalen Struktur* Brownscher Pfade.



folglich ist die Abbildung  $W : (\Omega, \mathcal{F}) \longrightarrow (\mathcal{C}, \mathcal{C})$  meßbar, und mit  $\mu_0 = W^{-1} \circ \mathfrak{P}$  erhält man ein Maß auf  $(\mathcal{C}, \mathcal{C})$ . Dieses Maß nennt man *Wiener-Maß*. Der Wahrscheinlichkeitsraum  $(\mathcal{C}, \mathcal{C}, \mu_0)$  heißt auch *Wiener-Raum*.  $\Gamma = \{\Gamma_t\}_{t \geq 0}$  ist damit ein Wiener-Prozeß auf  $(\mathcal{C}, \mathcal{C}, \mu_0)$ .

Wir deuten kurz an, wie man das Wiener-Maß explizit konstruieren kann und machen hierfür zunächst den Raum  $\mathcal{C}$  von oben vermöge der Supremumsnorm  $\|\gamma\| = \sup_{t \in \mathbb{R}_0^+} |\gamma(t)|$

zu einem separablen Banachraum.  $\mathcal{B}(\mathcal{C})$  sei die von allen offenen Kugeln in  $\mathcal{C}$  generierte  $\sigma$ -Algebra über  $\mathcal{C}$ . Außerdem betrachten wir die Menge  $\mathcal{Z} \subset \mathcal{B}(\mathcal{C})$  der *Zylindermengen* oder *Quasi-Intervalle*

$$\mathcal{A}(t_1, t_2, \dots, t_n, \mathcal{B}) := \{\gamma \in \mathcal{C} \mid \gamma(t_1) \in B_1, \gamma(t_2) \in B_2, \dots, \gamma(t_n) \in B_n\}$$

mit  $n \in \mathbb{N}$  und  $\mathcal{B} = B_1 \times B_2 \times \dots \times B_n$  für Borelmengen  $B_1, B_2, \dots, B_n \in \mathcal{B}(\mathbb{R}^d)$ . Die von  $\mathcal{Z}$  erzeugte  $\sigma$ -Algebra ist gerade wieder  $\mathcal{B}(\mathcal{C})$ . Für eine solche Zylindermenge  $\mathcal{A}$  gilt

$$\begin{aligned} \mu_0(\mathcal{A}) &= \int_{\mathcal{A}} d\mu_0(\gamma) \\ &= \int_{B_n} \dots \int_{B_2} \int_{B_1} \prod_{j=1}^n \frac{1}{[2\pi(t_j - t_{j-1})]^{1/2}} \exp \left[ \frac{(x_j - x_{j-1})^2}{2(t_j - t_{j-1})} \right] dx_1 dx_2 \dots dx_n. \end{aligned}$$

Das ist ganz anschaulich die Wahrscheinlichkeit, daß ein Brownsches Teilchen die Zylindermenge  $\mathcal{A}$  durchläuft<sup>31</sup>. Für diejenigen  $\mathcal{B} \in \mathcal{B}(\mathcal{C})$ , die als disjunkte Vereinigung  $\bigcup_{j=1}^{\infty} \mathcal{A}_j$  aus Zylindermengen darstellbar sind, führt das auf

$$\int_{\mathcal{B}} d\mu_0(\gamma) = \sum_{j=1}^{\infty} \int_{\mathcal{A}_j} d\mu_0(\gamma),$$

und damit läßt sich das so gebildete Maß in der üblichen Weise auf ganz  $\mathcal{B}(\mathcal{C})$  erweitern, indem man alle möglichen uneingeschränkten Durchschnitte und abzählbare Vereinigungen von beliebigen Zylindermengen betrachtet.

Mit dem Wiener-Maß kann man nun für meßbare Funktionen  $f : \mathcal{C} \longrightarrow \mathbb{R}^d$  ein Integral

$$I(f) = \int_{\mathcal{C}} f(\gamma) d\mu_0(\gamma)$$

definieren, bei dem über alle Brownschen Pfade im  $\mathbb{R}^d$  integriert wird. Solche Integrale nennt man *Wiener-Integrale*.

Als nächstes sortieren wir für zwei fest gewählte Zeiten  $t_a < t_b$  und zwei fest gewählte Punkte  $x_a, x_b \in \mathbb{R}^d$  alle Elemente  $\gamma$  von  $\mathcal{C}$  mit  $\gamma(t_a) = x_a$  und  $\gamma(t_b) = x_b$  heraus, das heißt, wir beschränken uns auf die Betrachtung aller Wege, auf denen ein Brownsches Teilchen im Zeitintervall  $[t_a, t_b]$  von  $x_a$  nach  $x_b$  gelangen kann. Die Menge aller dieser

<sup>31</sup>Glimm bezeichnet Zylindermengen sehr treffend als Slaloms.  $\mu_0(\mathcal{A})$  ist dann die Wahrscheinlichkeit, daß ein Brownsches Teilchen den entsprechenden Slalom fehlerfrei bewältigt [192].

Pfade nennen wir  $\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}$ . Auf diesem Raum soll nun ein Maß  $\mu_{x_a, x_b}^{[t_a, t_b]}$  definiert werden, das dessen Eigenschaft respektiert, genau alle Brownschen Pfade von  $(x_a, t_a)$  nach  $(x_b, t_b)$  zu enthalten. Die Normalverteilung aller Zuwächse eines Wiener-Prozesses legt nahe, für den ganzen Raum  $\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}$  das Maß

$$\mu_{x_a, x_b}^{[t_a, t_b]}(\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}) = \int_{\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}} d\mu_{x_a, x_b}^{[t_a, t_b]}(\gamma) = \frac{1}{[2\pi(t_b - t_a)]^{d/2}} \exp\left[-\frac{(x_b - x_a)^2}{2(t_b - t_a)}\right]$$

zu fordern<sup>32</sup>. Man erreicht das durch

$$d\mu_{x_a, x_b}^{[t_a, t_b]}(\gamma) := \delta(\gamma(t_a) - x_a) \delta(\gamma(t_b) - x_b) d\mu_0.$$

Das dadurch definierte Maß  $\mu_{x_a, x_b}^{[t_a, t_b]}$  heißt *bedingtes Wiener-Maß*. Wenn Anfangs- und Endpunkt der Brownschen Pfade klar sind, schreiben wir dafür abkürzend  $\mu_{x_a, x_b}^{[t_a, t_b]} \equiv \mu$ . Explizit erhalten wir dieses Maß völlig analog zu oben, wenn wir den Raum  $\mathcal{B}(\mathcal{C})$  durch  $\mathcal{B}(\mathcal{C}_{x_a, x_b}^{[t_a, t_b]})$  ersetzen und diese  $\sigma$ -Algebra wieder durch entsprechende Zylindermengen generieren. Für eine Zylindermenge der Form

$$\mathcal{A} = \{\gamma \in \mathcal{C}_{x_a, x_b}^{[t_a, t_b]} \mid \gamma(t_1) \in B_1, \gamma(t_2) \in B_2, \dots, \gamma(t_n) \in B_n\}$$

liefert das entsprechend

$$\int_{\mathcal{A}} d\mu(\gamma) = \int_{B_n} \cdots \int_{B_2} \int_{B_1} \prod_{j=1}^n \frac{1}{[2\pi(t_j - t_{j-1})]^{1/2}} \exp\left[\frac{(x_j - x_{j-1})^2}{2(t_j - t_{j-1})}\right] dx_1 dx_2 \cdots dx_n,$$

so daß sich  $\mu_{x_a, x_b}^{[t_a, t_b]}$  auf ganz  $\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}$  erweitern läßt. Das bedingte Wiener-Maß unterscheidet sich vom gewöhnlichen Wienermaß nur dadurch, daß bei ersterem die Pfade an zwei ausgewählten Punkten festgehalten werden und bei letzterem nicht.

Für meßbare Funktionen  $f : \mathcal{C}_{x_a, x_b}^{[t_a, t_b]} \longrightarrow \mathbb{R}^d$  definieren wir nun das *Pfadintegral*

$$J(f) = \int_{\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}} f(\gamma) d\mu(\gamma).$$

Dieses interpretieren wir analog zu Abschnitt 2.2 als ein Integral über alle möglichen Brownschen Pfade, die vom Ort  $x_a$  zur Zeit  $t_a$  zum Ort  $x_b$  zur Zeit  $t_b$  führen. Im Unterschied zu dort handelt es sich jedoch hier um ein wohldefiniertes mathematisches Objekt, dessen Existenz bewiesen werden kann. Das führt jedoch noch nicht auf direktem Weg zu einer mathematischen Konsolidierung der Pfadintegrale aus Abschnitt 2.2, einerseits, weil der im Integranden von Feynman-Integralen auftauchende Ausdruck  $e^{iS}$  keine meßbare Funktion darstellt und darüber hinaus im allgemeinen singularär ist, und

<sup>32</sup>Man beachte, daß es sich hierbei wegen  $\mu_{x_a, x_b}^{[t_a, t_b]}(\mathcal{C}_{x_a, x_b}^{[t_a, t_b]}) \neq 1$  nicht um ein Wahrscheinlichkeitsmaß handelt.

andererseits, weil ein Grenzwert in Gestalt eines Produkts aus einem Kontinuum von Kopien des translationsinvarianten Lebesgue-Maßes selbst translationsinvariant sein sollte und das Wiener-Maß diese Eigenschaft nicht aufweist. Wir verschieben die weitere Diskussion dieses Sachverhalts jedoch auf den übernächsten Abschnitt und bleiben zunächst bei Pfadintegralen auf Brownschen Bewegungen.

Bevor nun die für diesen Abschnitt namensgebende Formel im Mittelpunkt steht, sei nebenbei und weniger aufgrund seiner sehr wohl vorhandenen mathematischen Bedeutung als vielmehr zur Unterhaltung noch ein weiterer Aspekt erwähnt. Für Brownsche Pfade  $\omega(t)$  definieren wir die mengenwertige Zufallsvariable

$$W_{\delta, [t_a, t_e]}(\omega) := \{x \mid \|x - \omega(t)\| \leq \delta \text{ für irgendein } t \in [t_a, t_e]\}.$$

$W_\delta(t) \equiv W_{\delta, [0, t]}$  heißt dann *Wiener Würstchen* [125], eine Bezeichnung, die sich durch die geometrische Gestalt solcher Mengen geradezu aufdrängt. Wiener Würstchen sind unter anderem bei der Beschreibung der Wärmeleitung über stochastische Prozesse oder der Bose-Einstein-Kondensation von Bedeutung<sup>33</sup>.

### 2.4.2 Beweis der Feynman-Kac-Formel

Die zentrale Idee für die erste mathematisch exakte Definition von zumindest so etwas ähnlichem wie einem Feynmanschen Pfadintegral besteht darin, eine Integraldarstellung für die Halbgruppe  $\{e^{-t\hat{H}}\}_{t \geq 0}$  statt für die unitäre Gruppe  $\{e^{it\hat{H}}\}_{t \in \mathbb{R}}$  zu konstruieren<sup>34</sup>. Die wesentlichen Schwierigkeiten des heuristischen Feynmanschen Pfadintegrals liegen im Auftreten von Termen wie  $\exp[-i \int (\dot{\gamma}(t))^2 dt]$ , die divergieren, wenn  $\dot{\gamma}$  singular wird, sowie in der Tatsache, daß ein unendliches Produkt von Lebesgue-Maßen gewöhnlich selbst unendlich wird. Wird dagegen in Termen der Form  $\exp[-\int (\dot{\gamma}(t))^2 dt]$  der Integrand singular, dann werden diese zu Null, und das hebt sich im günstigsten Fall gegen die unendlich werdenden unendlichen Produkte von Lebesgue-Maßen gerade auf.

Die Verwendung von  $\{e^{-t\hat{H}}\}_{t \geq 0}$  wird zusätzlich auch durch die Möglichkeit nahegelegt, mit Hilfe der Brownschen Bewegung den physikalischen Prozeß der *Diffusion* zu modellieren. Die  $d$ -dimensionale *Diffusionsgleichung* lautet bekanntlich

$$\frac{\partial f}{\partial t} = D \nabla^2 f;$$

hier ist  $D$  die Diffusionskonstante und  $\nabla^2$  der  $d$ -dimensionale Laplace-Operator. Die Diffusionsgleichung ist formal identisch mit der *Wärmeleitungsgleichung*. Mit  $\hat{H}_0 = -\nabla^2$  und  $D = 1$  wird daraus

$$\frac{\partial f}{\partial t} = -\hat{H}_0 f,$$

und eine formale Integration des entsprechenden Anfangswertproblems liefert unmittelbar

$$f(x, t) = e^{-t\hat{H}_0} f(x, t_0).$$

<sup>33</sup>Siehe hierzu [480].

<sup>34</sup>Die Halbgruppe  $e^{-t\hat{H}}$  wird als *Schrödinger-Halbgruppe* bezeichnet.

Andererseits gilt<sup>35</sup>

$$f(x, t) = \frac{1}{(4\pi t)^{d/2}} \int_{\mathbb{R}^d} e^{-\|x-x'\|^2/4\pi t} f(x', t_0) dx',$$

was den Zusammenhang zwischen der Integraldarstellung der Halbgruppe  $\{e^{-t\hat{H}}\}_{t \geq 0}$  mit Gaußschen Prozessen zumindest für den freien Fall schon einmal offensichtlich werden läßt. Es gilt folglich nun, dies auf  $\hat{H} = \nabla^2 + V$  zu verallgemeinern oder anders formuliert, eine Integraldarstellung für Lösungen der Wärmeleitungsgleichung mit externem Kühlterm

$$\frac{\partial f}{\partial t} = (\nabla^2 - V) f \quad (2.30)$$

anzugeben. Genau das leistet die Feynman-Kac-Formel, die wir hier in einer sehr allgemeinen, von Simon in [480] angegebenen und bewiesenen Form vorstellen<sup>36</sup>. Dabei schreiben wir  $C_\infty(\mathbb{R}^d)$  für den Raum der stetigen Funktionen auf  $\mathbb{R}^d$ , die bei  $\infty$  verschwinden und  $(\cdot, \cdot)$  für das Skalarprodukt im  $\mathcal{L}^2(\mathbb{R}^d)$ <sup>37</sup>.

**Satz:** (Feynman-Kac-Formel) *Es sei  $V \in C_\infty(\mathbb{R}^d)$  und  $\hat{H} = -\nabla^2 + V$ , außerdem sei  $\mathbf{m}$  das Lebesgue-Maß auf  $\mathbb{R}^d$  und  $\mu_0$  das  $d$ -dimensionale Wiener-Maß auf  $\mathcal{C}$ . Für  $x \in \mathbb{R}^d$  und  $\omega \in \mathcal{C}$  sei  $\omega \equiv x + \gamma$ . Dann gilt für alle  $f, g \in \mathcal{L}^2(\mathbb{R}^d)$*

$$(f, e^{-t\hat{H}} g) = \int_{\mathbb{R}^d \times \mathcal{C}} f^*(\omega(0)) g(\omega(t)) \exp \left[ - \int_0^t V(\omega(t')) dt' \right] d(\mathbf{m} \otimes \mu_0).$$

BEWEIS: Da  $\hat{H}$  wesentlich selbstadjungiert ist, können wir die Trotter-Produktformel anwenden und erhalten damit zunächst

$$(f, e^{-t\hat{H}} g) = \lim_{n \rightarrow \infty} (f, (e^{-t\hat{H}_0/n} e^{-tV/n})^n g).$$

Nun zerlegen wir das Intervall  $[0, t]$  in  $n$  äquidistante Teile mit den Grenzen  $s_j = tj/n$ ,  $j = 0, 1, \dots, n$ . Weil die Zuwachsfunktionen eines Wiener-Prozesses paarweise unabhängige Gaußsche Zufallsvariablen sind, finden wir

$$\int_{\mathbb{R}^d \times \mathcal{C}} f^*(\omega(0)) g(\omega(t)) \exp \left[ - \frac{t}{n} \sum_{j=0}^{n-1} V(\omega(tj/n)) \right] d(\mathbf{m} \otimes \mu_0)$$

<sup>35</sup>Siehe (2.25)

<sup>36</sup>Noch allgemeinere Versionen findet man für unbeschränkte Schrödinger-Halbgruppen in [479] und für ebensolche mit zusätzlichem Vektorpotential in [77]. Siehe auch [399].

<sup>37</sup>Die hier verfolgte, verbreitete Marschrichtung, die Feynman-Kac-Formel mit Hilfe der Trotter-Produktformel zu beweisen, ist nicht die einzige Möglichkeit, es existieren wie fast immer auch hierfür mehrere unterschiedliche Beweise. Als Beispiele seien die Variante erwähnt, das Resultat mit Hilfe stochastischer Integrale, hier speziell sogenannter *Itô-Integrale*, zu beweisen; siehe dazu [100], [101] sowie auch [480], oder aber diejenige, die Markov-Eigenschaft Brownscher Bewegungen sowie die Halbgruppeneigenschaft von  $\{e^{-t\hat{H}}\}_{t \geq 0}$  zu verwenden [363].

$$\begin{aligned}
&= \int_{\mathbb{R}^d \times \mathcal{C}} f^*(\omega(0)) g(\omega(t)) \exp \left[ -\frac{t}{n} V(\omega(0)) \right] \exp \left[ -\frac{t}{n} V(\omega(t/n)) \right] \dots \times \\
&\quad \times \exp \left[ -\frac{t}{n} V(\omega(t)) \right] d(\mathbf{m} \otimes \mu_0) \\
&= \int_{\mathbb{R}^d} f^*(x) g(x) \exp \left\{ -\frac{t}{n} [\hat{H}_0 + V(x)] \right\} \exp \left\{ -\frac{t}{n} [\hat{H}_0 + V(x)] \right\} \dots \times \\
&\quad \times \exp \left\{ -\frac{t}{n} [\hat{H}_0 + V(x)] \right\} dx \\
&= \int_{\mathbb{R}^d} f^*(x) g(x) \left\{ \exp \left\{ -\frac{t}{n} [\hat{H}_0 + V(x)] \right\} \right\}^n dx = (f, (e^{-t\hat{H}_0/n} e^{-tV/n})^n g),
\end{aligned}$$

also auch

$$(f, e^{-t\hat{H}} g) = \lim_{n \rightarrow \infty} \int_{\mathbb{R}^d \times \mathcal{C}} f^*(\omega(0)) g(\omega(t)) \exp \left[ -\frac{t}{n} \sum_{j=0}^{n-1} V(\omega(tj/n)) \right] d(\mathbf{m} \otimes \mu_0).$$

Alle  $\gamma \in \mathcal{C}$  sind stetig, folglich gilt für fast alle  $\omega$

$$\lim_{n \rightarrow \infty} \left[ -\frac{t}{n} \sum_{j=0}^{n-1} V(\omega(tj/n)) \right] = \int_0^t V(\omega(t')) dt'.$$

Wegen

$$\int_{\mathcal{C}} |f(\omega(0))| |g(\omega(t))| d\mu_0(\omega) = \int_{\mathbb{R}^d} |f(x)| e^{-t\hat{H}_0} |g(x)| dx = (|f|, e^{-t\hat{H}_0} |g|) < \infty$$

ist  $|f(\omega(0))| |g(\omega(t))| e^{t\|V\|_\infty}$  in  $\mathcal{L}^1(\mathbb{R}^d)$ ; hierfür gilt

$$\left| f^*(\omega(0)) g(\omega(t)) \exp \left[ -\frac{t}{n} \sum_{j=0}^{n-1} V(\omega(tj/n)) \right] \right| \leq |f(\omega(0))| |g(\omega(t))| e^{t\|V\|_\infty}$$

für alle  $n \in \mathbb{N}$  und alle  $x \in \mathbb{R}^d$ . Der Satz von der dominierten Konvergenz liefert damit die Behauptung.  $\square$

Die Feynman-Kac-Formel erlaubt unmittelbar eine Integraldarstellung der Schrödinger-Halbgruppe; man erreicht das durch Verwenden des bedingten Wiener-Maßes. Das führt auf folgendes

**Korollar:** *Unter obigen Voraussetzungen ist  $e^{-t\hat{H}}$  ein Integraloperator mit Integralkern*

$$e^{-t\hat{H}}(x, x') = \int \exp \left[ -\int_0^t V(\gamma(t')) dt' \right] d\mu_{x,x'}^{[0,t]}(\gamma).$$

Eine unmittelbare Konsequenz daraus ist die Möglichkeit, Lösungen der Wärmeleitungsgleichung als Wiener-Integrale zu formulieren: Ist  $f(x', 0)$  eine Anfangsverteilung, dann gilt für die Lösung der Wärmeleitungsgleichung (2.30)

$$\begin{aligned} f(x, t) &= \int_{\mathbb{Z}^d} \int f(\gamma(t), 0) \exp \left[ - \int_0^t V(\gamma(t')) dt' \right] d\mu_{x, x'}^{[0, t]}(\gamma) dx' \\ &= \int_{\mathbb{Z}^d \times \mathcal{C}} f(x' + \gamma(t), 0) \exp \left[ - \int_0^t V(x' + \gamma(t')) dt' \right] dx' \otimes d\mu_0. \end{aligned}$$

Obwohl sie auf den ersten Blick einfach eine formale Lösung der Diffusions- beziehungsweise Wärmeleitungsgleichung darstellt, geht die Bedeutung der Feynman-Kac-Formel weit darüber hinaus. Das betrifft einerseits die statistische Mechanik, beispielsweise bei der Beschreibung der zeitlichen Entwicklung von Zufallsprozessen, andererseits aber auch die Quantenmechanik und Quantenfeldtheorie. Mit Hilfe der Feynman-Kac-Formel lassen sich unter anderem der Grundzustand des Hamilton-Operators quantenmechanischer Systeme bestimmen sowie dessen Eindeutigkeit beweisen und sehr weitgehende nicht-störungstheoretische Aussagen über die Eigenfunktionen des Hamilton-Operators machen<sup>38</sup>. Außerdem ist sie eine der Grundlagen der konstruktiven Quantenfeldtheorie, was im nächsten Abschnitt kurz zur Sprache kommen wird. Zuvor aber geht es nun darum, wie die Feynman-Kac-Formel eine mathematisch exakte Definition feynmanscher Pfadintegrale ermöglicht.

## 2.4.3 Analytische Fortsetzung

### 2.4.3.1 Operatorwertige Funktionalintegrale

Die auffällige Ähnlichkeit der Schrödingergleichung

$$i \frac{\partial \psi}{\partial t} = (-\nabla^2 + V) \psi$$

mit der Wärmeleitungsgleichung

$$\frac{\partial f}{\partial t} = (\nabla^2 - V) f$$

drängt die Einführung einer imaginären Zeitvariablen  $\tau \equiv it$  geradezu auf, denn mit

$$\frac{\partial \psi}{\partial \tau} = \frac{\partial \psi}{\partial(it)} = -i \frac{\partial \psi}{\partial t}$$

gehen die beiden Gleichungen in der Tat ineinander über. Dieser auf den ersten Blick rein formale Sachverhalt ist auf den zweiten Blick nichts weniger als die höchst nichttriviale Aussage, daß was immer durch die Schrödingergleichung beschrieben wird, mathematisch äquivalent zu einem Diffusionsprozeß ist. Als eine Konsequenz daraus wird die in den

<sup>38</sup>Siehe zum Beispiel [193].

Abschnitten 2.2.2 und 2.3.1 beschriebene Feynman-Nelsonsche Formel (2.24) nach einer solchen Substitution plötzlich im Sinne eines Wiener-Integrals interpretierbar, und alle in Abschnitt 2.3.2 beschriebenen und in der Originalfassung unlösbaren Schwierigkeiten, die beim Versuch, (2.24) direkt als Pfadintegral aufzufassen, verschwinden zumindest aus mathematischer Sicht plötzlich fast von alleine<sup>39</sup>. Man nennt die Substitution von  $t$  durch  $it$  *Wick-Rotation*<sup>40</sup> und bringt damit zum Ausdruck, daß dieselbe aufgrund der wohlbekannten Relation  $e^{i\pi/2} = i$  geometrisch in Form einer Rotation der reellen Zeitachse um  $\pi/2$  auf die imaginäre Achse interpretierbar ist.

Natürlich stellt sich bei der Einführung einer imaginären Zeitvariablen die Frage, ob man so etwas überhaupt darf, und ganz ohne weiteres darf man es auch nicht. Genauer gesagt ist das zwar ein mathematisch wohldefinierter Vorgang, der jedoch aus physikalischer Sicht völlig vom betrachteten Gegenstand wegführt; die Schrödingergleichung mit imaginärer Zeitvariable ist zunächst einmal *nicht quantenmechanisch interpretierbar*. Das heißt jedoch keineswegs, daß die Wick-Rotation in der Quantenmechanik ohne Bedeutung wäre; das Gegenteil ist der Fall. Wir kommen darauf wieder zurück, vorerst aber konzentrieren wir uns auf den dazu erforderlichen mathematischen Hintergrund.

Es sei  $\mathcal{C}_{[0,t]}(\mathbb{R}^d) \equiv \mathcal{C}_t$  die Menge der stetigen  $\mathbb{R}^d$ -wertigen Funktionen auf  $[0, t]$ , ausserdem seien  $f \in \mathcal{L}^2(\mathbb{R}^d)$  und  $x \in \mathbb{R}^d$ . Für Funktionen  $F : \mathcal{C}_t \rightarrow \mathbb{R}$  betrachten wir Ausdrücke der Form

$$(J_t(F) f)(x) = \int_{\mathcal{C}_t} F(\gamma + x) f(\gamma(t) + x) d\mu_0(\gamma). \quad (2.31)$$

Solche operatorwertige Funktionalintegrale stellen sinnvolle Ausdrücke dar, wenn dadurch lineare Transformationen  $J_t(F)$  auf  $\mathcal{L}^2(\mathbb{R}^d)$  definiert werden. Besitzt so ein  $J_t(F)$ , betrachtet als Funktion von  $t$ , darüberhinaus eine Fortsetzung auf eine analytische operatorwertige Funktion<sup>41</sup> auf  $\mathfrak{U}'_+ \equiv \{z \in \mathfrak{U}' \mid \Re z > 0\}$ , die dann notwendigerweise eindeutig ist, und ist diese zusätzlich auf  $\overline{\mathfrak{U}'}_+ \equiv \{z \in \mathfrak{U}' \mid \Re z \geq 0\}$  stark stetig<sup>42</sup>, dann sagen wir,  $J_z(F)$  existiert für alle  $z \in \overline{\mathfrak{U}'}_+$ .

### 2.4.3.2 Anwendung auf Feynman-Integrale

Bei den in der Feynman-Kac-Formel auftretenden Integralen handelt es sich um Wiener-Integrale und folglich um mathematisch wohldefinierte Objekte. Auf dem Weg des Formalismus der analytischen Fortsetzung läßt sich nun auch das heuristische Feynman-Integral in ein solches mathematisch wohldefiniertes Objekt verwandeln<sup>43</sup>.

<sup>39</sup>Die Ehre, diesen Zusammenhang entdeckt zu haben, gebührt ohne Einschränkung M. Kac, der dadurch unmittelbar auf die Feynman-Kac-Formel geführt wurde.

<sup>40</sup>Benannt nach dem italienischen Theoretiker Gian-Carlo Wick.

<sup>41</sup>Analytizität kann hier im Sinne der Operatornorm, der schwachen oder der starken Operator-Topologie auf dem Raum der linearen Operatoren auf  $\mathcal{L}^2(\mathbb{R}^d)$  verstanden werden; diese drei sind paarweise äquivalent, siehe dazu beispielsweise [247]. Weitergehende Informationen über analytische operatorwertige Funktionen findet man unter anderem ebenda und in [304].

<sup>42</sup>Das heißt, stetig im Sinne einer Funktion von  $z \in \overline{\mathfrak{U}'}_+$  und nicht nur als Element von  $\mathcal{L}^2(\mathbb{R}^d)$ .

<sup>43</sup>Diese Idee geht ursprünglich auf D. G. Babbitt [31] und J. Feldman [163] zurück; umfangreiche Erweiterungen stammen von Johnson und Lapidus [282], [283], [284].



**Satz:** Es sei  $\hat{H} = -\nabla^2 + V$  selbstadjungiert und nach unten beschränkt, außerdem sei  $F_V : \mathcal{C}_t \longrightarrow \mathbb{R}$  für alle  $t > 0$  definiert durch

$$F_V(\gamma) := \exp \left[ - \int_0^t V(\gamma(t')) dt' \right].$$

Dann existiert  $J_z(F_V)$  für alle  $z \in \overline{\mathfrak{U}}_+$  und ist analytisch für alle  $z \in \mathfrak{U}_+$ . Für  $t > 0$  ist  $J_t(F_V)$  dabei definiert durch

$$(J_t(F) f)(x) = \int_{\mathcal{C}_t} F_V(\gamma + x) f(\gamma(t) + x) d\mu_0(\gamma).$$

Für alle  $z \in \overline{\mathfrak{U}}_+$  gilt

$$J_z(F_V) = e^{-z\hat{H}}.$$

BEWEIS: Wir zeigen zunächst, daß die Funktion  $T(z) := e^{-z\hat{H}}$  wohldefiniert und für alle  $z \in \overline{\mathfrak{U}}_+$  stark stetig ist.  $\hat{H}$  ist nach unten beschränkt, folglich gilt für dessen Spektrum  $\inf \sigma(\hat{H}) < -\infty$ , und die Funktion  $f_z : \sigma(\hat{H}) \longrightarrow \mathfrak{U}^1$  mit  $f_z(u) := e^{-zu}$  ist auf  $\overline{\mathfrak{U}}_+$  beschränkt. Aus dem Spektralsatz folgt dann, daß  $T(z) := e^{-z\hat{H}}$  für alle  $z \in \overline{\mathfrak{U}}_+$  ein wohldefinierter linearer Operator auf  $\mathcal{L}^2(\mathbb{R}^d)$  ist. Nun sei  $\{z_n\}_{n \in \mathbb{N}}$  eine Folge in  $\overline{\mathfrak{U}}_+$  mit  $\lim_{n \rightarrow \infty} z_n = z$ . Daraus folgt  $\lim_{n \rightarrow \infty} f_{z_n}(u) = f_z(u)$  für alle  $u \in \sigma(\hat{H})$ . Die Folge  $\{\|f_{z_n}\|_\infty\}_{n \in \mathbb{N}}$  mit  $\|f_{z_n}\|_\infty = \sup\{|f_{z_n}(u)| \mid u \in \sigma(\hat{H})\}$  ist beschränkt; es folgt wieder aus dem Spektralsatz in der starken Operatortopologie  $\lim_{n \rightarrow \infty} f_{z_n}(\hat{H}) = f_z(\hat{H})$  und damit auch  $\lim_{n \rightarrow \infty} e^{-z_n \hat{H}} = e^{-z\hat{H}}$ .

Nun zeigen wir, daß  $T(z) := e^{-z\hat{H}}$  in  $\mathfrak{U}_+$  analytisch ist. Unter Verwendung der schwachen Operatortopologie ist das gleichbedeutend mit der Aussage,  $(e^{-z\hat{H}}\varphi, \psi)$  sei analytisch in  $\mathfrak{U}_+$  für alle  $\varphi, \psi \in \mathcal{L}^2(\mathbb{R}^d)$ . Dafür genügt es zu zeigen, daß  $(e^{-z\hat{H}}\psi, \psi)$  analytisch ist für alle  $\psi \in \mathcal{L}^2(\mathbb{R}^d)$  mit  $\|\psi\|_2 = 1$ . Es sei  $\mathfrak{P}_{\hat{H}}$  das Spektralmaß des selbstadjungierten Operators  $\hat{H}$ ; dieses definiert für alle Borelmengen  $\mathcal{B} \in \mathcal{B}(\mathbb{R})$  ein Wahrscheinlichkeitsmaß  $\mu_{\psi\psi}$  durch  $\mu_{\psi\psi} := (\psi, \mathfrak{P}_{\hat{H}}(\mathcal{B})\psi)$  und  $\mu_{\psi\psi}(\sigma(\hat{H})) = 1$ . Damit gilt

$$(e^{-z\hat{H}}\psi, \psi) = \int_{\mathbb{R}} e^{-zu} \mu_{\psi\psi}(du) = \int_{\inf \sigma(\hat{H})}^{\infty} e^{-zu} \mu_{\psi\psi}(du).$$

$e^{-z\hat{H}}$  ist stark stetig auf  $\overline{\mathfrak{U}}_+$ , folglich ist  $(e^{-z\hat{H}}\psi, \psi)$  stetig auf  $\overline{\mathfrak{U}}_+$ . Nun sei  $\Delta$  ein beliebiges Dreieck in  $\mathfrak{U}_+$ . Mit dem Satz von Fubini und dem Cauchyschen Integralsatz gilt dafür

$$\int_{\Delta} (e^{-z\hat{H}}\psi, \psi) dz = \int_{\Delta} \int_{\inf \sigma(\hat{H})}^{\infty} e^{-zu} \mu_{\psi\psi}(du) dz = \int_{\inf \sigma(\hat{H})}^{\infty} \int_{\Delta} e^{-zu} dz \mu_{\psi\psi}(du) = 0,$$

und mit dem Satz von Morera folgt hieraus, daß  $(e^{-z\hat{H}}\psi, \psi)$  analytisch ist.  $\square$

Die bis hierher diskutierten Pfadintegraldarstellungen der Schrödinger-Halbgruppe  $\{e^{-t\hat{H}}\}$  und der unitären Gruppe  $\{e^{-it\hat{H}}\}$  erweisen sich nun als Spezialfälle dieses Satzes. Dazu muß man sich lediglich auf geeignete Teilmengen von  $\overline{\mathfrak{U}}_+$  beschränken. Der Fall  $z = t > 0$  wurde bereits erwähnt, er führt genau auf die Feynman-Kac-Formel. Das ist allerdings insofern nichts neues, als die Existenz der dort auftretenden Wiener-Integrale bereits kraft ihrer Konstruktion über stochastische Prozesse gesichert ist. Interessanter ist folgendes

**Korollar:** *Es sei  $\hat{H}$  definiert wie oben.  $J_{it}(F_V)$  existiert für alle  $t \in \mathbb{R}$ , und es gilt*

$$J_{it}(F_V) = e^{-it\hat{H}}. \quad (2.32)$$

Für alle  $\varphi \in \mathcal{D}(\hat{H})$  ist  $\psi(x, t) = (J_{it}(F_V)\varphi)(x)$  die eindeutige Lösung der Differentialgleichung

$$i \frac{\partial \psi}{\partial t} = (-\nabla^2 + V) \psi \quad (2.33)$$

mit der Anfangsbedingung  $\psi(0, x) = \varphi$ .

(2.32) liefert exakt die der Feynman-Kac-Formel entsprechende Integraldarstellung für die unitären Gruppe  $\{e^{-it\hat{H}}\}$ , welche die gesamte Dynamik der Quantenmechanik beschreibt, und (2.33) ist natürlich nichts anderes als die Schrödingergleichung. Damit stellt  $J_{it}(F_V)$  eine mathematisch exakte analytische Version des Feynmanschen Pfadintegrals dar. Dieses Resultat läßt sich, zwar mit physikalischer Grobheit, dafür aber inhaltlich umso prägnanter, wie folgt formulieren: *Die zeitliche Entwicklung eines quantenmechanischen Objekts ist eine Brownsche Bewegung in imaginärer Zeit.* Dahinter verbirgt sich der bereits angedeutete Sachverhalt, daß die Dynamik, welche durch die Schrödingergleichung beschrieben wird, formal zu einem Diffusionsprozeß äquivalent ist<sup>44</sup>. Man bezeichnet diese Betrachtungsweise gelegentlich als *stochastische Interpretation der Quantenmechanik*.<sup>45</sup>

Die beschriebene Technik der analytischen Fortsetzung läßt sich weiter verallgemeinern. Man kann beispielsweise anstelle von selbstadjungierten Operatoren Sesquilinearformen auf  $\mathcal{L}^2(\mathbb{R}^d)$  verwenden und erwischt so eine wesentlich größere Klasse von Potentialen; insbesondere kann es dabei vorkommen, daß sich der Ausdruck  $\hat{H}_0 + V$  als Formen-Summe noch auf Definitionsmengen von sinnvoller Größe erklären läßt, wo das im Sinne einer Operatorsumme auf nichts vernünftiges mehr führt [284]. Eine noch weitergehende Verallgemeinerung stammt von Albeverio, Johnson und Ma und betrachtet

<sup>44</sup>Näheres dazu findet man sehr ausführlich in [380]. Der erste, der diese Idee hatte, war bereits 1931 Erwin Schrödinger [459]. Es gibt auch Verallgemeinerungen der Interpretation dynamischer quantenmechanischer Prozesse als Brownsche Bewegungen auf die relativistische Quantenmechanik, genauer gesagt auf die Dirac-Gleichung; siehe dazu beispielsweise [179], [180], [181]. Man beachte jedoch, daß der hier beschriebene Sachverhalt nicht mit dem Begriff der *quantenmechanischen Brownschen Bewegung* verwechselt werden darf, bei der man Verallgemeinerungen der Schrödingergleichung betrachtet, die zusätzliche stochastische Terme enthalten. Durch solche nichtlinearen Schrödingergleichungen beschriebene Systeme kommen unter anderem als Modelle zur Beschreibung von Dekohärenzprozessen zum Einsatz [75], [286], [337].

<sup>45</sup>Die Tatsache, daß in [2] von einer dubiosen „stochastischen Interpretation der Wellenfunktion“ zu lesen ist, hat hiermit nichts zu tun; es handelt sich dabei einmal mehr um ein typisches Beispiel fehlerhafter Verwendung von Fachsprache in Lehrplänen.

additive Funktionale auf dem Raum der Brownschen Bewegungen [11]. Eine ganz andere Variante wurde nach und nach von Gelfand und Yaglom [182], Cameron [89] und insbesondere Nelson [386] gefunden und weiterentwickelt<sup>46</sup>. Dabei wird die Masse als komplexe Variable betrachtet und eine analytischen Fortsetzung von Wiener-Integralen im Sinne von Funktionen von  $m$  konstruiert. Auch das führt auf mathematisch wohldefinierte Feynmansche Pfadintegrale.

Ein nicht unerheblicher Nachteil aller Pfadintegralzugänge über Wiener-Integrale sollte nicht unerwähnt bleiben. Da man sich hier generell von einer Integration mit der Gewichtung  $e^{-iS}$  verabschiedet, geht die Möglichkeit einer mathematisch strengen Rekonstruktion des Übergangs von der Quantenmechanik zur klassischen Physik vermöge eines Grenzprozesses mit  $\hbar \rightarrow 0$ , wie er in Abschnitt 2.2.4 angedeutet wurde, verloren. Wir werden im nächsten Abschnitt einen anderen mathematisch rigorosen Zugang kennenlernen, bei welchem das funktioniert; wenig überraschend hat dieser dafür andere Nachteile.

### 2.4.3.3 Euklidische Feldtheorie

Wir bleiben noch einen Moment bei der analytischen Fortsetzung in  $t$ , mit der sich nun zwei sehr unterschiedliche Möglichkeiten ergeben. Einerseits kann man mit dem im obigen Sinn konstruierten Feynman-Integral arbeiten; das bedeutet, daß man den dabei auftretenden Parameter  $t$  direkt als physikalische Zeit auffaßt. Andererseits kann man mit der Feststellung starten, daß Lösungen der Schrödingergleichung als analytische Fortsetzungen von Lösungen der Wärmeleitungs- oder Diffusionsgleichung auffaßbar sind. Letztere ist mathematisch als diejenige Differentialgleichung, welche den stochastischen Prozeß der Brownschen Bewegung beschreibt, wohlverstanden und ihre Dynamik über die Feynman-Kac-Formel vollständig erfaßt. Damit stellt sich die Frage, wie man die Mathematik der Wiener-Prozesse für die Quantenmechanik nutzbringend einsetzen kann, und wir haben weiter oben gesehen, wie das geht: Der Trick besteht darin, die erwähnte Substitution  $t \rightarrow it$  auf eine imaginäre Zeitvariable zu verwenden. Hierbei sind von vorneherein zwei Dinge zu beachten: Erstens kehrt sich die Interpretation der Zeitvariable nun exakt um, die rein imaginäre Zeitvariable in obigem Korollar wird als reelle physikalische Zeit  $t$  aufgefaßt, und die Ersetzung dieser durch  $it$  führt wie gesehen auf die Anwendbarkeit der Feynman-Kac-Formel, da hierdurch formal die unitäre Gruppe in die Schrödinger-Halbgruppe übergeht. Zweitens wird bei relativistischer Betrachtungsweise aus einem Koordinatensystem  $\{x^0, x^1, x^2, x^3\}$  mit Minkowskischer Metrik  $ds^2 = g_{\mu\nu} dx^\mu dx^\nu = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2$  durch den Übergang  $x^0 \rightarrow ix^0 = x^4$  ein Koordinatensystem  $\{x^1, x^2, x^3, x^4\}$  mit Euklidischer Metrik  $ds^2 = \delta_{\mu\nu} dx^\mu dx^\nu = (dx^1)^2 + (dx^2)^2 + (dx^3)^2 + (dx^4)^2$  und damit aus der Minkowski-Geometrie des Raums  $\mathfrak{M}_4$  die gewöhnliche Euklidische Geometrie des Raums  $\mathfrak{L}_4$ . Das hat natürlich nicht unwesentliche Konsequenzen; die scheinbare formale Kleinigkeit der Einführung der imaginären Koordinate  $x^4$  täuscht darüber hinweg, daß sich die Räume  $\mathfrak{M}_4$  und  $\mathfrak{L}_4$  stark unterscheiden. Insbesondere wird dabei aus der nichtkompakten

<sup>46</sup>Eine zusätzliche Erweiterung stammt von B. O. Haugsby [216]. Detaillierte Informationen hierzu und zu der Methode im allgemeinen findet man in [284].

Lorentz-Gruppe  $O(3, 1)$  die kompakte Drehgruppe  $SO(4)$ , außerdem werden aus Fourier-Transformationen Laplace-Transformationen.

Entsprechend wenig erstaunlich ist es, daß sich die relativistische Quantenfeldtheorie auf dem  $\mathbb{E}_4$  ganz anders benimmt, als man das von ihr im üblichen minkowskischen Fall gewohnt ist. Dieses Konzept wurde von Schwinger [467], [468] und Symanzik [503] eingeführt und ist unter der Bezeichnung *Euklidische Quantenfeldtheorie* bekannt geworden<sup>47</sup>. Dabei führt eine Verallgemeinerung der Feynman-Kac-Formel auf quantenfeldtheoretische Pfadintegrale, woran erkennbar ist, daß euklidische Quantenfelder als verallgemeinerte stochastische Prozesse aufgefaßt werden können<sup>48</sup>. Zusätzliches Gewicht erhält dieser Ansatz, wenn man die analytische Fortsetzung in der Zeitvariable in Form der Substitution  $\beta \rightarrow it$  oder ausführlich geschrieben  $1/k_B T \rightarrow it/\hbar$  vornimmt. Durch Einführung dieser imaginären Temperatur wird aus dem quantenmechanischen Zeitentwicklungsoperator  $e^{-it\hat{H}}$  der statistische Operator  $e^{-\beta\hat{H}}$  des thermodynamischen Gleichgewichts. Damit ergeben sich sehr weitgehende Anwendungsmöglichkeiten auch in der statistischen Mechanik<sup>49</sup>. Aus technischer Sicht profitiert man in der Euklidischen Quantenfeldtheorie wie oben von der Tatsache, daß aus den Feynman-Integralen der Minkowskischen Quantenfeldtheorie nun mathematisch wohldefinierte Wiener-Integrale werden. Ebenfalls von Vorteil ist, daß man es in letzterer wesentlich mit elliptischen anstelle von hyperbolischen Differentialgleichungen zu tun hat. Dem steht der Nachteil gegenüber, mit nicht unmittelbar physikalisch interpretierbaren Größen hantieren zu müssen. Die physikalische Interpretierbarkeit muß nach vermeintlich getaner Arbeit erst durch eine analytische Fortsetzung, die auf die reelle Achse zurückführt, wiederhergestellt werden.

Besondere Bedeutung erlangt die Euklidische Theorie im Zusammenhang mit der *konstruktiven Quantenfeldtheorie*, die den Versuch darstellt, Quantenfeldtheorien mathematisch exakt axiomatisch zu konstruieren und damit gleichzeitig deren Existenz zu beweisen<sup>50</sup>. Die Taktik besteht hierbei darin, Quantenfeldtheorien im euklidischen Bereich mit Hilfe von Wiener-Integralen zu formulieren und anschließend deren Existenz auch im Minkowskischen Bereich zu beweisen. Grundlage dafür sind die *Osterwalder-Schrader-Axiome*<sup>51</sup>, die festlegen, welche Voraussetzungen eine Euklidische Quantenfeldtheorie erfüllen muß, damit man daraus eine Minkowskische Quantenfeldtheorie konstruieren kann. Wie schwierig das ist, kann man daran erkennen, daß konkrete nichttriviale Beispiele bis-

<sup>47</sup>Einführende Texte hierzu sind unter anderem [436] und [540]; siehe auch [499] und [500].

<sup>48</sup>Näheres dazu steht in [436] und [503].

<sup>49</sup>Sehr ausführlich dazu [540].

<sup>50</sup>Wir befinden uns hier auf dem Gebiet der *axiomatischen Quantenfeldtheorie*. Dieses Forschungsprogramm hat sich seit den fünfziger Jahren des 20. Jahrhunderts zum Ziel gesetzt, die Quantenfeldtheorie in der umfassendsten Bedeutung dieses Begriffes aus möglichst allgemeinen Axiomen streng mathematisch abzuleiten. Die Motivation hierfür kommt aus der Feststellung, daß Quantenfeldtheorien, wie sie in der theoretischen Physik in vielfältiger Form zum Einsatz kommen, zwar einerseits teilweise unvorstellbar erfolgreich sind im Hinblick auf die Berechnung experimenteller Resultate, andererseits jedoch mathematisch in keiner Weise definiert sind und teilweise für meßbare Größen unendliche Werte liefern, die unter Verwendung trickreicher Verfahren weggemogelt werden müssen; eine formale Konsolidierung erscheint somit dringend erforderlich. Das Programm hat inzwischen beachtliche Erfolge erzielt, ist aber gleichwohl noch weit von einem Abschluß entfernt. Die Anfänge findet man in [496], die weitere Entwicklung, die unter den Bezeichnungen *Algebraische Quantenfeldtheorie* und *Local Quantum Physics* firmiert, in [142] und insbesondere in [16] und [208].

<sup>51</sup>Benannt nach Konrad Osterwalder und Robert Schrader, die dieselben 1973 veröffentlicht haben [401].

lang nur in zwei und in drei Raum-Zeit-Dimensionen konstruiert werden konnten<sup>52</sup>. In vier Raum-Zeit-Dimensionen gibt es bisher nur Modelle ohne Wechselwirkung. Die Konstruktion realistischer wechselwirkender Quantenfeldtheorien in vier Dimensionen und der Beweis ihrer mathematischen Existenz darf derzeit als eines der wichtigsten und gleichzeitig schwierigsten ungelösten mathematischen Probleme aufgefaßt werden<sup>53</sup>.

## 2.5 Fresnel-Integrale

Während bei den im vorigen Abschnitt beschriebenen Pfadintegral-Varianten zunächst der durch die unitäre Gruppe der Schrödinger-Dynamik verursachte oszillatorische Charakter der Integranden im Weg stand und es folglich im wesentlichen darum ging, diesen wegzutricksen, drehen die nun folgenden Varianten die Sache um und nutzen genau diese Eigenschaft, um auf andere Weise zu exakten Definitionen von Pfadintegralen zu gelangen. Damit kann man bei der unitären Gruppe  $\{e^{-it\hat{H}}\}$  bleiben und muß nicht auf die Schrödinger-Halbgruppe  $\{e^{-t\hat{H}}\}$  ausweichen. Wir beginnen mit den Fresnel-Integralen, einem Integral-Begriff, der auf der Banachalgebra der beschränkten komplexwertigen Maße lebt<sup>54</sup>.

### 2.5.1 Definition und Eigenschaften

Es sei  $\mu$  ein komplexwertiges Maß auf einem Maßraum  $\mathcal{E}$  mit  $\sigma$ -Algebra  $\mathcal{E} \subset \mathcal{P}(\mathcal{E})$ . Die *totale Variation*  $|\mu|$  von  $\mu$  ist ein positives Maß, das wie folgt definiert ist: Für  $A \subset \mathcal{E}$  setzt man

$$|\mu|(A) := \sup \sum_i |\mu(A_i)|;$$

dabei läuft das Supremum über alle Partitionen  $\bigcup A_i$  von  $A$  in meßbare Mengen  $A_i$ . Man beachte, daß  $|\mu|(A)$  und  $|\mu(A)|$  nicht übereinstimmen müssen. Die totale Variation eines Maßes  $\mu$  ermöglicht eine polare Darstellung desselben analog zu den komplexen Zahlen,

$$d\mu = f d|\mu|$$

mit  $|f| = 1$ .

<sup>52</sup>Protagonisten auf diesem Gebiet sind hauptsächlich James Glimm und Arthur Jaffe. Ausführliches dazu steht in [193].

<sup>53</sup>Die Aufgabe, den in der Elementarteilchenphysik fundamentalen Yang-Mills-Theorien die Gestalt mathematisch rigoros konstruierter Quantenfeldtheorien zu geben, wurde vom Clay Mathematics Institute in Cambridge, Massachusetts in die Liste der sieben als besonders bedeutend eingeschätzten mathematischen Probleme aufgenommen; dabei wurde für deren Lösung ein Preis von insgesamt sieben Millionen US-Dollar ausgesetzt, eine Million für jedes einzelne. Das einzige bisher gelöste Problem aus dieser Liste ist der Beweis der Poincaré-Vermutung. Diese lautet in verallgemeinerter Form wie folgt: Jede geschlossene  $n$ -dimensionale Mannigfaltigkeit mit dem Homotopietyp einer  $n$ -Sphäre ist zur  $n$ -Sphäre homöomorph. Der Beweis der Poincaré-Vermutung wurde 2002 von Grigori Perelman im Rahmen des Beweises eines allgemeineren Satzes über die Geometrisierung von 3-Mannigfaltigkeiten erbracht.

<sup>54</sup>Details und viele weitere Informationen dazu findet man in [10] und [284]. Die Bezeichnung *Fresnel-Integrale* stammt von Albeverio und Høegh-Krohn und wurde in Anlehnung an die in der klassischen Optik auftretenden Integrale gleichen Namens gewählt. Das sind Integrale der Form

$$\int_0^w \sin(\pi x^2/2) dx, \int_0^w \cos(\pi x^2/2) dx \text{ und } \int_0^w e^{i\pi x^2/2} dx. \text{ Siehe hierzu [487].}$$

Nun sei  $\mathcal{H}$  ein reeller separabler Hilbertraum mit Skalarprodukt  $(\cdot, \cdot)$  und  $\mathcal{M}(\mathcal{H})$  der Raum der beschränkten komplexwertigen Borel-Maße auf  $\mathcal{H}$ . Dieser Raum wird auf folgende Weise eine kommutative Banach-Algebra mit 1: Als Norm auf  $\mathcal{M}(\mathcal{H})$  setzen wir

$$\|\mu\| := |\mu|(\mathcal{H}),$$

Addition und skalare Multiplikation sind klar; die Multiplikation definieren wir als *Faltung*, das heißt, für Maße  $\mu, \nu \in \mathcal{M}(\mathcal{H})$  und eine Borelmenge  $B \in \mathcal{B}(\mathcal{H})$  sei

$$(\mu * \nu)(B) := \int_{\mathcal{H}} \mu(B - x) d\nu(x),$$

und das Einselement von  $\mathcal{M}(\mathcal{H})$  ist das Dirac-Maß  $\delta_0$ . Für beschränkte stetige Funktionen  $f$  auf  $\mathcal{H}$  gilt

$$\begin{aligned} \int_{\mathcal{H}} f(x) d(\mu * \nu)(x) &= \int_{\mathcal{H}} \int_{\mathcal{H}} f(x + y) d\mu(x) d\nu(y) \\ &= \int_{\mathcal{H}} \int_{\mathcal{H}} f(y + x) d\nu(x) d\mu(y) = \int_{\mathcal{H}} f(x) d(\nu * \mu)(x), \end{aligned}$$

und somit  $\mu * \nu = \nu * \mu$ , außerdem folgt aus

$$\int_{\mathcal{H}} f(x) d|\mu * \nu|(x) \leq \int_{\mathcal{H}} \int_{\mathcal{H}} f(x + y) d|\mu|(x) d|\nu|(y)$$

die Youngsche Ungleichung  $\|\mu * \nu\| \leq \|\mu\| \|\nu\|$ .

Das ist jedoch noch nicht der Raum, um den es hier in erster Linie geht; mit

$$f_{\mu}(x) = \int_{\mathcal{H}} e^{i\langle x, y \rangle} d\mu(y) \tag{2.34}$$

definieren wir die Fourier-Transformierte des Maßes  $\mu \in \mathcal{M}(\mathcal{H})$  und betrachten nun mit  $\mathcal{F}(\mathcal{H}) := \{f_{\mu} \mid \mu \in \mathcal{M}(\mathcal{H})\}$  den Raum aller Fourier-Transformierter auf  $\mathcal{M}(\mathcal{H})$ . Die Elemente dieses Raums sind jeweils auf ganz  $\mathcal{H}$  definiert und außerdem gleichmäßig stetig auf  $\mathcal{H}$ . Alle Eigenschaften von  $\mathcal{M}(\mathcal{H})$  übertragen sich auf  $\mathcal{F}(\mathcal{H})$ , so daß letzterer mit  $\|f_{\mu}\| := \|\mu\|$  für alle  $\mu \in \mathcal{M}(\mathcal{H})$  ebenfalls eine kommutative Banach-Algebra mit 1 ist. Offensichtlich ist die Fouriertransformation eine Bijektion zwischen  $\mathcal{F}(\mathcal{H})$  und  $\mathcal{M}(\mathcal{H})$ , und folglich sind diese beiden Algebren isometrisch isomorph zueinander.

Damit haben wir alle Voraussetzungen beisammen, um das angekündigte Fresnel-Integral auf  $\mathcal{H}$  zu definieren [9], [10]. Es soll von der allgemeinen Form

$$\mathcal{F}(f) = \widetilde{\int}_{\mathcal{H}} e^{i\|x\|^2/2} f(x) dx$$

sein, wobei die Schreibweise mit der Schlange darauf hinweisen soll, daß für Fresnel-Integrale die Gültigkeit zweier zusätzlicher spezieller Eigenschaften gefordert wird: Erstens sollen sie *normiert* sein im Sinne der Relation

$$\widetilde{\int}_{\mathcal{H}} e^{i\|x\|^2/2} dx = 1,$$



und zweitens *translationsinvariant*, es soll also

$$\widetilde{\int}_{\mathcal{H}} e^{i\|x+a\|^2/2} dx = \widetilde{\int}_{\mathcal{H}} e^{i\|x\|^2/2} e^{i(x,a)} e^{i\|a\|^2/2} dx = 1$$

gelten für alle  $a \in \mathcal{H}$ . Diese beiden Eigenschaften lassen sich nun verwenden, um das gewünschte Integral für Funktionen aus  $\mathcal{F}(\mathcal{H})$  zu konstruieren. Dabei betrachtet man wie üblich ein Integral über in gewisser Weise „einfachen“ Funktionen und erweitert dieses dann auf den gesamten Funktionenraum. Aus der Translationsinvarianz folgt

$$\widetilde{\int}_{\mathcal{H}} e^{i\|x\|^2/2} e^{i(x,a)} dx = e^{-i\|a\|^2/2},$$

womit die Wirkung von  $\mathcal{F}$  auf Funktionen der einfachen Gestalt  $\varphi_a(x) = e^{i(x,a)}$  schon einmal klar wäre. Das läßt sich aufgrund der sinnvollerweise vorauszusetzenden Linearität des Integrals sofort auf die komplexe lineare Hülle der Menge  $\{\varphi_a \mid a \in \mathcal{H}\}$  erweitern; diese ist nicht nur der von diesen Funktionen erzeugte Vektorraum, sondern gleichzeitig auch die von ihnen erzeugte Algebra. Die Elemente dieser Algebra sind von der Form

$$\varphi(x) = \sum_n c_n \varphi_{a_n},$$

und dafür ergibt sich

$$\mathcal{F}(\varphi) = \widetilde{\int}_{\mathcal{H}} e^{i\|x\|^2/2} \sum_n c_n \varphi_{a_n}(x) dx = \sum_n c_n e^{-i\|a_n\|^2/2},$$

oder unter Verwendung der Delta-Funktion

$$\mathcal{F}(\varphi) = \int_{\mathcal{H}} e^{-i\|a_n\|^2/2} \sum_n c_n \delta(a - a_n) da.$$

Mit dem hier auftretenden Maß  $d\mu_\varphi(x) \equiv \sum_n c_n \delta(a - a_n) dx$  gilt

$$\varphi(x) = \int_{\mathcal{H}} e^{i(x,a)} d\mu_\varphi(a),$$

das heißt,  $\varphi(x)$  ist die Fourier-Transformierte von  $\mu_\varphi$ , und für dessen Integral folgt

$$\mathcal{F}(\varphi) = \widetilde{\int}_{\mathcal{H}} e^{i\|x\|^2/2} \varphi(x) dx = \int_{\mathcal{H}} e^{-i\|a\|^2/2} d\mu_\varphi(a).$$

Da die  $e$ -Funktion im Integranden eine beschränkte stetige Funktion ist, ist hier die rechte Seite nicht nur für  $\mu_\varphi$ , sondern für jedes beschränkte komplexwertige Maß wohldefiniert, und wir können  $\mathcal{F}$  für die Fourier-Transformierte eines jeden solchen Maßes bilden. Das führt auf folgende



**Definition:** Für alle  $f_\mu \in \mathcal{F}(\mathcal{H})$  heißt

$$\mathcal{F}(f_\mu) := \int_{\mathcal{H}} e^{-i\|x\|^2/2} f_\mu(x) dx = \int_{\mathcal{H}} e^{-i\|x\|^2/2} d\mu(x) \quad (2.35)$$

*Fresnel-Integral* von  $f_\mu$ .

$\mathcal{F}(\mathcal{H})$  nennt man auch den Raum der *Fresnel-integrierbaren Funktionen*<sup>55</sup>, für Fresnel-Integrale ist auch die Bezeichnung *normierte Integrale* üblich<sup>56</sup>. Man prüft leicht nach, daß Fresnel-Integrale stetige lineare normierte Abbildungen auf  $\mathcal{F}(\mathcal{H})$  sind. Ebenso einfach überzeugt man sich davon, daß  $\mathcal{F}(\mathcal{H})$  weitreichende Abschlußeigenschaften besitzt. Produkte, ganze Funktionen und Exponentialausdrücke von Fresnel-integrierbaren Funktionen sind jeweils selbst wieder Fresnel-integrierbar. Fresnel-Integrale besitzen im wesentlichen die Eigenschaften, die man auch von Lebesgue- und Lebesgue-Stieltjes-Integralen gewohnt ist; insbesondere gilt der Satz von Fubini<sup>57</sup>.

### 2.5.2 Die Feynman-Itô-Formel

Wir wenden uns nun wieder der Schrödingergleichung

$$i \frac{\partial \psi}{\partial t} = (-\nabla^2 + V) \psi$$

und deren Lösungen

$$\psi(x, t) = (e^{-it\hat{H}} \varphi)(x) \quad (2.36)$$

zu, mit  $\varphi \in \mathcal{L}(\mathbb{F}^d)$  als Anfangszustand, also der Beschreibung von  $N$  nichtrelativistischen Teilchen im  $\mathbb{F}^n$  mit  $nN = d$ . Der Bequemlichkeit halber normieren wir wieder die Massen aller Teilchen auf  $2m = \hbar$ , was durch geeignete Wahl der Koordinaten für Teilchen beliebiger, auch sich unterscheidender Massen stets möglich ist.

Feynman-Integrale versuchen im nichtrelativistischen Fall, eine Integraldarstellung von (2.36) zu liefern, bei der eine Integration mit der Gewichtung  $e^{iS}$  über alle geometrisch möglichen stetigen Wege der Teilchen erfolgt, die zum Zeitpunkt  $t$  auf den Punkt  $x$  führen;  $S$  ist dabei die klassische Wirkung des Systems. Unter Verwendung von Fresnel-Integralen gelang es K. Itô 1967, in mathematischer Strenge genau eine solche Integraldarstellung und damit eine weitere rigorose Darstellung Feynmanscher Integrale herzuleiten [267]<sup>58</sup>. Hierbei handelt es sich um die erste Variante, bei der im wesentlichen

<sup>55</sup>Der Raum  $\mathcal{F}(\mathcal{H})$  läßt sich auch direkt konstruieren. Man betrachtet dazu die Menge aller positiv definiten stetigen Funktionen  $f$  auf  $\mathcal{H}$ , welche die zusätzliche Eigenschaft haben, daß es für alle  $\varepsilon > 0$  einen nuklearen Operator  $\hat{N}_\varepsilon$  gibt, so daß für alle  $x \in \mathcal{H}$  aus  $(x, \hat{N}_\varepsilon) \leq 1$  stets  $\Re(f(0) - f(x)) < \varepsilon$  folgt.  $\mathcal{F}(\mathcal{H})$  ist die komplexe lineare Hülle dieser Funktionen. Genaueres steht in [335].

<sup>56</sup>Ist  $\mathcal{H}$  endlich dimensional und damit isomorph zum  $\mathbb{F}^d$ , so gilt

$$\int_{\mathcal{H}} e^{i\|x\|^2/2} f(x) dx = (2\pi i)^{-d/2} \int_{\mathcal{H}} e^{i\|x\|^2/2} f(x) dx.$$

<sup>57</sup>Beweis in [10]. Die klassische Variante für Lebesgue-Integrale erscheint erstmals in [177].

<sup>58</sup>Vergleiche auch [266]. Einen ähnlichen Ansatz findet man bei Maslov und Chebotarev [360], [361]; siehe hierzu auch [317].

die Original-Struktur (2.15) der Feynman-Integrale rekonstruiert wird. Gleichzeitig wird damit für die Schrödingergleichung genau das geliefert, was die Feynman-Kac-Formel für die Wärmeleitungsgleichung darstellt. Man nennt dieses Resultat *Feynman-Itô-Formel*<sup>59</sup>. Der hier beschriebene Beweis derselben folgt im wesentlichen [10]<sup>60</sup>.

**Satz:** (*Feynman-Itô-Formel*) Es seien  $V$  und  $\varphi$  die Fourier-Transformierten zweier komplexwertiger Maße im  $\mathbb{F}\mathfrak{R}^d$ , außerdem sei  $\mathcal{H}$  der reelle Hilbertraum der stetigen Kurven  $\gamma : [0, t] \rightarrow \mathbb{F}\mathfrak{R}^d$  so daß  $\gamma(0) = 0$  und  $d\gamma/dt' \in \mathcal{L}^2([0, t]; \mathbb{F}\mathfrak{R}^d)$ , mit Skalarprodukt

$$(\gamma, \xi) = \int_0^t \left\langle \frac{d\gamma}{dt'}, \frac{d\xi}{dt'} \right\rangle dt' = \int_0^t \sum_{j=1}^d \left( \frac{d\gamma}{dt'} \right)_j \left( \frac{d\xi}{dt'} \right)_j dt'$$

für  $\gamma, \xi \in \mathcal{H}$ . Dann gelten

(i)

$$f(\gamma) = \exp \left[ -i \int_0^t V(\gamma(t') + x) dt' \right] \varphi(\gamma(0) + x) \in \mathcal{F}(\mathcal{H}),$$

(ii)

$$\psi(x, t) = \int_{\mathcal{H}} e^{i|\gamma|^2/2} f(\gamma) d\gamma \quad (2.37)$$

ist Lösung der Schrödingergleichung

$$i \frac{\partial \psi}{\partial t} = (-\nabla^2 + V) \psi \quad (2.38)$$

mit der Anfangsbedingung  $\psi(x, 0) = \varphi(x)$ .

**BEWEIS:** Wir beginnen mit der Darstellung der formalen Lösung

$$\psi(x, t) = e^{-it\hat{H}} \varphi(x)$$

der Schrödingergleichung (2.38) als Dyson-Reihe. Dazu entwickeln wir  $e^{-it\hat{H}}$  nach Potenzen von  $V$ . Es gilt

$$\frac{d}{dt} (e^{it\hat{H}_0} e^{-it\hat{H}}) = i\hat{H}_0 e^{it\hat{H}_0} e^{-it\hat{H}} - i e^{it\hat{H}_0} \hat{H} e^{-it\hat{H}} = -i e^{it\hat{H}_0} V e^{-it\hat{H}}.$$

Formale Integration dieser Gleichung liefert

$$e^{it\hat{H}_0} e^{-it\hat{H}} = \mathbf{1} - i \int_0^t e^{it_1\hat{H}_0} V e^{-it_1\hat{H}} dt_1$$

<sup>59</sup>Diese Bezeichnung wurde von Albeverio und Høegh-Krohn vorgeschlagen.

<sup>60</sup>Vergleiche auch [154]. Itôs ursprünglicher Beweis läuft anders und findet sich in [267].

$$= \mathbf{1} - i \int_0^t e^{it_1 \hat{H}_0} V e^{-it_1 \hat{H}_0} e^{it_1 \hat{H}_0} e^{-it_1 \hat{H}} dt_1,$$

und durch Iteration erhält man daraus

$$\begin{aligned} e^{it\hat{H}_0} e^{-it\hat{H}} &= \sum_{n=0}^{\infty} (-i)^n \int_0^{t_{n-1}} \cdots \int_0^{t_1} \int_0^t e^{it_n \hat{H}_0} V e^{-i(t_n - t_{n-1}) \hat{H}_0} \cdots \times \\ &\quad \times e^{-i(t_3 - t_2) \hat{H}_0} V e^{-i(t_2 - t_1) \hat{H}_0} V e^{-it_1 \hat{H}_0} dt_1 dt_2 \cdots dt_n, \end{aligned}$$

mit  $0 \leq t_1 \leq t_2 \leq \cdots \leq t_n$ , oder, nach beidseitiger Multiplikation mit  $-e^{it\hat{H}_0}$ ,

$$\begin{aligned} e^{-it\hat{H}} &= \sum_{n=0}^{\infty} (-i)^n \int_0^{t_{n-1}} \cdots \int_0^{t_1} \int_0^t e^{-i(t-t_n) \hat{H}_0} V e^{-i(t_n - t_{n-1}) \hat{H}_0} \cdots \times \\ &\quad \times e^{-i(t_3 - t_2) \hat{H}_0} V e^{-i(t_2 - t_1) \hat{H}_0} V e^{-it_1 \hat{H}_0} dt_1 dt_2 \cdots dt_n. \end{aligned} \quad (2.39)$$

Nach Voraussetzung sind das Potential  $V$  und die Anfangs-Wellenfunktion  $\varphi$  von der Form

$$V(x) = \int_{\mathbb{R}^d} e^{i\langle x, y \rangle} d\mu_V(y) \quad (2.40)$$

beziehungsweise

$$\varphi(x) = \int_{\mathbb{R}^d} e^{i\langle x, y \rangle} d\mu_\varphi(y), \quad (2.41)$$

dabei sind  $\mu_V$  und  $\mu_\varphi$  beschränkte komplexwertige Maße auf dem  $\mathbb{R}^d$ . Wegen  $\hat{H}_0 = -\nabla^2$  gilt außerdem

$$e^{-it\hat{H}_0} e^{i\langle x, y \rangle} = e^{-ity^2} e^{i\langle x, y \rangle}$$

im Sinne verallgemeinerter Eigenfunktionen, und damit finden wir nach Einsetzen von (2.40), (2.41) und (2.39) in (2.36)

$$\begin{aligned} \psi(x, t) &= \sum_{n=0}^{\infty} (-i)^n \int_0^{t_{n-1}} \cdots \int_0^{t_1} \int_{\mathcal{H}} \cdots \int_{\mathcal{H}} \exp \left\{ -i \left[ (t - t_n) \left( \sum_{j=0}^n y_j \right)^2 \right. \right. \\ &\quad \left. \left. + (t_n - t_{n-1}) \left( \sum_{j=0}^{n-1} y_j \right)^2 + \cdots + (t_2 - t_1) (y_0 + y_1)^2 + t_1 y_0^2 \right] \right\} \times \\ &\quad \times \exp \left[ i \left\langle \sum_{j=0}^n y_j, x \right\rangle \right] d\mu_\varphi(y_0) \prod_{j=1}^n d\mu_V(y_j) dt_j. \end{aligned}$$

Der Exponent der ersten  $e$ -Funktion ist nach Konstruktion bereits zeitgeordnet; damit und mit  $t_0 \equiv 0$  sowie  $t_{n+1} \equiv t$  ergibt sich folglich

$$\psi(x, t) = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_0^t \cdots \int_0^{t_1} \int_{\mathcal{H}} \cdots \int_{\mathcal{H}} \exp \left[ -i \sum_{i=0}^n (t_{i+1} - t_i) \left( \sum_{j=0}^i y_j \right)^2 \right] \times$$

$$\times \exp \left[ i \left\langle \sum_{j=0}^n y_j, x \right\rangle \right] d\mu_\varphi(y_0) \prod_{j=1}^n d\mu_V(y_j) dt_j.$$

Wie man mit Induktion nach  $n$  leicht nachweist, gilt für die Doppelsumme

$$\sum_{i=0}^n (t_{i+1} - t_i) \left( \sum_{j=0}^i y_j \right)^2 = \sum_{j=0}^n \sum_{k=0}^n (t - \max\{t_j, t_k\}) y_j y_k,$$

und so erhalten wir schließlich

$$\begin{aligned} \psi(x, t) &= \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_0^t \cdots \int_0^t \int_{\mathcal{H}} \cdots \int_{\mathcal{H}} \exp \left[ -i \sum_{j=0}^n \sum_{k=0}^n (t - \max\{t_j, t_k\}) y_j y_k \right] \times \\ &\quad \times \exp \left[ i \left\langle \sum_{j=0}^n y_j, x \right\rangle \right] d\mu_\varphi(y_0) \prod_{j=1}^n d\mu_V(y_j) dt_j. \end{aligned} \quad (2.42)$$

als Lösung der Schrödingergleichung.

Nun gilt es nachzuweisen, daß dieses Resultat auch durch das Fresnel-Integral (2.37) darstellbar ist. Dazu betrachten wir die auf  $\mathcal{H}$  definierten Funktionen

$$\varphi(\gamma(0) + x) = \int_{\mathbb{F}^d} e^{i\langle \gamma(0), y \rangle} e^{i\langle x, y \rangle} d\mu_\varphi(y) \quad (2.43)$$

und

$$\int_0^t V(\gamma(t') + x) dt' = \int_0^t \int_{\mathbb{F}^d} e^{i\langle \gamma(t'), y \rangle} e^{i\langle x, y \rangle} d\mu_V(y) dt'; \quad (2.44)$$

wir zeigen zunächst, daß diese Funktionen aus  $\mathcal{F}(\mathcal{H})$  sind.

Zu diesem Zweck führen wir auf  $\mathbb{F}^d$  das Maß  $\mu_{\varphi, x}$  ein durch

$$\mu_{\varphi, x}(A) = \int_A e^{i\langle x, y \rangle} d\mu_\varphi(y);$$

die Umkehrung vom Satz von Radon-Nykodym<sup>61</sup> garantiert, daß dieses absolut stetig bezüglich  $\mu_\varphi$  ist und damit insbesondere von derselben totalen Variation. Mit dessen Hilfe schreiben wir für (2.43)

$$\varphi(\gamma(0) + x) = \int_{\mathbb{F}^d} e^{i\langle \gamma(0), y \rangle} d\mu_{\varphi, x}(y). \quad (2.45)$$

Um das weiter umzuformen, verwenden wir das Skalarprodukt auf dem Raum  $\mathcal{H}$ . Definieren wir die Funktion  $G_\sigma(\tau) = (t - \max\{\sigma, \tau\})$ , so erhalten wir für alle  $\sigma \in [0, t]$  und alle  $x \in \mathcal{H}$  mit  $G_\sigma(t) y$  ebenfalls ein Element von  $\mathcal{H}$ . Dann gilt für alle  $\gamma \in \mathcal{H}$

$$(\gamma, G_0 y) = \int_0^t \left\langle \frac{d\gamma}{dt'}, \frac{d(G_0(t')) y}{dt'} \right\rangle dt' = - \int_\sigma^t \left\langle \frac{d\gamma}{dt'}, y \right\rangle dt'$$

<sup>61</sup>Siehe [395], [131].

$$= - \left\langle \int_{\sigma}^t \frac{d\gamma}{dt'} dt', y \right\rangle = \langle \gamma(0), y \rangle, \quad (2.46)$$

und (2.45) wird zu<sup>62</sup>

$$\varphi(\gamma(0) + x) = \int_{\mathbb{R}^d} e^{i\langle \gamma, G_0 y \rangle} d\mu_{\varphi, x}(y).$$

Nun definieren wir die Funktion  $\Gamma_{\tau} : \mathbb{R}^d \rightarrow \mathcal{H}$  durch  $\Gamma_{\tau}(x) := G_{\tau}x$  und außerdem  $\mu_{\varphi, \tau, x} : \mathcal{B}(\mathcal{H}) \rightarrow \mathfrak{L}^1$  durch  $\mu_{\varphi, \tau, x}(A) := \mu_{\varphi, x}(G_{\tau}^{-1}(A))$ . Weil  $\mu_{\varphi}$  ein komplexes Maß auf  $\mathbb{R}^d$  ist, ist  $\mu_{\varphi, \tau, x}$  gemäß dem Satz vom Bildmaß<sup>63</sup> ein solches auf  $\mathcal{H}$ . Wegen

$$\|\Gamma_{\tau}(x) - \Gamma_{\tau}(x')\| = |t - \tau| \|x - x'\|$$

für alle  $\tau \in [0, t]$  und alle  $x, x' \in \mathbb{R}^d$  ist  $\Gamma_{\tau}$  stetig, und damit ist  $\Gamma_{\tau}^{-1}(A)$  für alle offenen Teilmengen  $A$  von  $\mathcal{H}$  ebenfalls offen. Folglich ist  $\mu_{\varphi, \tau, x}$  sogar ein Borel-Maß, und der Satz vom Bildmaß liefert weiter

$$\varphi(\gamma(\tau) + x) = \int_{\mathcal{H}} e^{i\langle \gamma, \gamma' \rangle} d\mu_{\varphi, \tau, x}(\gamma'),$$

also insbesondere

$$\varphi(\gamma(0) + x) = \int_{\mathcal{H}} e^{i\langle \gamma, \gamma' \rangle} d\mu_{\varphi, 0, x}(\gamma'). \quad (2.47)$$

Analog erhalten wir mit dem Maß  $\mu_{V, x}$  aus (2.44)

$$\int_0^t V(\gamma(t') + x) dt' = \int_0^t \int_{\mathbb{R}^d} e^{i\langle \gamma(t'), y \rangle} d\mu_{V, x}(y) dt'.$$

Die Funktion  $g(\tau, x) = e^{i\langle \gamma(\tau), x \rangle}$  ist auf  $[0, t] \times \mathbb{R}^d$  stetig und beschränkt und somit integrierbar bezüglich des Maßes  $\mathbf{m} \otimes \mu_{V, x}$ , wobei  $\mathbf{m}$  das Lebesgue-Maß auf  $\mathbb{R}^d$  ist; mit dem Satz von Fubini<sup>64</sup> folgt somit

$$\int_0^t V(\gamma(t') + x) dt' = \int_{[0, t] \times \mathbb{R}^d} e^{i\langle \gamma(t'), y \rangle} d(\mathbf{m} \otimes \mu_{V, x})(t', y).$$

Diesesmal definieren wir die Funktion  $\Lambda : [0, t] \times \mathbb{R}^d \rightarrow \mathcal{H}$  durch  $\Lambda(\tau, x) := \gamma_{\tau}x$  und das Maß  $\bar{\mu}_{V, x} : \mathcal{B}(\mathcal{H}) \rightarrow \mathfrak{L}^1$  durch  $\bar{\mu}_{V, x}(A) := (\mathbf{m} \otimes \mu_{\varphi, x})(\gamma_{\tau}^{-1}(A))$ . Daß  $\Lambda$  stetig ist, folgt aus der Abschätzung

$$\|\Lambda(\tau, x) - \Lambda(\tau', x')\| \leq (\|x\| + \|x'\|) |\tau - \tau'| + (t - \max \tau, \tau') \|x - x'\|,$$

<sup>62</sup>Hierbei macht man sich eine besondere Eigenschaft von  $\mathcal{H}$  zunutze: Für alle  $\tau \in [0, t]$  ist das Funktional  $\gamma \mapsto \gamma(\tau)$  stetig, und nach dem *Satz von Aronszajn-Bergman* [19], [52] besitzt folglich der Hilbertraum  $\mathcal{H}$  einen eindeutig bestimmten reproduzierenden Kern; in der Tat ist letzterer genau durch  $G(\sigma, \tau)$  gegeben. (2.46) ist gerade die spezielle Eigenschaft des reproduzierenden Kerns von  $\mathcal{H}$ , durch welche dieser definiert wird. Näheres dazu findet man in [52].

<sup>63</sup>Siehe beispielsweise [131].

<sup>64</sup>Zum Satz von Fubini für komplexwertige Maße siehe ebenfalls [131].

damit ist  $\bar{\mu}_{V,x}$  ein beschränktes Borel-Maß auf  $\mathcal{H}$ , und der Satz vom Bildmaß liefert

$$\int_0^t V(\gamma(t') + x) dt' = \int_{\mathcal{H}} e^{i(\gamma,\gamma')} d\hat{\mu}_{V,x}(\gamma'). \quad (2.48)$$

Wie die Relationen (2.47) und (2.48) unmittelbar zeigen, sind die durch (2.43) und (2.44) definierten Funktionen Fourier-Transformierte von den beschränkten komplexwertigen Maßen  $\mu_{\varphi,0,x}$  beziehungsweise  $\hat{\mu}_{V,x}$  aus  $\mathcal{M}(\mathcal{H})$ , das heißt, sie sind aus  $\mathcal{F}(\mathcal{H})$ , und daher ist auch

$$f(\gamma) = \exp \left[ -i \int_0^t V(\gamma(t') + x) dt' \right] \varphi(\gamma(0) + x)$$

aus  $\mathcal{F}(\mathcal{H})$ . Somit ist das Integral

$$\psi(x, t) = \int_{\mathcal{H}} e^{i\|\gamma\|^2/2} f(\gamma) d\gamma \quad (2.49)$$

wohldefiniert.

Um zu zeigen, daß dieses Integral eine Lösung der Schrödingergleichung ist, berechnen wir es explizit und finden zunächst einmal

$$\begin{aligned} \psi(x, t) &= \int_{\mathcal{H}} e^{i\|\gamma\|^2/2} \exp \left[ -i \int_0^t V(\gamma(t') + x) dt' \right] \varphi(\gamma(0) + x) d\gamma \\ &= \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_{\mathcal{H}} e^{i\|\gamma\|^2/2} \left[ \int_0^t V(\gamma(t') + x) dt' \right]^n \varphi(\gamma(0) + x) d\gamma. \end{aligned}$$

Analog zu (2.47) und (2.48) erhalten wir daraus unter Verwendung der oben definierten Funktion  $G_{\sigma}(\tau)$  sowie des durch  $n+1$ -fache Faltung gebildeten Maßes  $\mu_{\varphi,0,x} * \bar{\mu}_{V,x} * \cdots * \bar{\mu}_{V,x} * \bar{\mu}_{V,x}$  weiter

$$\begin{aligned} \psi(x, t) &= \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_{\mathcal{H}} e^{-i\|\gamma\|^2/2} d(\mu_{\varphi,0,x} * \bar{\mu}_{V,x} * \cdots * \bar{\mu}_{V,x} * \bar{\mu}_{V,x})(\gamma) \\ &= \sum_{n=1}^{\infty} \frac{(-i)^n}{n!} \int_{\mathcal{H}} \cdots \int_{\mathcal{H}} \int_{\mathcal{H}} \exp \left( -\frac{i}{2} \left\| \sum_{j=1}^{n+1} \gamma_j \right\|^2 \right) d\mu_{\varphi,0,x}(\gamma_1) \prod_{j=2}^{n+1} d\bar{\mu}_{V,x}(\gamma_j) \\ &= \sum_{n=1}^{\infty} \frac{(-i)^n}{n!} \int_{\mathcal{H}} \cdots \int_{\mathcal{H}} \int_{\mathcal{H}} \exp \left[ -\frac{i}{2} \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} (\gamma_i, \gamma_j) \right] d\mu_{\varphi,0,x}(\gamma_1) \prod_{j=2}^{n+1} d\bar{\mu}_{V,x}(\gamma_j) \end{aligned}$$

$$\begin{aligned}
&= \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_0^t \cdots \int_0^t \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \\
&\quad \exp \left( -\frac{i}{2} \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} (G_{t_i} y_i, G_{t_j} y_j) + i \sum_{j=1}^{n+1} \langle x, y_j \rangle \right) d\mu_{\varphi, x}(y_1) \prod_{j=2}^{n+1} d\mu_{V, x}(y_j) dt_j \\
&= \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_0^t \cdots \int_0^t \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \exp \left( -\frac{i}{2} \left\| \sum_{j=1}^{n+1} G_{t_j} y_j \right\|^2 + i \sum_{j=1}^{n+1} \langle x, y_j \rangle \right) \times \\
&\quad \times d\mu_{\varphi, x}(y_1) \prod_{j=2}^{n+1} d\mu_{V, x}(y_j) dt_j;
\end{aligned}$$

hierfür können wir unter Verwendung der Definitionen von  $G_{\sigma}(\tau)$  und der Maße  $\mu_{\varphi, x}$  und  $\mu_{V, x}$  auch

$$\begin{aligned}
\psi(x, t) &= \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_0^t \cdots \int_0^t \int_{\mathcal{H}} \cdots \int_{\mathcal{H}} \exp \left[ -i \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} (t - \max\{t_i, t_j\}) y_i y_j \right] \times \\
&\quad \times \exp \left[ i \left\langle \sum_{j=1}^{n+1} y_j, x \right\rangle \right] d\mu_{\varphi}(y_1) \prod_{j=2}^{n+1} d\mu_V(y_j) dt_j
\end{aligned}$$

schreiben, was identisch ist mit (2.42). Damit ist der Satz bewiesen.  $\square$

Die Analogie der Feynman-Itô-Formel zur Feynman-Kac-Formel wird noch deutlicher, wenn man die Fresnel-Integral-Darstellung (2.37) der Lösung der Schrödingergleichung in der ausführlichen Form

$$\psi(x, t) = \int_{\mathcal{H}} \widetilde{\exp} \left( i \int_0^t \left| \frac{d\gamma}{dt'} \right|^2 dt' \right) \exp \left[ -i \int_0^t V(\gamma(t')) dt' \right] \varphi(\gamma(0) + x) d\gamma$$

betrachtet, aber das ist noch nicht alles. Schaut man sich nämlich die beiden Exponenten genauer an, erkennt man darin die klassische Wirkung

$$S_{\text{kl}}(\gamma) = \int_0^t \left| \frac{d\gamma}{dt'} \right|^2 dt' - \int_0^t V(\gamma(t')) dt'$$

längs des Weges  $\gamma$  im Zeitintervall  $[0, t]$ . Verwendet man nun noch die Translationsinvarianz von Fresnel-Integralen, ersetzt  $\gamma$  durch  $\omega = \gamma - x$  und beachtet  $\omega(\tau) = (\gamma - x)(\tau) = \gamma(\tau) - x$ , erhält man

$$\psi(x, t) = \int_{\mathcal{H}} \widetilde{\exp} e^{iS_{\text{kl}}(\gamma)} \varphi(\gamma(0)) d\gamma$$



und damit eine präzise Rekonstruktion des Feynmanschen Pfadintegrals (2.15). Dabei ist das keinesfalls eine nur rein formale Ähnlichkeit, denn auch hier erfolgt die Integration über alle geometrisch möglichen stetigen Wege, die das beschriebene System im Konfigurationsraum klassisch beschreiten könnte. Aus dieser Perspektive ist die Interpretation der Feynman-Itô-Formel als exakte Formulierung heuristischer Feynman-Integrale wohlbegründet. Von einer vollständigen Übersetzung in die mathematische Sprache kann jedoch keine Rede sein; das verhindert die Bedingung

$$\int_0^t |\dot{\gamma}(t')|^2 dt' < \infty,$$

die in der Feynman-Itô-Formel an die Pfade zu stellen ist, über die integriert wird. Denn damit werden, anders als etwa bei Wiener-Integralen, stark nichtdifferenzierbare Pfade aus der Integration ausgeschlossen, wogegen der heuristische Zugang natürlich nahelegt, daß gerade solche Pfade wesentlich in die Integration einzubeziehen sind. Integration über alle *stetigen* Kurven zwischen zwei Punkten bedeutet, daß dabei *fast alle* Kurven, über die integriert wird, *nirgends differenzierbar* sind.

Damit soll jedoch die praktische Bedeutung dieses Zugangs nicht kleingeredet werden. Fresnel-Integrale erlauben eine Vielzahl von Anwendungen auf konkrete quantenmechanische Systeme, beispielsweise in der Streutheorie, im Zusammenhang mit harmonischen und anharmonischen Oszillatoren oder mit der abelschen wie auch nicht-abelschen Chern-Simons-Theorie, einer speziellen topologischen Quantenfeldtheorie, die unter anderem bei der Beschreibung des fraktionalen Quanten-Hall-Effekts zum Einsatz kommt<sup>65</sup>. Sie erlauben überdies eine mathematisch strenge Beschreibung des Übergangs von der Quantenmechanik zur klassischen Physik als Grenzübergang für  $\hbar \rightarrow 0$ . Mit Wiener-Integralen ist das wie erwähnt nicht möglich. Diesen Vorteilen stehen jedoch auch erhebliche Nachteile gegenüber, die mit der oben erwähnten Einschränkung der betrachteten Pfade und insbesondere mit der ziemlich eingeschränkten Auswahl an Potentialen und Anfangszuständen zu tun haben, für welche der Formalismus funktioniert. Die Potentiale dürfen kaum Singularitäten aufweisen, und da sie ebenso wie die Anfangszustände stets als Fourier-Transformierte komplexwertiger Maße vorliegen müssen, sieht man sich zusätzlich zu dieser weiteren Einschränkung auch noch mit der Schwierigkeit konfrontiert, jeweils geeignete Maße aufzutreiben und teilweise nichttriviale Konvergenzfragen klären zu müssen. Alle anderen Pfadintegral-Varianten sind hier sehr viel robuster.

### 2.5.3 Unendlichdimensionale oszillatorische Integrale

Wir betrachten noch kurz eine Verallgemeinerung, die unter anderem für die erwähnte Modellierung des klassischen Grenzfalles von Bedeutung ist. Fresnel-Integrale sind spezielle Beispiele für Integrale der Form  $\int e^{i\Phi(x)} f(x) dx$ ; solche Integrale nennt man in Anbetracht der komplexen  $e$ -Funktion im Integranden *oszillatorische Integrale*. Während Ausdrücke dieser Form im  $\mathbb{R}^d$  und für genügend glatte Funktionen seit langem wohlbekannt sind, ist die Situation für nicht Lebesgue-integrierbare Funktionen und insbesondere im unendlichdimensionalen Fall auf den ersten Blick weit weniger klar.

<sup>65</sup>Einen Überblick findet man in [10].

Ausgangspunkt für den endlichdimensionalen Fall ist die folgende, auf Hörmander zurückgehende Definition [252]. Gegeben seien eine Borel-Funktion  $f : \mathbb{R}^d \rightarrow \mathbb{C}$  sowie eine genügend glatte Funktion  $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}$ . Für  $\varepsilon > 0$  und alle Testfunktionen  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  mit  $\varphi(0) = 1$  sei außerdem

$$I_\varepsilon^\Phi(f, \varphi) := \int_{\mathbb{R}^d} e^{i\Phi(x)} f(x) \varphi(\varepsilon x) dx.$$

Solche Integrale existieren aufgrund der angenehmen Eigenschaften der schnell fallenden Funktionen im Lebesgueschen Sinn häufig auch dann, wenn  $f$  für sich nicht absolut integrierbar ist. Damit definiert man nun das oszillatorische Integral von  $f$  zur Phasenfunktion  $\Phi$  durch

$$\int_{\mathbb{R}^d} e^{i\Phi(x)} f(x) dx := \lim_{\varepsilon \rightarrow 0} I_\varepsilon^\Phi(f, \varphi).$$

Natürlich ist diese Definition nur sinnvoll, wenn sowohl das Integral  $I_\varepsilon$  für alle  $\varepsilon > 0$  als auch der Grenzwert  $\lim_{\varepsilon \rightarrow 0} I_\varepsilon$  existieren, und das jeweils auch unabhängig von  $\varphi$ <sup>66</sup>. Es ist bis jetzt noch nicht bekannt, wie die Räume der Funktionen  $f$  und der Phasenfunktionen  $\Phi$ , für welche solche oszillierenden Integrale existieren, im Detail aussehen.

Die Erweiterung auf den unendlichdimensionalen Fall stammt von Elworthy und Truman [141]<sup>67</sup>. Wir betrachten einen unendlichdimensionalen separablen Hilbertraum  $\mathcal{H}$ , dazu eine Funktion  $f : \mathcal{H} \rightarrow \mathbb{C}$  sowie eine Funktion  $\Phi : \mathcal{H} \rightarrow \mathbb{R}$  und außerdem die Menge  $\mathcal{P}^\infty(\mathcal{H})$  der Folgen  $\{\hat{P}_n\}_{n \in \mathbb{N}}$  von Projektionsoperatoren auf  $n$ -dimensionale Unterräume von  $\mathcal{H}$  mit  $\hat{P}_n \leq \hat{P}_{n+1}$  für alle  $n \in \mathbb{N}$  und dem starken Grenzwert  $\lim_{n \rightarrow \infty} \hat{P}_n = \mathbb{1}$ . Damit kann man im Sinne obiger Definition die normierten oszillatorischen Integrale

$$I_n^\Phi(f, \hat{P}_n) := \frac{1}{(2\pi i)^n} \int_{\hat{P}_n \mathcal{H}} e^{i\Phi(\hat{P}_n x)} f(\hat{P}_n x) d(\hat{P}_n x)$$

bilden. Nun definieren wir das *unendlichdimensionale oszillatorische Integral* von  $f$  zur Phasenfunktion  $\Phi$  durch

$$\widetilde{\int}_{\mathcal{H}} e^{i\Phi(x)} f(x) dx := \lim_{n \rightarrow \infty} I_n^\Phi(f, \hat{P}_n).$$

Wieder haben wir es nur dann mit einem sinnvollen Ausdruck zu tun, wenn sowohl  $\{I_n^\Phi(f, \hat{P}_n)\}_{n \in \mathbb{N}}$  für beliebige Folgen aus  $\mathcal{P}^\infty(\mathcal{H})$  als auch der Grenzwert  $\lim_{n \rightarrow \infty} I_n^\Phi$  existieren und letzterer unabhängig von  $\{\hat{P}_n\}_{n \in \mathbb{N}}$  ist. Die Schlange über dem Integral steht wieder

<sup>66</sup>Auch hier lassen sich die Integrale durch den Übergang

$$\int_{\mathbb{R}^d} e^{i\Phi(x)} f(x) dx \longrightarrow \frac{1}{(2\pi i)^d} \int_{\mathbb{R}^d} e^{i\Phi(x)} f(x) dx$$

normieren.

<sup>67</sup>Siehe auch [10] und [12].

für die Normierung. Auch hier ist eine genaue Beschreibung der Räume der Funktionen  $f$  und Phasenfunktionen  $\Phi$ , für welche solche Integrale existieren, noch unbekannt.

Unendlichdimensionale oszillatorische Integrale gestatten eine alternative Definition der Fresnel-Integrale des vorigen Abschnitts [10]. Hierzu setzt man speziell  $\Phi(x) \equiv \|x\|^2/2$  und definiert

$$\tilde{\mathcal{F}}_{\hat{P}_n}(f) = \int_{\hat{P}_n\mathcal{H}} e^{i\|\hat{P}_n x\|^2} f(\hat{P}_n x) d(\hat{P}_n x) \left[ \int_{\hat{P}_n\mathcal{H}} e^{i\|\hat{P}_n x\|^2} d(\hat{P}_n x) \right]^{-1}.$$

Sofern alle Integrale sowie der Grenzwert  $\tilde{\mathcal{F}}(f) := \lim_{n \rightarrow \infty} \tilde{\mathcal{F}}_{\hat{P}_n}(f)$  im obigen Sinne existieren, schreiben wir

$$\tilde{\mathcal{F}}(f) = \int_{\mathcal{H}} e^{i\|x\|^2/2} f(x) dx \quad (2.50)$$

und erhalten damit eine Definition für Integrale, die eine Verallgemeinerung der Fresnel-Integrale darstellen.  $\tilde{\mathcal{F}}(\mathcal{H})$  ist die Klasse der Funktionen, für die Integrale der Form (2.50) existieren. Man kann zeigen, daß die Klasse  $\mathcal{F}(\mathcal{H})$  der Fresnel-integrierbaren Funktionen eine echte Teilmenge von  $\tilde{\mathcal{F}}(\mathcal{H})$  ist<sup>68</sup>.

Besondere Bedeutung kommt den unendlichdimensionalen oszillatorischen Integralen unter anderem auch deshalb zu, weil sie eine spezielle Beschreibung des Übergangs von der Quantenmechanik zur klassischen Mechanik liefern können. Das wurde im Abschnitt 2.2.4 bereits kurz beschrieben, dort jedoch im Rahmen der heuristischen Feynman-Integrale und damit wie letztere auch ohne irgendwelche formale Grundlagen. Diese sogenannte Methode der stationären Phase läßt sich in mathematischer Strenge rekonstruieren, einerseits mit Hilfe von Fresnel-Integralen [8], [10], [431], andererseits mit unendlichdimensionalen oszillatorischen Integralen [4]<sup>69</sup>.

Ein weiterer, auf den ersten Blick ganz anderer Zugang, der eine Verbindung zwischen den in den Abschnitt 2.4.3 und dem vorliegenden Abschnitt beschriebenen Gegenständen bildet, sollte zusätzlich erwähnt werden. Es handelt sich um eine Variante, die von Cameron und Storvick vorgeschlagen wurde; diese hat starke Ähnlichkeit mit den in Abschnitt 2.4.3 beschriebenen operatorwertigen Funktionalintegralen [90], [91], [539]. Die dort beiläufig erwähnte Möglichkeit, die Masse als komplexe Variable aufzufassen und eine entsprechende analytische Fortsetzung vorzunehmen, ist ein Spezialfall dieses Zugangs. Eine Verallgemeinerung wurde von Johnson und Skoug gefunden [285]. Johnson wies später die Äquivalenz der Cameron-Storvick-Integrale zu den Fresnel-Integralen nach [281]. Im übrigen lassen sich auch die Fresnel-Integrale selbst weiter verallgemeinern [295], [296], [534].

<sup>68</sup>Siehe dazu [10].

<sup>69</sup>Vergleiche auch [307], wo ein anderer, auf verallgemeinerten kohärenten Zuständen aufbauender Ansatz verfolgt wird.

## 2.6 Weißes Rauschen

Eine weitere Technik, Feynmansche Pfadintegrale mathematisch streng zu definieren, verwendet das sogenannte *weiße Rauschen*. Darunter kann man sich anschaulich einen verallgemeinerten Gaußschen Prozeß vorstellen, der durch die zeitliche Ableitung eines Wiener-Prozesses gegeben ist und folglich die Geschwindigkeit einer Brownschen Bewegung beschreibt. Anders und ebenso unpräzise formuliert versteht man unter weißem Rauschen einen stochastischen Prozeß, der zu unterschiedlichen Zeiten unabhängig und außerdem gleichverteilt ist, mit verschwindendem Mittelwert und unendlicher Varianz<sup>70</sup>. Dieser Zugang, der gewisse Ähnlichkeiten zu demjenigen über Fresnel-Integrale aufweist, stammt ursprünglich von Streit und Hida [497] und wurde seither vielfach weiterentwickelt und verallgemeinert<sup>71</sup>. Bevor wir auf eine präzise Definition des weißen Rauschens und des darauf aufbauenden Pfadintegrals kommen, ist erst wieder einmal etwas Vorarbeit erforderlich.

### 2.6.1 Gelfandsche Raumtripel

Da Brownsche Pfade bekanntlich fast immer nirgends differenzierbar sind, ist es klar, daß die obige Charakterisierung des weißen Rauschens im üblichen Sinn nicht funktioniert. Man muß die dabei auftretenden zeitlichen Ableitungen allgemeiner auffassen und landet so automatisch bei *Distributionen*<sup>72</sup>. In diesem Zusammenhang von besonderem Interesse ist der Sachverhalt, daß übliche Distributionenräume wie  $\mathcal{D}'(\mathbb{R}^d)$ ,  $\mathcal{S}'(\mathbb{R}^d)$  oder  $\mathcal{E}'(\mathbb{R}^d)$ , wobei  $\mathcal{D}(\mathbb{R}^d)$ ,  $\mathcal{S}(\mathbb{R}^d)$  und  $\mathcal{E}(\mathbb{R}^d)$  die Räume der unendlich oft differenzierbaren Funktionen mit kompaktem Träger beziehungsweise der schnell fallenden Funktionen beziehungsweise aller unendlich oft differenzierbarer Funktionen auf  $\mathbb{R}^d$  sind<sup>73</sup>, jeweils als Bestandteile sogenannter *Gelfandscher Raumtripel* aufgefaßt werden können<sup>74</sup>, weswegen wir uns zunächst mit diesem Begriff beschäftigen.

**Definition:** Unter einem *Gelfandschen Raumtripel* verstehen wir ein Tripel  $\Phi \subset \mathcal{H} \subset \Phi'$ , wobei  $\mathcal{H}$  ein Hilbertraum ist,  $\Phi$  ist ein dichter Teilraum von  $\mathcal{H}$  mit einer lokalkonvexen Topologie, die feiner ist als diejenige, die von der Normtopologie auf  $\mathcal{H}$  stammt, und  $\Phi'$  ist der Raum der stetigen linearen Funktionale auf  $\Phi$ .

Die zweite Inklusion ist dabei sinnvoll, da gemäß dem *Rieszschen Darstellungssatz* der Hilbertraum  $\mathcal{H}$  zu seinem Dualraum  $\mathcal{H}'$  isomorph ist und folglich mit diesem identifiziert werden kann<sup>75</sup>. Gelfandsche Raumtripel stellen, wie die Inklusionen in der Definition zeigen, Erweiterungen von Hilberträumen in Form der Dualräume dichter Teilräume dieser Hilberträume dar, weswegen in der englischsprachigen Literatur die Bezeichnung *Rigged*

<sup>70</sup>Ausführliche Darstellungen dieses Gegenstands sind [242] und [333]; vergleiche auch [336]. Einen kurzen einführenden Überblick liefert [334].

<sup>71</sup>Siehe beispielsweise [161], [241], [332], [476], [521] beziehungsweise [206].

<sup>72</sup>Das Standardwerk überhaupt über Distributionen und damit zusammenhängende Dinge bilden die vier Bände von Gelfand und Shilov [183] - [186], in denen die gesamte Theorie entwickelt wird.

<sup>73</sup>Wegen  $\mathcal{D}(\mathbb{R}^d) \subset \mathcal{S}(\mathbb{R}^d) \subset \mathcal{E}(\mathbb{R}^d)$  gilt umgekehrt  $\mathcal{D}'(\mathbb{R}^d) \supset \mathcal{S}'(\mathbb{R}^d) \supset \mathcal{E}'(\mathbb{R}^d)$ .

<sup>74</sup>Die Theorie der Gelfandschen Raumtripel wurde, wie der Name zumindest zu einem Drittel schon andeutet, von Gelfand, Shilov und Wilenkin entwickelt [185], [186]. Vergleiche auch [529].

<sup>75</sup>Ausführliches zum Rieszschen Darstellungssatz samt Beweis findet man beispielsweise in [204].

*Hilbert Spaces* üblich ist<sup>76</sup>. Daß es sich bei Gelfandschen Raumtripeln wirklich um Inklusionen handelt, garantiert der folgende

**Satz:** (*Banach-Steinhaus*)  $\mathcal{E}$  und  $\mathcal{F}$  seien Banachräume,  $\mathcal{G}$  sei eine in  $\mathcal{E}$  dichte Menge. Eine Folge  $\{\hat{A}_n\}_{n \in \mathbb{N}}$  stetiger linearer Abbildungen von  $\mathcal{E}$  nach  $\mathcal{F}$  konvergiert genau dann punktweise gegen eine stetige lineare Abbildung  $\hat{A}$  von  $\mathcal{E}$  nach  $\mathcal{F}$ , wenn sowohl die Folge  $(\|\hat{A}_n\|)_{n \in \mathbb{N}}$  beschränkt ist, als auch die Folgen  $(\hat{A}_n x)_{n \in \mathbb{N}}$  für alle  $x \in \mathcal{E}$  konvergieren.

Ein Beweis dieses Satzes steht beispielsweise in [239]<sup>77</sup>. – Ist nun  $\Phi$  ein bezüglich irgend-einer Metrik  $d$  vollständiger, in Bezug auf die durch ein gemäß dieser Metrik stetiges Skalarprodukt  $\langle \cdot, \cdot \rangle$  induzierte Metrik jedoch nicht vollständiger Raum und der Hilbert-raum  $\mathcal{H}$  die zugehörige Vervollständigung von  $\Phi$ , dann wird offensichtlich durch jedes  $f \in \mathcal{H}$  vermöge dem Funktional  $l_f(g) := \langle f, g \rangle$  ein bezüglich  $\langle \cdot, \cdot \rangle$  stetiges lineares Funktional auf  $\Phi$  definiert.  $\Phi$  ist dicht in  $\mathcal{H}$  bezüglich der  $\langle \cdot, \cdot \rangle$ -Metrik, deshalb existiert in  $\Phi$  zu jedem  $f \in \mathcal{H}$  eine gegen  $f$  konvergente Folge  $(g_n)_{n \in \mathbb{N}}$ . Es gilt also einerseits  $\lim_{n \rightarrow \infty} \langle g_n, \chi \rangle = \langle f, \chi \rangle$  für alle  $\chi \in \Phi$ , und andererseits gibt es damit für alle  $\chi \in \Phi$  Konstanten  $C_\chi > 0$ , so daß  $l_{g_n}(\chi) < C_\chi$  gilt für alle  $n \in \mathbb{N}$ . Aus dem Satz von Banach-Steinhaus ergibt sich dann, daß alle  $l_f$  über  $\Phi$  stetig bezüglich  $d$  sind. Daraus folgt sofort  $\Phi \subset \mathcal{H} \subset \Phi'$ .

Wenn die Topologien der Teilräume der betrachteten Hilberträume durch Normen induziert werden, sind ihre in den zugehörigen Gelfandschen Raumtripeln auftretenden Dualräume *Distributionenräume*. Genau in diese Kategorie gehören die bereits erwähnten, hier interessierenden Beispiele  $\mathcal{D}(\mathbb{R}^d) \subset \mathcal{L}^2(\mathbb{R}^d) \subset \mathcal{D}'(\mathbb{R}^d)$  und  $\mathcal{S}(\mathbb{R}^d) \subset \mathcal{L}^2(\mathbb{R}^d) \subset \mathcal{S}'(\mathbb{R}^d)$ ; letzteres wird uns gleich wieder begegnen<sup>78</sup>.

<sup>76</sup>Genaugenommen versteht man unter einem Rigged Hilbert Space einen Tripel  $\Phi \subset \mathcal{H} \subset \Phi^\times$ , wobei  $\Phi^\times$  der Raum der stetigen antilinearen Funktionale auf  $\Phi$  ist. Da dieser jedoch isomorph zu  $\Phi'$  ist, ergibt sich so dieselbe Struktur. Siehe hierzu beispielsweise [62] und [63].

<sup>77</sup>Er ist benannt nach Stefan Banach und Hugo Steinhaus, die ihn 1927 entdeckten [41].

<sup>78</sup>Der Spezialfall der Distributionenräume liefert eine weitere und eigentlich die hauptsächliche Bedeutung Gelfandscher Raumtripel für die Quantenmechanik, da hierdurch auch Zustände wie Orts- oder Impulseigenfunktionen, die zum kontinuierlichen Spektrum selbstadjungierter Operatoren gehören und dadurch innerhalb üblicher, zum Raum  $\mathcal{L}^2$  isomorpher Zustandsräume nicht enthalten sind, erfaßt werden. Sie erweisen sich stattdessen als *Eigendistributionen*, die Elemente des Raums  $\Phi'$  eines geeigneten Tripels  $\Phi \subset \mathcal{L}^2(\mathbb{R}^d) \subset \Phi'$  sind. Man spricht hierbei auch von *verallgemeinerten Eigenfunktionen* und nennt die zugehörigen Elemente des kontinuierlichen Spektrums des betrachteten Operators entsprechend *verallgemeinerte Eigenwerte*. Damit werden solche Zustände im Rahmen des üblichen quantenmechanischen Formalismus handhabbar, wodurch auch ansonsten mathematisch dubiose Dinge wie zum Beispiel Wellenpakete streng definierbar werden. Insbesondere lassen sich auf diesem Wege die bekannten Schwierigkeiten, die im Zusammenhang mit unbeschränkten Operatoren auftreten und den Diracschen Bra-Ket-Formalismus unbrauchbar machen, elegant beseitigen. Betrachtet man die Quantenmechanik in ihrer Standardformulierung mit den Augen des mathematischen Physikers, stellt man im übrigen fest, daß ganz generell ein Großteil der dort auftauchenden Differentialgleichungen in Wirklichkeit Distributionen-Differentialgleichungen und die zugehörigen Lösungen schwache Lösungen sind. Näheres dazu in [62], [63], [204] und [362].

### 2.6.2 Analysis des weißen Rauschens

Als nächstes benötigen wir einen Zusammenhang zwischen komplexwertigen Funktionen und Maßen. Dieser wird durch den *Satz von Minlos* geliefert, für dessen Beweis wir zunächst zwei Definitionen benötigen<sup>79</sup>. Die erste verallgemeinert den Begriff der *Zylindermengen* auf Räume von linearen Funktionalen.

**Definition:**  $A$  sei eine beliebige Teilmenge des  $\mathbb{F}^d$ , außerdem sei  $\Phi$  ein topologischer Vektorraum,  $\{\varphi_1, \varphi_2, \dots, \varphi_d\}$  eine Familie von Elementen aus  $\Phi$  und  $f$  ein stetiges lineares Funktional aus dem Dualraum  $\Phi'$ . Dann heißt

$$\mathcal{Z}(\varphi_1, \varphi_2, \dots, \varphi_d; A) := \{f \in \Phi' \mid (f(\varphi_1), f(\varphi_2), \dots, f(\varphi_d)) \in A\}$$

die von  $\{\varphi_1, \varphi_2, \dots, \varphi_d\}$  und  $A$  definierte *Zylindermenge* in  $\Phi'$ .

$A$  nennt man auch die *Basis* der Zylindermenge. Die Menge  $\mathcal{Z}(\Phi')$  aller Zylindermengen in  $\Phi'$  ist eine Mengenalgebra und heißt auch *Zylinderalgebra*. – Die zweite Definition liefert einen speziellen Maßbegriff auf Familien von Zylindermengen über Borelmengen.

**Definition:** Eine skalare Funktion  $\mu$  über der Familie  $\mathcal{Z}(\Phi'; \mathcal{B}(\mathbb{F}^d))$  der Zylindermengen in  $\Phi'$  mit Basen aus der Borel algebra  $\mathcal{B}(\mathbb{F}^d)$  heißt *Zylindermaß*, wenn sie folgende Eigenschaften hat:

- (i)  $0 \leq \mu(\mathcal{Z}) \leq 1$  für alle  $\mathcal{Z} \in \mathcal{Z}(\Phi'; \mathcal{B}(\mathbb{F}^d))$ ,
- (ii)  $\mu(\Phi') = 1$ ,
- (iii) Für jede Folge  $\{\mathcal{Z}_n\}_{n \in \mathbb{N}}$  von paarweise disjunkten Zylindermengen aus  $\mathcal{Z}(\Phi'; \mathcal{B}(\mathbb{F}^d))$ , die von einer gemeinsamen Familie  $\{\varphi_1, \varphi_2, \dots, \varphi_d\} \subset \Phi$  definiert werden, gilt

$$\mu\left(\bigcup_{n=0}^{\infty} \mathcal{Z}_n\right) = \sum_{n=0}^{\infty} \mu(\mathcal{Z}_n),$$

- (iv) Ist  $\mathcal{Z}$  eine Zylindermenge aus  $\mathcal{Z}(\Phi'; \mathcal{B}(\mathbb{F}^d))$  und  $\mathcal{U}(\mathcal{Z})$  die Menge aller offenen Zylindermengen  $\mathcal{U}$  mit  $\mathcal{U} \supset \mathcal{Z}$ , dann gilt

$$\mu(\mathcal{Z}) = \inf\{\mu(\mathcal{U}) \mid \mathcal{U} \in \mathcal{U}(\mathcal{Z})\}.$$

Folgt zusätzlich für jede beschränkte stetige Funktion  $g : \mathbb{F}^d \rightarrow \mathbb{F}$  aus  $\lim_{n \rightarrow \infty} \varphi_{nj} = \varphi_j$  für  $j = 1, 2, \dots, d$  in  $\Phi'$

$$\lim_{n \rightarrow \infty} \int_{\Phi'} g(f(\varphi_{n1}), f(\varphi_{n2}), \dots, f(\varphi_{nd})) d\mu(f) = \int_{\Phi'} g(f(\varphi_1), f(\varphi_2), \dots, f(\varphi_d)) d\mu(f),$$

dann heißt das Zylindermaß  $\mu$  *stetig*. – Damit kommen wir zum ersten der beiden zentralen Sachverhalte dieses Abschnitts. Der folgende Satz liefert einem Mechanismus, mit

<sup>79</sup>Der Satz von Minlos verallgemeinert den Satz von Bochner vom  $\mathbb{F}^d$  auf unendlichdimensionale nukleare Räume. Letzterer findet sich zuerst in [59], ersterer in [367]; vergleiche auch [186] und [465]. Wir folgen hier im wesentlichen [280].



dem man sich unter sehr allgemeinen Bedingungen Maße in unendlichdimensionalen Räumen beschaffen kann.

**Satz:** (Minlos) Eine komplexe Funktion  $f$  auf einem reellen nuklearen Raum  $\Phi$  ist genau dann die Fourier-Transformierte eines  $\sigma$ -additiven positiven normierten Maßes auf dem Dualraum  $\Phi'$ , wenn folgendes gilt:

(i)  $f(0) = 1$ ,

(ii)  $f$  ist stetig,

(iii) für beliebige  $z_1, z_2, \dots, z_n \in \mathbb{C}$  und  $\varphi_1, \varphi_2, \dots, \varphi_n \in \Phi$  gilt

$$\sum_{j=1}^n \sum_{k=1}^n z_j z_k^* f(\varphi_j - \varphi_k) \geq 0.$$

BEWEIS: („ $\implies$ “) Zunächst sei  $\mu$  ein Zylindermaß auf  $\Phi'$  mit den oben erwähnten Eigenschaften und  $f$  dessen Fourier-Transformierte, das heißt

$$f(\varphi) = \int_{\Phi} e^{i(\psi, \varphi)} d\mu(\psi) \quad \text{für alle } \varphi \in \Phi.$$

Wir zeigen hierfür die Gültigkeit der Eigenschaften (i), (ii) und (iii). Zunächst gilt offensichtlich

$$f(0) = \int_{\Phi} d\mu = 1.$$

Definiert man weiter für  $x \in \mathbb{R}$ ,  $\varphi \in \Phi$  und  $\mathcal{C}_x = \{f \in \Phi' \mid (f, \varphi) \leq x\}$

$$\mu_{\varphi}(x) = \mu(\mathcal{C}_x) = \int_{\mathcal{C}_x} d\mu,$$

so erhält man

$$f(\varphi) = \int e^{ix} d\mu_{\varphi}(x).$$

Außerdem seien  $\{\varphi_n\}_{n \in \mathbb{N}}$  eine konvergente Folge in  $\Phi$  mit  $\lim_{n \rightarrow \infty} \varphi_n = \varphi$  und  $\mu_{\varphi_n}$ ,  $n \in \mathbb{N}$  die den  $\varphi_n$  im obigen Sinn zugeordneten Maße. Hierfür findet man

$$f(\varphi_n) = \int e^{ix} d\mu_{\varphi_n}(x) \quad \text{für } n \in \mathbb{N}$$

und folglich auch

$$\lim_{n \rightarrow \infty} f(\varphi_n) = \lim_{n \rightarrow \infty} \int e^{ix} d\mu_{\varphi_n}(x) = \int e^{ix} d\mu_{\varphi}(x) = f(\varphi),$$

das heißt  $f$  ist stetig.

Nun sei  $\mathcal{A} \subset \Phi$  die lineare Hülle einer Familie  $\{\varphi_1, \varphi_2, \dots, \varphi_d\}$  von Elementen aus  $\Phi$  und



$\mathcal{A}^\circ := \{f \in \Phi' \mid \forall \varphi \in \mathcal{A} \ f(\varphi) = 0\}$ . Das  $\mu$  entsprechende Maß auf  $\Phi'/\mathcal{A}^\circ$  heie  $\mu_{\mathcal{A}}$ . Damit erhlt man

$$\begin{aligned} \sum_{j=1}^n \sum_{k=1}^n z_j z_k^* f(\varphi_j - \varphi_k) &= \sum_{j=1}^n \sum_{k=1}^n z_j z_k^* \int_{\Phi} e^{i(\psi, \varphi_j - \varphi_k)} d\mu(\psi) \\ &= \sum_{j=1}^n \sum_{k=1}^n z_j z_k^* \int_{\Phi'/\mathcal{A}} e^{i(\psi, \varphi_j - \varphi_k)} d\mu_{\mathcal{A}}(\psi) \\ &= \int_{\Phi'/\mathcal{A}} \left| \sum_{j=1}^n z_j e^{i(\psi, \varphi_j)} \right|^2 d\mu_{\mathcal{A}}(\psi) \geq 0. \end{aligned}$$

(„ $\Leftarrow$ “) Jetzt sei  $f$  ein lineares Funktional aus  $\Phi'$  mit den Eigenschaften (i), (ii) und (iii). Wir zeigen, da  $f$  die Fourier-Transformierte eines  $\sigma$ -additiven positiven Maes auf  $\Phi'$  ist.

$\mathcal{A}$  sei ein endlichdimensionaler Unterraum von  $\Phi$  mit  $f_{\mathcal{A}} = f \upharpoonright \mathcal{A}$ . Gem dem Satz von Bochner ist  $f_{\mathcal{A}}$  die Fourier-Transformierte eines positiven Maes  $\mu_{\mathcal{A}}$  auf  $\mathcal{A}'$ ; dabei gilt  $\mathcal{A}' \cong \Phi'/\mathcal{A}^\circ$ . Durch ein Zylinderma  $\mu$  wird fr jeden Quotientenraum  $\Phi'/\mathcal{A}^\circ$  auf dessen Borelalgebra  $\mathcal{B}(\Phi'/\mathcal{A}^\circ)$  ein positives normiertes regulres Ma  $\nu_{\mathcal{A}}$  induziert, das man erhlt, indem man man fr jede Borelmenge  $\mathcal{B} \in \mathcal{B}(\Phi'/\mathcal{A}^\circ)$  eine Zylindermenge  $\mathcal{Z}$  mit Basis  $\mathcal{B}$  whlt und

$$\nu_{\mathcal{A}}(\mathcal{B}) = \mu(\mathcal{Z}). \quad (2.51)$$

setzt. Fr je zwei endlichdimensionale Unterrume  $\mathcal{A}_1$  und  $\mathcal{A}_2$  von  $\Phi$  mit  $\mathcal{A}_1 \subset \mathcal{A}_2$  sei auerdem  $\pi : \Phi'/\mathcal{A}_2^\circ \rightarrow \Phi'/\mathcal{A}_1^\circ$  die kanonische Einbettung der zugehrigen Quotientenrume; diese ist definiert durch  $\pi(f + \mathcal{A}_2^\circ) = f + \mathcal{A}_1^\circ$ . Aus  $\mathcal{A}_1 \subset \mathcal{A}_2$  folgt fr beliebige Borelmengen  $\mathcal{B} \in \mathcal{B}(\Phi'/\mathcal{A}^\circ)$

$$\nu_{\mathcal{A}_1}(\mathcal{B}) = \nu_{\mathcal{A}_2}(\pi^{-1}(\mathcal{B})) \quad (2.52)$$

Ist diese Relation fr je zwei Mae  $\nu_1$  und  $\nu_{\mathcal{A}_2}$  auf  $\Phi'/\mathcal{A}_1^\circ$  beziehungsweise  $\Phi'/\mathcal{A}_2^\circ$  und fr alle Borelmengen aus  $\mathcal{B}(\Phi'/\mathcal{A}^\circ)$  erfllt, so wird durch (2.51) eindeutig ein Zylinderma  $\mu$  auf  $\Phi'$  definiert<sup>80</sup>.

Nun sei  $\nu$  das durch  $\mu_{\mathcal{A}_2}$  auf  $\Phi'/\mathcal{A}_1^\circ$  induzierte Ma; dessen Fourier-Transformierte ist gegeben durch

$$g(\varphi) := \int_{\Phi'/\mathcal{A}_1^\circ} e^{i(\psi, \varphi)} d\nu(\psi).$$

Auerdem seien  $f_1 := f \upharpoonright \mathcal{A}_1$  und  $f_2 := f \upharpoonright \mathcal{A}_2$  die Fourier-Transformierten von  $\mu_{\mathcal{A}_1}$  beziehungsweise  $\mu_{\mathcal{A}_2}$ . Mit Hilfe von (2.52), angewandt auf  $\nu$  und  $\mu_{\mathcal{A}_2}$ , folgt daraus

$$g(\varphi) = \int_{\Phi'/\mathcal{A}_2^\circ} e^{i(\psi, \varphi)} d\mu_{\mathcal{A}_2}(\psi) = f_2(\varphi) = f \upharpoonright \mathcal{A}_2,$$

<sup>80</sup>Siehe zum Beispiel [186].

also  $g = f_2$  und daher auch  $\nu = \mu_1$ , da zwei Maße mit der gleichen Fourier-Transformierten auch selbst übereinstimmen. Das führt wiederum auf

$$\mu_1(\mathcal{B}) = \mu_2(\pi^{-1}(\mathcal{B}))$$

für alle  $\mathcal{B} \in \mathcal{B}(\Phi'/\mathcal{A}^\circ)$ , das heißt auf (2.52) für beliebige  $\mathcal{A}_1 \subset \mathcal{A}_2$  und deren zugehörige Maße  $\mu_1$  und  $\mu_2$ . Folglich gehört zu  $f$  ein eindeutig bestimmtes Zylindermaß  $\mu$  auf  $\Phi'$ . Als nächstes zeigen wir, daß dieses Zylindermaß stetig ist. Dazu sei wieder  $\{\varphi_n\}_{n \in \mathbb{N}}$  eine konvergente Folge in  $\Phi$  mit  $\lim_{n \rightarrow \infty} \varphi_n = \varphi$  und  $\mu_{\varphi_n}$ ,  $n \in \mathbb{N}$  beziehungsweise  $\mu$  die den  $\varphi_n$  beziehungsweise  $\varphi$  zugeordneten Maße. Für jedes  $\alpha \in \mathbb{R}$  gilt

$$f(\alpha\varphi_n) = \int e^{i\alpha x} d\mu_{\varphi_n}(x), \quad n \in \mathbb{N}$$

sowie

$$f(\alpha\varphi) = \int e^{i\alpha x} d\mu(x)$$

und aufgrund der Stetigkeit von  $f$  und der daraus folgenden Relation  $\lim_{n \rightarrow \infty} f(\varphi_n) = f(\varphi)$  weiter

$$\lim_{n \rightarrow \infty} \int e^{i\alpha x} d\mu_{\varphi_n}(x) = \int e^{i\alpha x} d\mu(x)$$

für alle  $\alpha \in \mathbb{R}$ . Damit konvergiert die Folge  $\{\mu_{\varphi_n}\}_{n \in \mathbb{N}}$  schwach gegen  $\mu$ <sup>81</sup>, und letzteres ist damit stetig.

Zusammengefaßt haben wir nun in Gestalt des Maßes  $\mu$ , dessen Fourier-Transformierte  $f$  ist, ein normiertes positiv definites stetiges Zylindermaß auf dem Dualraum  $\Phi'$  des nuklearen Raums  $\Phi$ . Ein solches Maß ist stets  $\sigma$ -additiv<sup>82</sup>, und damit ist der Satz bewiesen.  $\square$

Der Satz von Minlos erweist sich nun als wesentliches Hilfsmittel zur Konstruktion der mathematischen Umgebung, in der sich die Prozesse abspielen, die man als weißes Rauschen bezeichnet. Um das zu sehen, betrachten wir das Gelfandsche Raumtripler

$$\mathcal{S}(\mathbb{R}) \subset \mathcal{L}^2(\mathbb{R}) \subset \mathcal{S}'(\mathbb{R}),$$

bestehend aus dem Raum der schnell fallenden Testfunktionen, der quadratintegrablen Funktionen<sup>83</sup> und der temperierten Distributionen, jeweils über  $\mathbb{R}$ . Außerdem definieren

<sup>81</sup>Siehe beispielsweise [427].

<sup>82</sup>Siehe [186].

<sup>83</sup>Es sei daran erinnert, daß es sich bei den Räumen  $\mathcal{L}^2(\mathbb{E}^d)$  mit  $d \in \mathbb{N}^*$  geneaugenommen um Räume aus Äquivalenzklassen auf den eigentlichen Räumen  $L^2(\mathbb{E}^d)$  der quadratisch Lebesgue-integrierbaren Funktionen handelt, und zwar bezüglich der Äquivalenzrelation  $\sim$ , die durch

$$f \sim g \quad :\iff \quad \int_{\mathbb{R}^d} [f(x) - g(x)] dx = 0$$

definiert ist, das heißt, es gilt  $\mathcal{L}^2(\mathbb{E}^d) = L^2(\mathbb{E}^d)/\sim$ . Anschaulich wird  $L^2(\mathbb{E}^d)$  zu  $\mathcal{L}^2(\mathbb{E}^d)$ , indem man Funktionen, die sich jeweils nur auf einer Menge mit Lebesgue-Maß Null unterscheiden, als identisch

wir das Funktional

$$C(f) = e^{-\|f\|_2^2/2} = \exp \left[ \int_{-\infty}^{\infty} |f(x)|^2 dx \right], \quad f \in \mathcal{S}(\mathbb{R}).$$

Es gilt  $C(0) = 1$ , und  $C$  ist stetig und positiv definit, wie man sich leicht überzeugen kann. Gemäß dem Satz von Minlos gibt es daher ein eindeutig bestimmtes  $\sigma$ -additives positives normiertes Maß  $\mu$  auf  $\mathcal{S}'(\mathbb{R})$ , so daß

$$\int_{\mathcal{S}'(\mathbb{R})} e^{i\langle \varphi, f \rangle} d\mu(\varphi) = e^{-\|f\|_2^2/2}, \quad f \in \mathcal{S}(\mathbb{R})$$

gilt; hier ist  $\langle \cdot, \cdot \rangle$  die Erweiterung des Standard-Skalarprodukts von  $\mathcal{L}^2(\mathbb{R})$  auf  $\mathcal{S}'(\mathbb{R})$ . Zusammen mit der Menge  $\mathcal{B}(\mathcal{S}'(\mathbb{R}))$  der Borelmengen auf  $\mathcal{S}'(\mathbb{R})$  definiert  $\mu$  dank seiner Eigenschaften einen Wahrscheinlichkeitsraum  $(\mathcal{S}'(\mathbb{R}), \mathcal{B}(\mathcal{S}'(\mathbb{R})), \mu)$ ; man nennt  $C$  auch das *charakteristische Funktional* dieses Wahrscheinlichkeitsraums. Diesem Raum kann man eine anschauliche Interpretation geben [333]. Jede Funktion  $f \in \mathcal{S}(\mathbb{R})$  liefert mit  $\langle \cdot, f \rangle : \mathcal{S}'(\mathbb{R}) \rightarrow \mathbb{R}_0^+$  eine auf diesem Wahrscheinlichkeitsraum definierte Zufallsvariable, die auf ganz  $\mathcal{S}'(\mathbb{R})$  definiert und normalverteilt ist und die Varianz  $\|f\|_2^2$  hat. Das läßt sich auf  $\mathcal{L}^2(\mathcal{S}'(\mathbb{R}), \mathcal{B}(\mathcal{S}'(\mathbb{R})), \mu) \equiv \mathcal{L}^2(\mu)$  erweitern. Es sei  $g \in \mathcal{L}(\mathbb{R})$  und  $\{f_n\}_{n \in \mathbb{N}}$  eine Folge in  $\mathcal{S}(\mathbb{R})$  mit  $\lim_{n \rightarrow \infty} f_n = g$ . Diese liefert gleichzeitig eine Cauchy-Folge  $\{\langle \cdot, f_n \rangle\}_{n \in \mathbb{N}}$  in  $\mathcal{L}^2(\mu)$  mit dem Grenzwert  $\lim_{n \rightarrow \infty} \langle \cdot, f_n \rangle = \langle \cdot, g \rangle \in \mathcal{L}^2(\mu)$ . Die so definierte Zufallsvariable ist unabhängig von der verwendeten Folge und normalverteilt mit Erwartungswert 0 und Varianz  $\|g\|_2^2$ . Außerdem können wir mit dem durch

$$B_t(\varphi) = \langle \varphi, \chi_{[0,t)} \rangle = \int_0^t \varphi(t') dt', \quad \varphi \in \mathcal{S}'(\mathbb{R}), \quad (2.53)$$

definierten stochastischen Prozeß eine Brownsche Bewegung definieren;

$$\chi_{\mathcal{A}}(x) = \begin{cases} 1 & \text{falls } x \in \mathcal{A}, \\ 0 & \text{falls } x \notin \mathcal{A} \end{cases}$$

steht dabei für die charakteristische Funktion der im Index auftauchenden Menge. Ableiten nach der Zeit liefert so etwas wie  $\dot{B}_t = \varphi(t)$ , und damit haben wir all die Eigenschaften beisammen, die zu Beginn des vorliegenden Abschnitts als charakteristisch für

auffaßt. Der Grund für diese Definition liegt darin, daß mit dem Standard-Skalarprodukt

$$\langle f, g \rangle = \int_{\mathbb{R}^d} f^*(x) g(x) dx$$

$\mathcal{L}^2(\mathbb{R}^d)$  zu einem Hilbertraum wird,  $\mathcal{L}^2(\mathbb{R}^d)$  jedoch nur zu einem Prä-Hilbertraum. Da die Eigenschaft der Vollständigkeit im allgemeinen wesentlich für typische Anwendungsgebiete quadratintegrabler Funktionen ist, greift man stets zu diesem Trick.

weißes Rauschen beschrieben wurden, wo wir letzteres etwas ungenau als zeitliche Ableitung eines Wiener-Prozesses charakterisiert haben. Entsprechend können wir  $\mathcal{S}'(\mathbb{F}\mathfrak{i})$  als Raum der Pfade des weißen Rauschens auffassen. Das führt auf folgende

**Definition:** Der Wahrscheinlichkeitsraum  $(\mathcal{S}'(\mathbb{F}\mathfrak{i}), \mathcal{B}(\mathcal{S}'(\mathbb{F}\mathfrak{i})), \mu)$  heißt *Raum des weißen Rauschens*,  $\mu$  heißt *Standard-Gaußsches Maß* auf  $\mathcal{S}'(\mathbb{F}\mathfrak{i})$ .

Natürlich ist auch eine Verallgemeinerung auf  $d$ -dimensionales weißes Rauschen möglich [332]. Dazu betrachten wir den Raum  $\mathcal{L}_d^2 := \mathcal{L}^2(\mathbb{F}\mathfrak{i}) \otimes \mathbb{F}\mathfrak{i}^d$  für  $d \in \mathbb{N}$ . Dieser wird mit dem Skalarprodukt

$$\langle f, g \rangle = \sum_{j=1}^d \int_{-\infty}^{\infty} f^*(x) g(x) dx, \quad f, g \in \mathcal{L}_d^2$$

zu einem komplexen separablen Hilbertraum, auf dem in üblicher Weise eine Norm  $\| \cdot \|_2$  definiert werden kann.  $\mathcal{L}_d^2$  ist unitär äquivalent zum Raum  $\bigoplus_{j=1}^d \mathcal{L}^2(\mathbb{F}\mathfrak{i})$ , der direkten Summe

aus  $d$  Kopien des Raums  $\mathcal{L}^2(\mathbb{F}\mathfrak{i})$ . Die Elemente dieses Hilbertraums sind  $d$ -dimensionale Vektoren mit quadratisch Lebesgue-integrierbaren komplexen Funktionen als Einträgen. Entsprechend definieren wir den Raum  $\mathcal{S}_d = \mathcal{S}(\mathbb{F}\mathfrak{i}) \otimes \mathbb{F}\mathfrak{i}^d$  und dessen Dualraum  $\mathcal{S}'_d = \mathcal{S}'(\mathbb{F}\mathfrak{i}) \otimes \mathbb{F}\mathfrak{i}^d$ , womit wir das Gelfandschen Raumtripel

$$\mathcal{S}_d \subset \mathcal{L}_d^2 \subset \mathcal{S}'_d$$

konstruieren. Alles weitere verläuft völlig analog wie oben. Mit Hilfe des Satzes von Minlos konstruieren wir auf  $\mathcal{S}'_d$  das Maß  $\mu$  vermöge des charakteristischen Funktionals

$$\int_{\mathcal{S}'_d} e^{i\langle \varphi, f \rangle} d\mu(\varphi) = e^{-\|f\|_2^2/2}, \quad f \in \mathcal{S}_d$$

und damit den Wahrscheinlichkeitsraum  $(\mathcal{S}'_d, \mathcal{B}(\mathcal{S}'_d), \mu)$ , den *Raum des vektorwertigen weißen Rauschens*. Es lassen sich analog wie oben  $d$ -dimensionale normalverteilte Zufallsvariable auf  $\mathcal{L}_d^2(\mathcal{S}'_d, \mathcal{B}(\mathcal{S}'_d), \mu) \equiv \mathcal{L}_d^2(\mu)$  definieren,  $d$ -dimensionale Brownsche Bewegung erhalten wir in Form des stochastischen Prozesses

$$B_t(\varphi) = \langle \varphi, \chi_{[0,t]} \rangle \equiv (\langle \varphi, \chi_{[0,t]} \otimes \vec{e}_1 \rangle, \langle \varphi, \chi_{[0,t]} \otimes \vec{e}_2 \rangle, \dots, \langle \varphi, \chi_{[0,t]} \otimes \vec{e}_d \rangle), \quad \varphi \in \mathcal{S}'_d,$$

und wieder gilt  $\dot{B}_t = (\varphi(t) \otimes \vec{e}_1, \varphi(t) \otimes \vec{e}_2, \dots, \varphi(t) \otimes \vec{e}_d)$ , so daß wir die Elemente von  $\mathcal{S}'_d$  als Pfade des  $d$ -dimensionalen weißen Rauschens auffassen können.

Die soeben beschriebene Modellierung des weißen Rauschens stellt nur den einfachsten denkbaren Fall dar. Außerdem erweisen sich die dabei auftretenden Raumtripel in vielen Fällen der praktischen Anwendung als zu klein. Daher betrachten wir im folgenden Abschnitt einige Verallgemeinerungen.

### 2.6.3 Verallgemeinerte Funktionenräume

#### 2.6.3.1 Distributionen und Fock-Raum

Wir beschäftigen uns nun mit einer allgemeineren Definition des weißen Rauschens, die auf den unendlichdimensionalen Analogien der Tripel  $\mathcal{S}(\mathbb{R}) \subset \mathcal{L}^2(\mathbb{R}) \subset \mathcal{S}'(\mathbb{R})$  und  $\mathcal{S}_d \subset \mathcal{L}_d^2 \subset \mathcal{S}'_d$  aufbaut [321] - [324]<sup>84</sup>. Insbesondere wird es dadurch möglich, einen sehr viel größeren Raum verallgemeinerter Funktionen bereitzustellen.

Dazu sei  $\mathcal{E}$  ein reeller separabler nuklearer Raum und  $\{\|\cdot\|_p\}_{p \in \mathbb{N}}$  eine abzählbare aufsteigende Folge miteinander verträglicher Normen auf  $\mathcal{E}$ . Bezeichnet man für alle  $p \in \mathbb{N}$  mit  $\mathcal{H}_p$  jeweils die Vervollständigung von  $\mathcal{E}$  bezüglich der Norm  $\|\cdot\|_p$ , so gilt

$$\mathcal{E} = \bigcap_{p=0}^{\infty} \mathcal{H}_p,$$

das heißt,  $\mathcal{E}$  ist ein abzählbarer Hilbertraum. Diesen kann man mit der durch die von den Normen der Hilbertraumfamilie  $\{\mathcal{H}_p\}_{p \in \mathbb{N}}$  gegebenen Topologien induzierten projektiven Topologie zu einem topologischen Vektorraum machen. Weiter sei  $\mathcal{H} \equiv \mathcal{H}_0$  und  $\mathcal{E}'$  der Dualraum von  $\mathcal{E}$ ; damit liegt wiederum ein Gelfandsches Raumtripel

$$\mathcal{E} \subset \mathcal{H} \subset \mathcal{E}'$$

vor. Schließlich seien mit  $(\cdot, \cdot)_p$  die Skalarprodukte der Hilberträume  $\mathcal{H}_p$  für  $p \in \mathbb{N}$  und mit

$$\mathcal{H}_{p, \cdot} \equiv \{f = g + ih \mid g, h \in \mathcal{H}_p\}$$

die *Komplexifizierung* von  $\mathcal{H}_p$  bezeichnet. Auf dieser ist in natürlicher Weise ein Skalarprodukt definiert durch

$$(f_1, f_2)_{\mathcal{H}_{p, \cdot}} = (g_1 + ih_1, g_2 + ih_2)_{\mathcal{H}_{p, \cdot}} = (g_1, g_2)_p + (h_1, h_2)_p + i[(h_1, g_2)_p - (g_1, h_2)_p]$$

für  $g_1, g_2, h_1, h_2 \in \mathcal{H}$ ,  $p \in \mathbb{N}$ .

Definiert man auf  $\mathcal{E}$  das Funktional

$$C(f) = e^{-\|f\|_2^2/2} = \exp \left[ \int_{-\infty}^{\infty} |f(x)|^2 dx \right], \quad f \in \mathcal{E},$$

so erhält man wie im vorigen Abschnitt gemäß dem Satz von Minlos mit der Fourier-Transformierten von  $f$  ein Gaußsches Maß  $\mu$  auf  $\mathcal{E}'$ , das heißt, es gilt

$$e^{-\|f\|_2^2/2} = \int_{\mathcal{E}'} e^{i\langle g, f \rangle} d\mu(g) \quad \text{für alle } f \in \mathcal{E}.$$

Wir betrachten nun zwei lokalkonvexe Vektorräume  $\mathfrak{X}$  und  $\mathfrak{Y}$  sowie deren algebraisches Tensorprodukt  $\mathfrak{X} \otimes_{\text{alg}} \mathfrak{Y}$ ; außerdem seien  $\{\|\cdot\|_{\alpha}\}_{\alpha \in A}$  und  $\{\|\cdot\|_{\beta}\}_{\beta \in B}$  die definierenden Familien von Halbnormen für  $\mathfrak{X}$  und  $\mathfrak{Y}$ , mit geeigneten Indexmengen  $A$  und  $B$ . Auf  $\mathfrak{X} \otimes_{\text{alg}} \mathfrak{Y}$  definieren wir die  $\pi$ -*Topologie* durch die Halbnormen

$$\|x\|_{\alpha, \beta} = \inf \sum_j \|f_j\|_{\alpha} \|g_j\|_{\beta}, \quad x \in \mathfrak{X} \otimes_{\text{alg}} \mathfrak{Y}, \quad \alpha \in A, \quad \beta \in B,$$

<sup>84</sup>Vergleiche auch [110], [318], [398].

wobei das Infimum über alle endlichen Darstellungen  $x = \sum_j f_j \otimes g_j$  von  $x$  läuft, mit  $f_j \in \mathfrak{X}$  und  $g_j \in \mathfrak{Y}$ <sup>85</sup>. Die Vervollständigung von  $\mathfrak{X} \otimes_{\text{alg}} \mathfrak{Y}$  bezüglich der  $\pi$ -Topologie heißt  $\pi$ -Tensorprodukt und wird mit  $\mathfrak{X} \otimes_{\pi} \mathfrak{Y}$  bezeichnet. Sind  $\mathfrak{X}$  und  $\mathfrak{Y}$  nukleare Räume, dann ist auch  $\mathfrak{X} \otimes_{\pi} \mathfrak{Y}$  nuklear. Man beachte, daß für unendlichdimensionale Hilberträume das gewöhnliche Hilbertraum-Tensorprodukt, also die Vervollständigung des algebraischen Tensorprodukts der Hilberträume bezüglich dem kanonischen tensoriellen Skalarprodukt, und das  $\pi$ -Tensorprodukt nicht übereinstimmen. Da hier ausschließlich letzteres zum Einsatz kommt, schreiben wir im folgenden  $\mathfrak{X} \otimes \mathfrak{Y}$  anstelle von  $\mathfrak{X} \otimes_{\pi} \mathfrak{Y}$ . Sind  $f_1, f_2, \dots, f_n \in \mathfrak{X}$  und ist  $\mathfrak{S}_n$  die Gruppe der Permutationen von  $\{1, 2, \dots, n\}$ , dann heißt

$$f_1 \widehat{\otimes} f_2 \widehat{\otimes} \dots \widehat{\otimes} f_n := \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} f_{\sigma(1)} \otimes f_{\sigma(2)} \otimes \dots \otimes f_{\sigma(n)}$$

die *Symmetrisierung* von  $f_1 \otimes f_2 \otimes \dots \otimes f_n$ . Wir schreiben außerdem

$$f^{\otimes n} \equiv \underbrace{f \otimes \dots \otimes f}_{n \text{ mal}}$$

sowie

$$\mathfrak{X}^{\otimes n} \equiv \underbrace{\mathfrak{X} \otimes \dots \otimes \mathfrak{X}}_{n \text{ mal}}$$

und

$$\mathfrak{X}^{\widehat{\otimes} n} \equiv \overline{\text{span} \{ f_1 \widehat{\otimes} f_2 \widehat{\otimes} \dots \widehat{\otimes} f_n \mid f_1, f_2, \dots, f_n \in \mathfrak{X} \}} = \overline{\text{span} \{ f^{\otimes n} \mid f \in \mathfrak{X} \}};$$

dabei setzen wir  $\mathfrak{X}^{\otimes 0} = \mathfrak{X}^{\widehat{\otimes} 0} = \mathbb{K}$  beziehungsweise  $\mathfrak{X}^{\otimes 0} = \mathfrak{X}^{\widehat{\otimes} 0} = \mathbb{K}'$ , je nachdem, ob es sich bei  $\mathfrak{X}$  um einen reellen oder komplexen Raum handelt.  $\mathfrak{X}^{\widehat{\otimes} n}$  ist ein Hilbert-Unterraum von  $\mathfrak{X}^{\otimes n}$ ; für die jeweiligen kanonischen Normen und beliebige  $\varphi \in \mathfrak{X}^{\widehat{\otimes} n}$  gilt

$$\|\varphi\|_{\mathfrak{X}^{\widehat{\otimes} n}}^2 = n! \|\varphi\|_{\mathfrak{X}^{\otimes n}}^2.$$

Mit  $(\mathfrak{X}^{\widehat{\otimes} n})'$  bezeichnen wir den Dualraum von  $\mathfrak{X}^{\widehat{\otimes} n}$ ; dazu sei  $\langle \cdot, \cdot \rangle$  die kanonische Bilinearform auf  $(\mathfrak{X}^{\widehat{\otimes} n})' \times \mathfrak{X}^{\widehat{\otimes} n}$ .

Als *Fockraum*  $\mathcal{F}(\mathcal{H})$  über  $\mathcal{H}$  bezeichnet man die direkte Summe

$$\mathcal{F}(\mathcal{H}) = \bigoplus_{n=0}^{\infty} \mathcal{H}_n^{\otimes n}.$$

Seine Elemente sind Folgen der Form  $\{f^{(n)}\}_{n \in \mathbb{N}}$  mit  $f^{(n)} \in \mathcal{H}_n^{\otimes n}$ . Entsprechend definiert man den bosonischen oder symmetrischen Fock-Raum über  $\mathcal{H}$ ,

$$\Gamma(\mathcal{H}) := \bigoplus_{n=0}^{\infty} \mathcal{H}_n^{\widehat{\otimes} n},$$

<sup>85</sup>Die  $\pi$ -Topologie ist die stärkste lokalkonvexe Topologie, für welche die kanonische Abbildung  $\mathfrak{X} \times \mathfrak{Y} \longrightarrow \mathfrak{X} \otimes_{\text{alg}} \mathfrak{Y}$  stetig ist.

also als direkte Summe aller symmetrischer tensoriellen Potenzen von  $\mathcal{H}$ .<sup>86</sup> Der Fock-Raum selbst ist ein Hilbertraum. Das zuhörige Skalarprodukt erhalten wir folgendermaßen: Zunächst definieren wir

$$(f^{\otimes n}, g^{\otimes n})_{\mathcal{H}^{\otimes n}} = n! (f, g)_{\mathcal{H}}^n;$$

damit konstruieren wir für beliebige

$$\mathbf{f} = \{f^{(n)} \mid n \in \mathbb{N}\} \in \Gamma(\mathcal{H}), \quad \mathbf{g} = \{g^{(n)} \mid n \in \mathbb{N}\} \in \Gamma(\mathcal{H})$$

mit  $f^{(n)}, g^{(n)} \in \mathcal{H}^{\widehat{\otimes} n}$ ,  $n \in \mathbb{N}$  das Fockraum-Skalarprodukt durch

$$(\mathbf{f}, \mathbf{g})_{\Gamma(\mathcal{H})} := \sum_{n=0}^{\infty} n! (f^{(n)}, g^{(n)})_{\mathcal{H}^{\otimes n}}.$$

Der symmetrischen Fockraum erlaubt eine schön anschauliche Interpretation; es gilt nämlich  $\Gamma(\mathcal{H}) \cong \mathcal{L}^2(\mathcal{E}, \mu, \mathfrak{U})$ . Für den Raum  $\mathcal{L}^2(\mathcal{E}, \mu, \mathfrak{U})$  schreiben wir im folgenden der üblichen Konvention folgend  $(\mathcal{L}^2)$ <sup>87</sup>.

Für  $n, p \in \mathbb{N}$  seien  $\mathcal{H}_{:,p}^{\otimes n}$  und  $\mathcal{H}_{:,p}^{\widehat{\otimes} n}$  das  $n$ -fache beziehungsweise symmetrische  $n$ -fache Tensorprodukt von  $\mathcal{H}_p$ . Wie oben ist  $\mathcal{H}_{:,p}^{\widehat{\otimes} n}$  ein Hilbert-Unterraum von  $\mathcal{H}_{:,p}^{\otimes n}$ . Außerdem seien für  $q \in \mathbb{N}$  die Hilberträume, die man durch Vervollständigung der direkten Summe  $\bigoplus_{n=0}^q \mathcal{H}_{:,p}^{\widehat{\otimes} n}$  erhält, mit  $\Gamma_q(\mathcal{H}_p)$  bezeichnet. Ihre Skalarprodukte erhält man analog zu oben durch

$$(f^{\otimes n}, g^{\otimes n})_{\mathcal{H}_{:,p}^{\otimes n}} = 2^{nq} n! (f, g)_{\mathcal{H}_{:,p}}^n,$$

für ihre Dualräume gilt  $\Gamma_q(\mathcal{H}_p)' = \Gamma_{-q}(\mathcal{H}_{-p})$ . Ebenfalls analog zu oben definiert man nun

$$(\mathcal{E}) := \bigcap_{p,q=0}^{\infty} \Gamma_q(\mathcal{H}_p).$$

$\mathcal{E}$  wurde als nuklear vorausgesetzt; daraus folgt, daß  $(\mathcal{E})$  ebenfalls ein nuklearer Raum ist. Für den dazu dualen Raum  $(\mathcal{E})'$  erhalten wir<sup>88</sup>

$$(\mathcal{E})' = \bigcup_{p,q=0}^{\infty} \Gamma_q(\mathcal{H}_p)' = \bigcup_{p,q=0}^{\infty} \Gamma_{-q}(\mathcal{H}_{-p}),$$

womit wir das Gelfandsche Raumtripel

$$(\mathcal{E}) \subset \Gamma(\mathcal{H}) \subset (\mathcal{E})'$$

konstruiert haben. Die kanonische Bilinearform  $\langle\langle \cdot, \cdot \rangle\rangle$  auf  $(\mathcal{E})' \times (\mathcal{E})$  ist gegeben durch

$$\langle\langle \varphi, \mathbf{f} \rangle\rangle = \sum_{n=0}^{\infty} n! \langle \varphi^{(n)}, f^{(n)} \rangle.$$

<sup>86</sup>Der Fock-Raum ist insbesondere in der Quantenfeldtheorie von herausragender Bedeutung, weswegen Anhang C einige weitere Informationen dazu liefert.

<sup>87</sup>Man beachte auch hier, daß  $\mathcal{L}^2(\mathcal{E}, \mu, \mathfrak{U})$  wie in Anmerkung 79 beschrieben derjenige Raum ist, der aus dem Raum  $L^2(\mathcal{E}, \mu, \mathfrak{U})$  der quadratisch  $\mu$ -integrierbaren Funktionen entsteht, wenn man dort jeweils alle Funktionen identifiziert, die sich nur auf einer  $\mu$ -Nullmenge unterscheiden.

<sup>88</sup>Siehe beispielsweise [444].



### 2.6.3.2 Weitere Räume verallgemeinerter Funktionen

Die beschriebenen Gelfand-Tripel lassen sich weiter verallgemeinern; man erhält dabei noch beträchtlich größere verallgemeinerte Funktionenräume.  $\mathcal{E}$  sei nun ein beliebiger komplexer Hilbertraum. Eine erste Verallgemeinerung stammt von Kondratiev und Streit [320]<sup>89</sup> und verwendet die für alle Elemente  $\mathfrak{f}$  von  $(\mathcal{L}^2)$  in eindeutiger Weise mögliche *Wiener-Itô-Segal-Zerlegung*. Dazu brauchen wir den Begriff des *mehrfachen Wiener-Integrals*  $\mathfrak{I}_n$  auf  $\mathcal{H}^{\otimes n}$ ; dieses ist für alle  $f^{(n)} \in \mathcal{H}^{\otimes n}$  eindeutig definiert durch die Eigenschaft

$$\mathfrak{I}_n(f_1^{\otimes n_1} \widehat{\otimes} f_2^{\otimes n_2} \widehat{\otimes} \dots \widehat{\otimes} f_l^{\otimes n_l}) = \mathcal{H}_{n_1}(\langle \cdot, f_1 \rangle) \mathcal{H}_{n_2}(\langle \cdot, f_2 \rangle) \dots \mathcal{H}_{n_l}(\langle \cdot, f_l \rangle)$$

für alle  $n_1 + n_2 + \dots + n_l = n$  und alle  $f_1, f_2, \dots, f_l \in \mathcal{H}$ . Dabei sind die  $\mathcal{H}_n$  verallgemeinerte Hermite-Polynome, die durch

$$\mathcal{H}_n(x) = (-1)^n e^{x^2/2} D_x^n e^{-x^2/2}$$

gegeben sind;  $D_x$  ist die Gâteaux-Ableitung in  $x$ -Richtung. Die Wiener-Itô-Segal-Zerlegung eines beliebigen Vektors  $\mathfrak{f} \in (\mathcal{L}^2)$  schreibt sich damit<sup>90</sup>

$$\mathfrak{f} = \sum_{n=0}^{\infty} \mathfrak{I}_n(\mathfrak{f}^{(n)}) \quad (2.54)$$

mit eindeutig bestimmten  $\mathfrak{f}^{(n)} \in \mathcal{H}^{\otimes n}$ . Das liefert gleichzeitig für die  $(\mathcal{L}^2)$ -Norm von  $\mathfrak{f}$  den Ausdruck

$$\|\mathfrak{f}\|_0^2 = \sum_{n=0}^{\infty} n! \|\mathfrak{f}^{(n)}\|_0^2.$$

Außerdem seien  $\hat{A}$  ein Operator auf  $\mathcal{E}$  und  $\{\xi_j\}_{j \in \mathbb{N}^*}$  eine Orthonormalbasis von  $\mathcal{E}$  aus Eigenvektoren von  $\hat{A}$  mit den Eigenschaften

- (i)  $\hat{A}\xi_j = \lambda_j \xi_j, \quad j \in \mathbb{N}^*,$
- (ii)  $1 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n \leq \dots,$
- (iii)  $\sum_{j=1}^{\infty} \lambda_j^{-\alpha} < \infty$  mit einer geeigneten Konstanten  $\alpha > 0.$

Mit Hilfe der für eine beliebige Zahl  $0 \leq \beta < 1$  und für alle  $p \in \mathbb{N}$  durch

$$\|\mathfrak{f}\|_{p,\beta}^2 := \sum_{n=0}^{\infty} (n!)^{1+\beta} \|(\hat{A}^p)^{\otimes n} \mathfrak{f}^{(n)}\|_0^2$$

<sup>89</sup>vergleiche auch [333].

<sup>90</sup>Die Aussage, eine solche Zerlegung sei stets eindeutig möglich, wurde zuerst von Wiener bewiesen [527]; später führte Itô die mehrfachen Wiener-Integrale ein und konnte das Resultat damit rekonstruieren [265], weswegen es als *Satz von Wiener-Itô* bekannt geworden ist. Äquivalent dazu ist die bereits erwähnte Isomorphie von  $\Gamma(\mathcal{H})$  und  $(\mathcal{L}^2)$ . Die Komponenten der direkten Summe  $(\mathcal{L}^2) = \bigoplus_{n=0}^{\infty} \mathcal{H}^{\otimes n}$  werden somit für alle  $n \in \mathbb{N}$  jeweils aufgespannt durch alle verallgemeinerten Hermite-Polynome  $\mathcal{H}_n(x)$ , deren Variablen  $x(t)$  verallgemeinerte Funktionen aus  $\mathcal{E}'$  sind. Näheres dazu steht in [398].

gegebenen Normenfamilie  $\{\|\cdot\|_{p,\beta}\}_{p \in \mathbb{N}}$  definieren wir die Räume

$$(\mathcal{E}_p)_\beta := \{f \in (\mathcal{L}^2) \mid \|f\|_{p,\beta} < \infty\}$$

und deren Dualräume  $(\mathcal{E}_p)'_\beta$ ; auf diesen sind die zugehörigen Normen durch

$$\|\varphi\|_{p,\beta}^2 = \sum_{n=0}^{\infty} (n!)^{1-\beta} \|(\hat{A}^{-p})^{\otimes n} f^{(n)}\|_0^2.$$

gegeben. Dabei gilt  $(\mathcal{E}_0)_0 = (\mathcal{L}^2)$ , aber  $(\mathcal{E}_0)_\beta \neq (\mathcal{L}^2)$  für  $\beta \neq 0$ . Außerdem definieren wir

$$(\mathcal{E})_\beta := \bigcap_{p=0}^{\infty} (\mathcal{E}_p)_\beta$$

und dessen Dualraum  $(\mathcal{E})'_\beta$ . Den Raum  $(\mathcal{E})_\beta$  kann man wieder mit der aus der Familie  $\{(\mathcal{E}_p)_\beta\}_{p \in \mathbb{N}}$  kommenden projektiven Topologie versehen. Da  $(\mathcal{E})_\beta$  ein nuklearer Raum ist, konstruieren wir damit wiederum ein Gelfand-Tripel

$$(\mathcal{E})_\beta \subset (\mathcal{L}^2) \subset (\mathcal{E})'_\beta.$$

Dabei gilt für alle  $p \in \mathbb{N}$

$$(\mathcal{E})_\beta \subset (\mathcal{E}_p)_\beta \subset (\mathcal{L}^2) \subset (\mathcal{E}_p)'_\beta \subset (\mathcal{E})'_\beta$$

und außerdem für  $0 \leq \beta < 1$

$$(\mathcal{E})_\beta \subset (\mathcal{E}) \subset (\mathcal{L}^2) \subset (\mathcal{E})' \subset (\mathcal{E})'_\beta.$$

Wir erwähnen noch, daß die kanonische Bilinearform  $\langle\langle \cdot, \cdot \rangle\rangle$  auf  $(\mathcal{E})_\beta \times (\mathcal{E})'_\beta$  gegeben ist durch

$$\langle\langle \Phi, \mathfrak{F} \rangle\rangle = \sum_{n=0}^{\infty} (n!)^{1+\beta} \langle (\hat{A}^p)^{\otimes n} \Phi^{(n)}, (\hat{A}^p)^{\otimes n} \mathfrak{F}^{(n)} \rangle.$$

Eine noch weitergehende Verallgemeinerung wurde von Cochran, Kuo und Sengupta vorgeschlagen [106]<sup>91</sup>. Dazu sei  $\{\alpha_n\}_{n \in \mathbb{N}}$  eine Folge mit

(i)  $\alpha_0 = 1$ ,

(ii) Es gibt ein  $\sigma \geq 1$  so daß  $\inf_{n \geq 0} \alpha_n \sigma^n > 0$ ,

(iii)  $\lim_{n \rightarrow \infty} (\alpha_n/n!)^{1/n} = 0$ ,

(iv) Es gibt eine Konstante  $C > 0$  so daß  $\alpha_m \alpha_n \leq C^{n+m} \alpha_{m+n}$  gilt für alle  $m, n \in \mathbb{N}$ .

Außerdem sei  $f$  ein Element von  $(\mathcal{L}^2)$  mit Wiener-Itô-Segal-Zerlegung (2.54) und  $p$  eine natürliche Zahl. Hierfür definieren wir die Norm

$$\|f\|_{p,\alpha}^2 := \sum_{n=0}^{\infty} n! \alpha_n \|f^{(n)}\|_p^2;$$

<sup>91</sup>Vergleiche auch [20].

für alle  $p \in \mathbb{N}$  ist dann

$$[\mathcal{E}_p]_\alpha := \{ \mathbf{g} \in (\mathcal{L}^2) \mid \|\mathbf{g}\|_{p,\alpha} < \infty \}$$

bezüglich der Norm  $\|\cdot\|_{p,\alpha}$  ein Hilbertraum. Aus Eigenschaft (ii) folgt  $[\mathcal{E}_p]_\alpha \subset [\mathcal{E}_q]_\alpha$  für alle  $p \geq \frac{\ln \sigma}{2 \ln \lambda_1}$ . Wieder betrachten wir den Raum

$$[\mathcal{E}]_\alpha := \bigcap_{p=0}^{\infty} [\mathcal{E}_p]_\alpha$$

und dessen Dualraum  $[\mathcal{E}]'_\alpha$ . Wir können eine anschauliche Darstellung dieses Dualraums konstruieren, indem wir die Tensoralgebra

$$T(\mathcal{E}) = \bigoplus_{n=0}^{\infty} \mathcal{E}^{\otimes n}$$

und deren Dualraum  $T(\mathcal{E})'$  betrachten. Definieren wir auf ersterer die Normenfamilie

$$\|(\varphi_0, \varphi_1, \dots)\|_{-q,\alpha}^2 := \sum_{n=0}^{\infty} \frac{\|\varphi_n\|_{-q}^2}{n! \alpha_n}, \quad q \in \mathbb{N}$$

und damit die Räume

$$[\mathcal{E}_{-q}]_\alpha := \{ (\varphi_0, \varphi_1, \dots) \mid \text{für alle } n \in \mathbb{N} \text{ ist } \varphi_n \text{ ein symmetrisches lineares Funktional auf } \mathcal{E}^{\otimes n}, \text{ und es gilt } \|(\varphi_0, \varphi_1, \dots)\|_{-q,\alpha}^2 < \infty \},$$

dann erhalten wir

$$[\mathcal{E}]'_\alpha = \bigcup_{q=0}^{\infty} [\mathcal{E}_{-q}]_\alpha.$$

Damit haben wir für alle  $p \geq \frac{\ln \sigma}{2 \ln \lambda_1}$  das Gelfandsche Raumtripel

$$[\mathcal{E}]_\alpha \subset (\mathcal{L}^2) \subset [\mathcal{E}]'_\alpha.$$

Solche Räume werden nach ihren Entdeckern *Cochran-Kuo-Sengupta-Räume* oder kurz *CKS-Räume* genannt. Dabei ergibt sich die zusätzliche Inklusionskette

$$[\mathcal{E}]_\alpha \subset (\mathcal{E})_\beta \subset (\mathcal{L}^2) \subset (\mathcal{E}_p)'_\beta \subset [\mathcal{E}]'_\alpha.$$

Speziell liefert  $\alpha_n \equiv 1$  die Räume  $(\mathcal{E}) \subset (\mathcal{L}^2) \subset (\mathcal{E})'$  und  $a_n = (n!)^\beta$ ,  $n \in \mathbb{N}$  die Räume  $(\mathcal{E})_\beta \subset (\mathcal{L}^2) \subset (\mathcal{E})'_\beta$  von oben.

Verzichtet man beim Ausgangsraum auf die Eigenschaft der Nuklearität, lassen sich noch größere Distributionenräume für weißes Rauschen konstruieren. Eine Beispiel hierfür wurde von Meyer und Yan gefunden [365]; eine leicht abgewandelte Version stammt von Kondratiev und Streit [320]. Wir betrachten diese hier abschließend. Dazu sei

$$\mathcal{P}_{\text{cyl}}(\mathcal{E}') = \{ P(\langle \cdot, f_1 \rangle, \langle \cdot, f_2 \rangle, \dots, \langle \cdot, f_n \rangle) \mid f_i \in \mathcal{E}; i = 1, 2, \dots, n; P \in \mathbb{R}[X]; n \in \mathbb{N} \}$$

die Algebra der Zylinder-Polynome auf  $\mathcal{E}'$ . Außerdem seien für Elemente  $f$  von  $(\mathcal{L}^2)$  mit Wiener-Itô-Segal-Zerlegung (2.54) und  $p, q \in \mathbb{N}$  diesmal Normen der Form

$$\|f\|_{\mathcal{M}_{p,q}}^2 := \sum_{n=0}^{\infty} (n!)^2 2^{-qn} \|f^{(n)}\|_p^2 = \sum_{n=0}^{\infty} (n!)^2 2^{-qn} \|(\hat{A}^{\otimes n})^p f^{(n)}\|^2$$

gegeben; dazu seien  $\mathcal{M}_{p,q}$ ,  $p, q \in \mathbb{N}$ , die Räume, die durch Vervollständigung von  $\mathcal{P}_{\text{cyl}}(\mathcal{E}')$  in Bezug auf die Normen  $\|\cdot\|_{\mathcal{M}_{p,q}}$  entstehen. Deren Dualräume sind gegeben durch  $\mathcal{M}'_{p,q} = \mathcal{M}_{-p,-q}$ . Diese sind die Vervollständigungen von  $\mathcal{P}_{\text{cyl}}(\mathcal{E}')$  bezüglich der dualen Normen  $\|\cdot\|_{\mathcal{M}'_{p,q}}$ . Zusätzlich sei

$$\mathcal{M}_p = \bigcup_{q=0}^{\infty} \mathcal{M}_{p,q}.$$

Nun bilden wir den Raum

$$\mathcal{M} = \bigcap_{p=0}^{\infty} \mathcal{M}_p = \bigcap_{p=0}^{\infty} \bigcup_{q=0}^{\infty} \mathcal{M}_{p,q}$$

und dessen Dualraum  $\mathcal{M}'$ . Für diesen gilt<sup>92</sup>

$$\mathcal{M}' = \bigcap_{p=0}^{\infty} \mathcal{M}'_p = \bigcap_{p=0}^{\infty} \bigcup_{q=0}^{\infty} \mathcal{M}'_{p,q}.$$

$\mathcal{M}$  wird mit der durch die Raumfamilie  $\{\mathcal{M}_p\}_{p \in \mathbb{N}}$  gegebene projektiven Topologie zu einem topologischen Vektorraum und heißt *Meyer-Yan-Raum*. Den dualen Raum  $\mathcal{M}'$  nennt man den *Raum der Meyer-Yan-Distributionen*. Insgesamt erhalten wir damit das Raumtripel

$$\mathcal{M} \subset (\mathcal{L}^2) \subset \mathcal{M}'$$

und die Inklusionskette

$$\mathcal{M} \subset [\mathcal{E}]_{\alpha} \subset (\mathcal{E})_{\beta} \subset (\mathcal{L}^2) \subset (\mathcal{E}_p)'_{\beta} \subset [\mathcal{E}]'_{\alpha} \subset \mathcal{M}'.$$

Die Bedeutung dieser Hierarchie verallgemeinerter Distributionenräume geht weit über deren Anwendung im Zusammenhang mit weißem Rauschen hinaus. Insbesondere erlauben sie eine Klassifikation einschließlich tiefgreifender Charakterisierungstheoreme für Distributionen; eines davon wird uns gleich noch weiter beschäftigen<sup>93</sup>.

### 2.6.3.3 Hida- und Kondratiev-Distributionen

Um zu der angekündigten Verallgemeinerung des Raums des weißen Rauschens zu gelangen, betrachten wir nun den Spezialfall  $\mathcal{E} \equiv \mathcal{S}(\mathbb{R})$  und das Gelfand-Tripel

$$(\mathcal{S}) \subset \Gamma(\mathcal{H}) \subset (\mathcal{S})'$$

<sup>92</sup>Siehe wieder [444].

<sup>93</sup>In [320] wird ein weitergehender Überblick hierzu gegeben.

oder äquivalent dazu

$$(\mathcal{S}) \subset (\mathcal{L}^2) \subset (\mathcal{S})'.$$

Die Elemente von  $(\mathcal{S})'$  heißen *Hida-Distributionen*. Sie sind stetige lineare Funktionale auf dem symmetrischen Fock-Raum über  $\mathcal{H}$  und stellen die unendlichdimensionale Version der temperierten Distributionen dar. Analog erhalten wir für den mehrdimensionalen Fall mit  $\mathcal{E} \equiv \mathcal{S}_d$  das Gelfand-Tripel

$$(\mathcal{S}_d) \subset (\mathcal{L}_d^2) \subset (\mathcal{S}_d)';$$

die Elemente von  $(\mathcal{S}_d)'$  nennt man *Kondratiev-Distributionen*. Diese Konstruktionen lassen sich wie oben weiter verallgemeinern, womit wir die Gelfandschen Inklusionsketten

$$(\mathcal{S})_\beta \subset (\mathcal{S}) \subset (\mathcal{L}^2) \subset (\mathcal{S})' \subset (\mathcal{S})'_\beta$$

beziehungsweise

$$(\mathcal{S}_d)_\beta \subset (\mathcal{S}_d) \subset (\mathcal{L}_d^2) \subset (\mathcal{S}_d)' \subset (\mathcal{S}_d)'_\beta.$$

bekommen. Zusammen mit dem oben beschriebenen Gaußschen Maß  $\mu$  haben wir damit die verallgemeinerten Weißes-Rauschen-Räume  $((\mathcal{S})', \mathcal{B}[(\mathcal{S})'], \mu)$  und  $((\mathcal{S}_d)', \mathcal{B}[(\mathcal{S}_d)'], \mu)$  sowie  $((\mathcal{S})_\beta, \mathcal{B}[(\mathcal{S})_\beta], \mu)$  und  $((\mathcal{S}_d)_\beta, \mathcal{B}[(\mathcal{S}_d)_\beta], \mu)$ , welche die Bühne für den folgenden Abschnitt darstellen. Bevor wir diesen einläuten, müssen wir uns zuerst noch das entscheidende Hilfsmittel dafür verschaffen, nämlich das angedeutete Theorem zur Charakterisierung verallgemeinerter Distributionen.

#### 2.6.3.4 Das Charakterisierungstheorem

Das im folgenden diskutierte Resultat erweist sich als der Schlüssel zur Interpretation von Pfadintegralen in der Sprache der Mathematik des weißen Rauschens. Um es zu formulieren, brauchen wir zunächst einen weiteren neuen Begriff.

**Definition:** (*S*-Transformation) Unter der *S*-Transformation von  $\Phi = \{f_n\}_{n \in \mathbb{N}} \in [\mathcal{E}]'_\alpha$  versteht man die Funktion

$$\hat{S} \Phi(x) = \sum_{n=0}^{\infty} \frac{1}{n!} f_n(x^{\otimes n}), \quad x \in \mathcal{E}. \quad (2.55)$$

Man überzeugt sich sofort, daß aus der Wiener-Itô-Segal-Zerlegung

$$\Phi = \sum_{n=0}^{\infty} \mathcal{I}_n(\Phi^{(n)})$$

einer verallgemeinerten Funktion  $\Phi \in (\mathcal{E})'_\beta$  deren *S*-Transformation in der Gestalt

$$\hat{S} \Phi(f) = \sum_{n=0}^{\infty} \langle \Phi^{(n)}, f^{\otimes n} \rangle, \quad f \in \mathcal{E},$$

folgt. Außerdem ist die *S*-Transformation injektiv, denn die lineare Hülle der Familie  $\{x^{\otimes n} \mid x \in \mathcal{E}\}$  liefert den gesamten Raum  $\mathcal{E}^{\otimes n}$ .

Da die Eigenschaft der Nuklearität für die betrachteten Räume weiterhin vorausgesetzt werden soll, ist der allgemeinste Rahmen, innerhalb dessen wir uns bewegen können, derjenige der CKS-Räume aus Abschnitt 2.6.2.3. Wieder betrachten wir eine Folge  $\{\alpha_n\}_{n \in \mathbb{N}}$  mit

- (i)  $\alpha_0 = 1$ ,
- (ii) Es gibt ein  $\sigma \geq 1$  so daß  $\inf_{n \geq 0} \alpha_n \sigma^n > 0$ ,
- (iii)  $\lim_{n \rightarrow \infty} (\alpha_n/n!)^{1/n} = 0$ ,

Zusätzlich definieren wir die Funktion

$$G_\alpha(z) := \sum_{n=0}^{\infty} \frac{\alpha_n z^n}{n!};$$

man überzeugt sich mit dem Wurzelkriterium leicht, daß Bedingung (iii) genau dann zutrifft, wenn diese Reihe in der ganzen komplexen Ebene konvergiert, und damit genau dann, wenn  $G_\alpha$  eine ganze Funktion von  $z \in \mathbb{C}$  ist.

Damit läßt sich das Charakterisierungstheorem in der angekündigten allgemeinen Form folgendermaßen formulieren [106]<sup>94</sup>.

**Satz:** (Cochran, Kuo, Sengupta) *Es seien  $\{\alpha_n\}_{n \in \mathbb{N}}$  eine Folge mit den Eigenschaften (i), (ii) und (iii),  $\mathcal{E}$  ein komplexer Hilbertraum und  $F$  eine Funktion auf  $\mathcal{E}$ . Dann ist  $F$  genau dann die  $S$ -Transformierte einer eindeutig bestimmten verallgemeinerten Distribution  $\Phi$  aus  $[\mathcal{E}'_\alpha]$ , wenn folgendes gilt:*

- (a) Für jedes  $g$  und  $h$  aus  $\mathcal{E}$  ist  $F(g + zh)$  eine ganze Funktion von  $z \in \mathbb{C}$ .
- (b) Es gibt ein  $p \in \mathbb{N}$ , so daß

$$\limsup_{n \rightarrow \infty} \left\{ \frac{n}{\alpha_n^{1/n}} \left[ \inf_{r > 0} \frac{1}{r} \sup_{|x|_p = \sqrt{r}} |F(x)|^{2/n} \right] \right\} < \infty.$$

**BEWEIS:** („ $\implies$ “) Es sei  $\Phi \in [\mathcal{E}'_\alpha]$  und  $F \equiv \hat{S} \Phi$ . Eigenschaft (a) folgt direkt aus der Definition (2.55) der  $S$ -Transformation. Zum Nachweis von Eigenschaft (b) leiten wir zwei Abschätzungen her, aus welchen diese Eigenschaft folgt. Für die erste Abschätzung verwenden wir die Zerlegung

$$[\mathcal{E}'_\alpha] = \bigcup_{p \in \mathbb{N}} [\mathcal{E}_{-p}]_\alpha;$$

$\Phi$  muß in einer der Komponenten enthalten sein, es gibt also ein  $p \in \mathbb{N}$ , so daß  $\Phi \in [\mathcal{E}_{-p}]_\alpha$  gilt. Daraus folgt unter Verwendung der Definitionen der  $S$ -Transformation und der Funktion  $G_\alpha$

$$|F(g)| \leq \|\Phi\|_{-p} \sqrt{G_\alpha(\|g\|_p^2)}, \quad g \in \mathcal{E}. \quad (2.56)$$

<sup>94</sup>Die allgemeinste Form dieses Theorems wurde von Ji und Obata bewiesen [278], die ursprüngliche speziellere Fassung stammt von Pothhoff und Streit [419].

Für die zweite Abschätzung stellen wir einerseits fest, daß  $\alpha_n/n! \geq 0$  gilt für alle  $n \in \mathbb{N}$  und daraus

$$G_\alpha(x) \geq \frac{\alpha_n x^n}{n!}$$

folgt für alle  $x \geq 0$ . Damit erhalten wir

$$1 \leq \frac{n!}{\alpha_n} \inf_{x>0} \frac{G_\alpha(x)}{x^n}$$

für alle  $x \geq 0$  als untere Abschätzung. Andererseits ergibt sich aus der Divergenz der

$$\text{Folge } a_n := \frac{\alpha_{n-1} n!}{\alpha_n (n-1)!}$$

$$\frac{\alpha_k}{k!} = \frac{\alpha_n}{n!} \prod_{j=k+1}^n a_j \leq a_n^{n-k} \frac{\alpha_n}{n!} \quad \text{falls } k \leq n$$

beziehungsweise

$$\frac{\alpha_n}{n!} = \frac{\alpha_k}{k!} \prod_{j=n+1}^k a_j \geq a_{n+1}^{k-n} \frac{\alpha_k}{k!} \leq a_n^{k-n} \frac{\alpha_k}{k!} \quad \text{falls } k > n,$$

zusammen also

$$\frac{\alpha_k}{k!} \leq \frac{\alpha_n}{n!} a_n^{n-k} \quad \text{für alle } k \in \mathbb{N}.$$

Auf dem Intervall  $(0, a_n)$  folgt daraus

$$G_\alpha(x) \leq \frac{\alpha_n a_n^n}{n! \left(1 - \frac{x}{a_n}\right)}$$

und mit der Substitution  $t = x/a_n$  auf  $(0, 1)$  weiter

$$\frac{G_\alpha(x)}{x^n} \leq \frac{\alpha_n}{n! t^n (1-t)} \equiv g(t).$$

Das Infimum der linken Seite kann man abschätzen, indem man das Minimum der durch die rechte Seite definierten Funktion  $g$  berechnet. Nullsetzen der Ableitung

$$g'(t) = \frac{\alpha_n t^{n-1} [(n+1)t - n]}{n! [t^n (1-t)]^2}$$

liefert  $t = n/(n+1)$  als einzige Extremstelle, und wegen  $\lim_{t \rightarrow 0} g(t) = \lim_{t \rightarrow 1} g(t) = +\infty$  ist das in der Tat ein Minimum. Dadurch erhalten wir

$$\inf_{x>0} \frac{G_\alpha(x)}{x^n} \leq \frac{\alpha_n}{n!} (n+1) \left(1 + \frac{1}{n}\right)^n$$

und mit  $\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e$  weiter

$$\inf_{x>0} \frac{G_\alpha(x)}{x^n} \leq \frac{\alpha_n}{n!} (n+1) e$$



als obere Abschätzung. Zusammengenommen ergibt das

$$1 \leq \frac{n!}{\alpha_n} \inf_{x>0} \frac{G_\alpha(x)}{x^n} \leq \frac{\alpha_n}{n!} (n+1) e.$$

Daraus wiederum folgt

$$\lim_{n \rightarrow \infty} \left\{ \left( \frac{n!}{\alpha_n} \right)^{1/n} \inf_{x>0} \frac{[G_\alpha(x)]^{1/n}}{x} \right\} = 1$$

und damit weiter

$$\limsup_{n \rightarrow \infty} \left\{ \frac{n}{\alpha_n^{1/n}} \left[ \inf_{x>0} \frac{(G_\alpha(x))^{1/n}}{x} \right] \right\} < \infty. \quad (2.57)$$

Aus (2.56) und (2.57) folgt unmittelbar die Eigenschaft (b).

(„ $\Leftarrow$ “) Nun sei  $F$  eine Funktion auf  $\mathcal{E}$  mit den Eigenschaften (a) und (b). Aus Eigenschaft (a) und dem Satz von Hartogs<sup>95</sup> folgt, daß  $F$  auf jedem endlichdimensionalen komplexen Unterraum von  $\mathcal{E}$  analytisch ist und folglich auf jedem solchen Unterraum eine eindeutige Taylorreihenentwicklung besitzt. Das bedeutet auch, daß es für jedes  $n \in \mathbb{N}$  ein lineares Funktional  $D^n F(0)$  auf  $\mathcal{E}^{\otimes n}$  gibt, so daß

$$F(g) = \sum_{n=0}^{\infty} \frac{1}{n!} \langle D^n F(0), g^{\otimes n} \rangle \quad (2.58)$$

gilt; dabei ist  $D^n F(0)$  definiert durch

$$\langle D^n F(0), f_1 \otimes f_2 \otimes \dots \otimes f_n \rangle = \left( \frac{\partial^n}{\partial x_1 \partial x_2 \dots \partial x_n} \right)_{x_1=x_2=\dots=x_n=0} F(x_1 f_1 + x_2 f_2 + \dots + x_n f_n)$$

für beliebige  $f_1, f_2, \dots, f_n \in \mathcal{E}$ . Mit Hilfe der Definition (2.55) der  $S$ -Transformation läßt sich (2.58) auch in der Form

$$F = S \left( \sum_{n=0}^{\infty} D^n F(0) \right)$$

schreiben; folglich ist nun zu zeigen, daß  $\Psi := (F(0), DF(0), D^2 F(0), \dots, D^n F(0), \dots)$  ein Element von  $[\mathcal{E}]'_\alpha$  ist. Im einzelnen erfordert das den Nachweis, daß erstens die  $D^n F(0)$ ,  $n \in \mathbb{N}$ , stetige lineare Funktionale auf den Räumen  $\mathcal{E}^{\otimes n}$ ,  $n \in \mathbb{N}$ , sind und zweitens die Ungleichung

$$\sum_{n=0}^{\infty} \frac{\|D^n F(0)\|_{-q}^2}{n! \alpha_n} < \infty \quad (2.59)$$

besteht.

Die Linearität der  $D^n F(0)$  ist klar. Für die Stetigkeit genügt es zu zeigen, daß es sich dabei um beschränkte lineare Funktionale handelt. Da  $F(z_1 x_1 + z_2 x_2 + \dots + z_n x_n)$  analytisch ist, folgt mit der Cauchyschen Integralformel

$$\langle D^n F(0), x_1 \otimes x_2 \otimes \dots \otimes x_n \rangle$$

<sup>95</sup>Siehe beispielsweise [60].

$$= \frac{1}{(2\pi n)^n} \int_{|z_1| < r} \int_{|z_2| < r} \cdots \int_{|z_n| < r} \frac{F(z_1 x_1 + z_2 x_2 + \cdots + z_n x_n)}{z_1^2 + z_2^2 + \cdots + z_n^2} dz_1 dz_2 \cdots dz_n$$

für beliebige  $x_1, x_2, \dots, x_n \in \mathcal{E}$  und  $r > 0$ . Beschränkt man sich zunächst auf den Spezialfall  $\|x_1\|_p = \|x_2\|_p + \cdots + \|x_n\|_p = 1$ , liefert das Maximumsprinzip der Funktionentheorie

$$|\langle D^n F(0), x_1 \otimes x_2 \otimes \cdots \otimes x_n \rangle| \leq \frac{1}{r^n} \sup_{\|z\|_p = nr} |F(z)|.$$

Die  $D^n F(0)$  sind linear, folglich gilt für beliebige  $y_1, y_2, \dots, y_n \in \mathcal{E}$

$$|\langle D^n F(0), y_1 \otimes y_2 \otimes \cdots \otimes y_n \rangle| \leq \left[ \frac{1}{r^n} \sup_{\|z\|_p = nr} |F(z)| \right] \|y_1\|_p \|y_2\|_p \cdots \|y_n\|_p.$$

Nun sei  $i_{qp} : \mathcal{E}_q \rightarrow \mathcal{E}_p$  die lineare Abbildung, die durch Fortsetzung der Identität auf  $\mathcal{E}$  entsteht und außerdem  $\{e_j \mid j \in \mathcal{J}\}$  eine Orthonormalbasis von  $\mathcal{E}_q$  mit einer geeigneten Indexmenge  $\mathcal{J}$ <sup>96</sup>. Für  $i_{qp}$  definieren wir die Hilbert-Schmidt-Norm

$$\|i_{qp}\|_{HS} := \sqrt{\sum_{j \in \mathcal{J}} |i_{qp}(e_j)|_p^2}.$$

Wählt man  $q > p$  so, daß  $\|i_{qp}\|^2 < \infty$  gilt, dann folgt für beliebige orthonormale Familien  $\{u_j \mid j \in \mathcal{J}_q\}$  in  $\mathcal{E}_q$

$$\sum_{j_1, j_2, \dots, j_n \in \mathcal{J}_q} |\langle D^n F(0), u_{j_1} \otimes u_{j_2} \otimes \cdots \otimes u_{j_n} \rangle|^2 \leq \frac{1}{r^{2n}} \|i_{qp}\|_{HS}^{2n} \sup_{\|z\|_p = nr} |F(z)|^2.$$

Daher ist für jeden beliebigen endlichdimensionalen Unterraum  $\mathcal{W}$  von  $\mathcal{E}$  die Einschränkung von  $D^n F(0)$  auf  $\mathcal{W}^{\otimes n}$  bezüglich der Norm  $\|\cdot\|_q$  eine stetige Abbildung. Jedes Element von  $\mathcal{E}^{\otimes n}$  ist auch in einem der  $\mathcal{W}^{\otimes n}$ , sodaß alle  $D^n F(0)$  beschränkte und somit stetige lineare Funktionale auf  $\mathcal{E}^{\otimes n}$  sind.

Damit fehlt noch der Nachweis der Ungleichung (2.59). Für die  $D^n F(0)$  gilt zusätzlich

$$\|D^n F(0)\|_{-q} \leq \frac{1}{r^{2n}} \|i_{qp}\|_{HS}^{2n} \sup_{\|z\|_p = nr} |F(z)|^2$$

und damit weiter

$$\sum_{n=0}^{\infty} \frac{\|D^n F(0)\|_{-q}^2}{n! \alpha_n} \leq \sum_{n=0}^{\infty} \frac{\|i_{qp}\|_{HS}^{2n}}{n! \alpha_n} \left\{ \inf_{r>0} \left[ \frac{1}{r^{2n}} \sup_{\|z\|_p = nr} |F(z)|^2 \right] \right\}. \quad (2.60)$$

<sup>96</sup>Bei separablen Hilberträumen ist  $\mathcal{J} = \mathbb{N}$ . Bei überseparablen Hilberträumen ist  $\mathcal{J}$  größer und insbesondere überabzählbar; die subtilen Details, die hier auftreten, spielen bei obigem Beweis keine Rolle, weswegen sie dort unberücksichtigt bleiben. Werden sie berücksichtigt, so muß insbesondere sehr sorgfältig definiert werden, was bei Indexmengen, die beliebig überabzählbar sein können, unter einer Summe zu verstehen ist. Genauereres dazu findet man beispielsweise in [249].

Diese Abschätzung taugt natürlich nur dann etwas zum Beweis der Behauptung, wenn die rechte Seite konvergent ist. Um sie abzuschätzen, verwenden wir einerseits die aus der Theorie der lokalkonvexen Vektorräume bekannte Relation

$$\lim_{q \rightarrow \infty} \|i_{qp}\| = 0 \quad (2.61)$$

und andererseits die Voraussetzung (b). Aus letzterer erhält man unter Verwendung der aus der Stirlingschen Formel folgenden Beziehung  $(n!)^{1/n} \sim n/e$  für große  $n$  und gleichzeitigem Ersetzen von  $\sqrt{r}$  durch  $nr$

$$\limsup_{n \rightarrow \infty} \left\{ \frac{1}{n \alpha_n} \inf_{r > 0} \left[ \frac{1}{r^{2n}} \sup_{|x|_p = nr} |F(x)|^2 \right] \right\} < \infty. \quad (2.62)$$

(2.61) und (2.62) liefern zusammen die Konvergenz der Reihe auf der rechten Seite von (2.60) und damit auch die Ungleichung (2.59).  $\square$

Der Vollständigkeit halber sei erwähnt, daß es solche Charakterisierungstheoreme nicht nur für verallgemeinerte Funktionen, sondern auch für Testfunktionen gibt<sup>97</sup>. – Die Bedeutung von Charakterisierungstheoremen wie dem soeben beschriebenen geht natürlich über die Verwendung bei der Definition von Pfadintegralen weit hinaus. Das betrifft insbesondere die Theorie der stochastischen Differentialgleichungen; näheres dazu steht unter anderem in [106], [333] und [419]. Die mathematische Physik macht davon in vielfältiger Weise Gebrauch, insbesondere in der Quantenmechanik, wofür der nächste Abschnitt nur ein Beispiel unter vielen ist, aber beispielsweise auch in der Hydrodynamik.

### 2.6.4 Feynman-Integrale und weißes Rauschen

Der zur Definition Feynmanscher Pfadintegrale mit Hilfe von weißem Rauschen notwendige mathematische Apparat steht nunmehr bereit [333], [342]. Dazu betrachten wir einen Brownschen Prozeß  $B(t)$  und schreiben denselben mit Hilfe des Raums  $(\mathcal{S}'_d, \mathcal{B}(\mathcal{S}'_d), \mu)$  in der Form (2.53); genauer gesagt soll die Brownsche Bewegung am Ort  $x$  starten, sodaß wir dafür

$$y(t) = x + \int_{t_0}^t \varphi(t') dt' \equiv x + B(t)$$

schreiben können, mit  $\varphi \in \mathcal{S}'_d$ . Setzen wir diesen Brownschen Pfad versuchsweise in den Ausdruck (2.28) des heuristischen Pfadintegrals aus Abschnitt 2.3.1.2 ein, so erhalten wir zunächst rein formal

$$\begin{aligned} \psi(x, t) &= \int_{\Gamma_x} e^{iS(x-B(t))} f(x-B(t)) \mathcal{D}\dot{B} \\ &= \lim_{n \rightarrow \infty} \left( \frac{4\pi it}{n} \right)^{-dn/2} \int_{\mathcal{S}'_d} \exp \left\{ i \int_0^t [\dot{B}(t')]^2 dt' \right\} \times \end{aligned}$$

<sup>97</sup>Ein Beispiel dafür liefert [20].

$$\times \exp \left[ -i \int_0^t V(x - B(t')) dt' \right] f(x - B(t)) \prod_{j=1}^{\infty} d\dot{B}(jt/n). \quad (2.63)$$

Natürlich ist damit allein nichts gewonnen, denn dieser Ausdruck ist zuerst einmal genauso und aus den selben Gründen undefiniert wie (2.28). Erst die Neuinterpretation des Terms  $\exp \left\{ i \int_0^t [\dot{B}(t')]^2 dt' \right\}$  im Rahmen der Theorie des weißen Rauschens weist den Weg, (2.63) eine vernünftige mathematische Bedeutung zu geben. Dazu muß das nicht existierende unendlichdimensionale Lebesgue-Maß wieder durch ein Gaußsches Maß  $\mu$  ersetzt werden, um überhaupt auf unendlichdimensionalen Räumen integrieren zu können, doch damit ist es nicht getan, denn die formale Relation

$$\mathcal{D}\dot{B} = (2\pi)^{\infty} \exp \left\{ -\frac{1}{2} \int_0^t [\dot{B}(t')]^2 dt' \right\} d\mu(\dot{B}) \quad (2.64)$$

macht für sich allein immer noch keinen Sinn. Den entscheidenden zusätzlichen Beitrag leistet das folgende Resultat [319], [320], [333].

**Lemma:** *Es seien  $\{e_n\}_{n \in \mathfrak{N}}$  eine Orthonormalbasis in  $\mathcal{L}^2([0, t])$  und  $z \in \mathfrak{U}^*$ . Dann ist*

$$\Phi := \prod_{j=0}^{\infty} \sqrt{1 - 2z} \exp(z \langle \dot{B}, e_j \rangle^2)$$

eine verallgemeinerte Funktion aus  $(\mathcal{S}_d)'$ , und es gilt

$$\hat{S} \Phi(f) = \exp \left\{ \frac{z}{1 - 2z} \int_0^t [f(t')]^2 dt' \right\} \quad (2.65)$$

für  $f \in \mathcal{S}_d$ .

BEWEIS: Zunächst gilt für die Partialprodukte

$$\Phi_n := \prod_{j=0}^n \sqrt{1 - 2z} \exp(z \langle \dot{B}, e_j \rangle^2)$$

von  $\Phi$  für alle  $n \in \mathfrak{N}$  und  $f \in \mathcal{S}_d$  einerseits  $\Phi_n \in (\mathcal{L}^2)$  und andererseits

$$\hat{S} \Phi_n(f) = \exp \left[ \frac{z}{1 - 2z} \sum_{j=1}^n \langle f, e_j \rangle^2 \right]$$

woraus sich bereits die Einschränkung  $z \neq -1/2$  ergibt. Es folgt weiter für alle  $f \in \mathcal{S}_d$

$$\lim_{n \rightarrow \infty} \hat{S} \Phi_n(f) = \exp \left\{ \frac{z}{1 - 2z} \int_0^t [f(t')]^2 dt' \right\}. \quad (2.66)$$

Die Eigenschaft der Funktionen  $f$ , für Argumente  $x \rightarrow \pm\infty$  schnell fallend zu sein, führt außerdem auf die Abschätzung

$$|\hat{S} \Phi_n(f)| \leq \exp \left( \left| \frac{z}{1-2z} \right| \|f\|_2^2 \right).$$

Mit diesen beiden Relationen folgt durch Anwenden des Charakterisierungstheorems aus dem vorigen Abschnitt die starke Konvergenz der Folge  $\{\Phi_n\}_{n \in \mathbb{N}}$  in  $(\mathcal{S}_d)'$  gegen den Grenzwert  $\Phi$ ; aus (2.66) folgt für dessen  $S$ -Transformation (2.65).  $\square$

Die Grenzfunktion bezeichnen wir mit

$$\Phi = \mathcal{N} \exp \left\{ z \int_0^t [\dot{B}(t')]^2 dt' \right\},$$

ein Ausdruck, der üblicherweise *Renormierung* von  $\exp \left\{ z \int_0^t [\dot{B}(t')]^2 dt' \right\}$  genannt wird. Es ist hilfreich, sich diese Renormierung mathematisch schlampig als

$$\mathcal{N} \exp \left\{ z \int_0^t [\dot{B}(t')]^2 dt' \right\} = \sqrt{1-2z}^\infty \exp \left\{ z \int_0^t [\dot{B}(t')]^2 dt' \right\} \quad (2.67)$$

vorzustellen.

Mit diesem Lemma deutet sich die Möglichkeit an, Feynman-Integrale als Integrale über Hida-Distributionen aufzufassen. Als Ergebnis heben sich die Probleme der einzelnen Bestandteile des ansonsten undefinierten Ausdrucks (2.63) gegenseitig weg. Dazu betrachten wir die Renormierung des Ausdrucks  $\exp \left\{ i \int_0^t [\dot{B}(t')]^2 dt' \right\}$  und erhalten mit (2.64) sowie (2.67) zunächst

$$\begin{aligned} & \lim_{n \rightarrow \infty} \left( \frac{4\pi i t}{n} \right)^{-dn/2} \exp \left\{ i \int_0^t [\dot{B}(t')]^2 dt' \right\} \\ &= \left\{ \mathcal{N} \exp \left[ \left( i + \frac{1}{2} \right) \int_0^t (\dot{B}(t'))^2 dt' \right] \right\} \left( \frac{1}{2\pi} \right)^\infty \exp \left[ \frac{1}{2} \int_0^t (\dot{B}(t'))^2 dt' \right] \end{aligned}$$

und damit weiter für (2.63)

$$\begin{aligned} \psi(x, t) &= \int_{\mathcal{S}'_d} \left\{ \mathcal{N} \exp \left[ \left( i + \frac{1}{2} \right) \int_0^t [\dot{B}(t')]^2 dt' \right] \right\} \times \\ &\quad \times \exp \left[ -i \int_0^t V(x - B(t')) dt' \right] f(x - B(t)) d\mu(\dot{B}). \quad (2.68) \end{aligned}$$

Nach obigem Lemma ist der erste Faktor des Integranden auf der rechten Seite eine Hida-Distribution. Um dem gesamten Integral eine mathematisch sinnvolle Bedeutung geben zu können, sollte der gesamte Integrand

$$\begin{aligned} \mathcal{F}_{V,f}(x,t) \equiv & \left\{ \mathcal{N} \exp \left[ \left( i + \frac{1}{2} \right) \int_0^t [\dot{B}(t')]^2 dt' \right] \right\} \times \\ & \times \exp \left[ -i \int_0^t V(x - B(t')) dt' \right] f(x - B(t)) \end{aligned}$$

eine Distribution aus  $(\mathcal{S})'_\beta$  sein, denn dann kann man (2.68) als Faltung dieser Distribution mit der Testfunktion  $\varphi \equiv 1$  auffassen und erhält den wohldefinierten Ausdruck

$$\int_{\mathcal{S}'_d} \mathcal{F}_{V,f}(x,t) d\mu = \langle\langle \mathcal{F}_{V,f}(x,t), 1 \rangle\rangle.$$

Für die Anwendung auf konkrete physikalische Systeme muß man daher jeweils nachweisen, daß  $\mathcal{F}_{V,f}(x,t)$  für das jeweils vorliegende Potential  $V$  und die Anfangsbedingung  $f$  eine verallgemeinerte Funktion aus dem Raum  $(\mathcal{S})'_\beta$  für eine Zahl  $\beta$  ist – worin der schwierigste Teil der ganzen Angelegenheit liegt. Immerhin genügt dazu die Betrachtung der Anfangsbedingung  $f = \delta_0$ , wobei  $\delta_0$  die Diracsche Delta-Distribution am Ursprung ist, denn aus einer Lösung  $\psi_0(x,t)$  der Schrödingergleichung zu dieser Anfangsbedingung erhält man bekanntlich eine Lösung  $\psi(x,t)$  derselben zu einer beliebigen Anfangsbedingung  $f$  durch

$$\psi(x,t) = \int_{\mathcal{Z}^d} \delta(x - x') \psi_0(x',t) dx'.$$

Der Nachweis von  $\mathcal{F}_{V,\delta_0}(x,t) \in (\mathcal{S})'_\beta$  für ein  $\beta$ , gelang inzwischen für eine Reihe spezieller Potentiale, beispielsweise für freie Teilchen [162], [162], für quadratische Potentiale [110], [162], [205], anharmonische Oszillator-Potentiale [110], [521], verallgemeinerte Kepler-Potentiale [206] und auch gewisse unbeschränkte und sogar schnell anwachsende Potentiale [161], [332]. Der Aufwand ist dabei jeweils erheblich. Als allgemeines Verfahren für beliebige Potentiale läßt sich dieser Zugang jedoch genauso wie die anderen beschriebenen nicht etablieren.

## 2.7 Weitere Varianten

Es gibt eine Vielzahl weiterer mathematisch strenger Rekonstruktionen Feynmanscher Pfadintegrale, deren Verwandtschaft zu den obigen Versionen von sehr eng bis nicht vorhanden reicht. Wir werden, schon um den Rahmen nicht zu sprengen, aber auch, weil dadurch keine weiteren für das vorliegende Buch wesentlichen Erkenntnisse zu gewinnen wären, auf eine ausführliche Betrachtung verzichten und nur einige weitere Beispiele ganz kurz ansprechen. Diese lassen sich grob in vier Kategorien einteilen, womit in keiner Weise Anspruch auf Vollständigkeit erhoben werden soll.

### 2.7.1 Promaße und Prodistributionen

Die erste davon ist im wesentlichen das Werk von Cécile DeWitt-Morette, seit Jahrzehnten eine Protagonistin auf dem Gebiet der Pfadintegrale in der mathematischen Physik. Dieser Zugang ist gleichzeitig der erste Versuch zur Definition von Feynman-Integralen, der auf einer soliden mathematischen Basis aufbaut und nicht gleichzeitig auf Wiener-Integrale ausweicht [94], [95], [115], [120]. Dennoch wird dabei an keiner Stelle irgendetwas diskretisiert; es handelt sich von Anfang an um echte Funktionalintegrale. Allerdings gibt man als Gegenleistung die Maßtheorie im engeren Sinn als Basis auf und sieht sich mit der Notwendigkeit konfrontiert, Integrationstheorie auch auf nicht lokal-konvexen Räumen zu betreiben beziehungsweise zu erklären, was das überhaupt sein soll.

Man startet dazu mit dem Raum  $\mathcal{C}_{(t_a, t_b)}(\mathbb{R}^d) \equiv \mathcal{C}$  der stetigen Kurven auf dem Intervall  $(t_a, t_b)$  sowie dem dazu dualen Raum  $\mathcal{C}'$  der beschränkten komplexen Maße auf  $(t_a, t_b)$ . Zu jedem Element  $\gamma$  von  $\mathcal{C}$  gibt es ein dazu duales Maß  $\mu$  aus  $\mathcal{C}'$ ; diese beiden sind durch die bilineare Relation

$$\langle \mu, \gamma \rangle = \int_{t_a}^{t_b} \gamma(t) d\mu(t) \in \mathfrak{A}'$$

verknüpft sind. Für beliebige  $\mu \in \mathcal{C}'$  sei außerdem  $\mathfrak{w}_\mu$  das beschränkte Gaußsche Maß auf  $\mathcal{C}$ , dessen Fourier-Transformation die Relation

$$\int_{\mathcal{C}} \exp(i\langle \gamma', \gamma \rangle) d\mathfrak{w}_\mu(\gamma) = \exp \left[ -\frac{i}{2} \int_{t_a}^{t_b} \int_{t_a}^{t_b} \text{inf}(t, t') d\mu(t) d\mu(t') \right]$$

erfüllt. Solche Maße nennt man *Pseudomaße*. Damit und mit der klassischen Wirkung  $S$  des zu beschreibenden Systems definiert man nun direkt ein Pfadintegral auf  $\mathcal{C}$  gemäß

$$I(f) = \int_{\mathcal{C}} e^{iS(\gamma)} f(\gamma) d\gamma$$

und erhält damit als Lösung der Schrödingergleichung

$$i \frac{\partial \psi}{\partial t} = \hat{H} \psi$$

mit Anfangsbedingung  $\psi(x, t_0) = \varphi(x) \in \mathcal{L}^2(\mathbb{R}^d)$  den Ausdruck

$$\psi(x, t) = \int_{\mathcal{C}} e^{iS(\gamma(x), t)} \varphi(x) d\mathfrak{w}(\gamma).$$

In der Tat kann man zeigen, daß für das Maß  $\mathfrak{w}$  die Relation

$$\int_{\mathcal{C}} d\mathfrak{w}(\gamma) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \dots \int_{\mathbb{R}^d} \exp \left\{ i \sum_{j=1}^n \frac{[\gamma(x_j) - \gamma(x_{j-1})]^2}{2(t_j - t_{j-1})} \right\} \prod_{j=1}^n \frac{d\gamma(x_j)}{\sqrt{2\pi i(t_j - t_{j-1})}}$$



gilt, wenn man auf der rechten Seite sogenannte *Promaße* oder *zylindrische Maße* verwendet, eine Verallgemeinerung des Maßbegriffs, die in geeigneter Weise jeder Zerlegung eines Intervalls ein Wahrscheinlichkeitsmaß zuordnet<sup>98</sup>. Damit erreicht man eine formale und insbesondere auch numerische Übereinstimmung mit Feynmans heuristischer Pfadintegral-Definition, wodurch konkrete Berechnungen in der üblichen Weise durchführbar sind. Von besonderer Bedeutung ist dabei, daß auch semiklassische Approximationen möglich sind [116]. Explizite Berechnungen von Pfadintegralen findet man zum Beispiel in [94] und [95]. Im Rahmen dieses Zugangs lassen sich auch Pfadintegrale in gekrümmten Raum-Zeiten mathematisch streng definieren [117], [118]. Der Nachteil ist abgesehen von der Notwendigkeit, den Begriff der Maßtheorie stark zu verallgemeinern, auch hier die Einschränkung auf spezielle Potentiale, wenn diese auch weite Gebiete realer Systeme abdecken.

Pseudomaße sind keine Maße im eigentlichen Sinn. Stattdessen kann man sie als Distributionen auffassen und spricht dann auch von *Prodistributionen*. Verwendet man Funktionen  $\varphi \in \mathcal{C}^\infty(\mathbb{R}^d)$  als Testfunktionen, dann ist

$$\langle \mathfrak{w}, \varphi \rangle = \int_{\mathcal{C}} \varphi(\gamma) d\mathfrak{w}(\gamma) \in \mathbb{R}$$

stets definiert, und die Distributionen-Eigenschaften dieser Abbildung lassen sich leicht nachweisen.

Ein auf den ersten Blick ganz anderer Weg geht auf Truman zurück; hierbei werden ganz ähnlich wie im heuristischen Originalzugang von Feynman polygonale Pfade verwendet, um über einen geeigneten Grenzübergang Pfadintegrale auf rigorose Weise zu definieren [510], [511], [512]. Genau besehen handelt es sich hierbei um eine Kombination von Fresnel-Integralen und Promaßen, so daß diese Theorie ebenfalls der ersten Kategorie zuzurechnen ist. Man startet mit einem reflexiven, separablen Banachraum  $\mathcal{E}$  und dessen Dualraum  $\mathcal{E}'$ . Wie bei den Fresnel-Integralen werden auch hier integrierbare Funktionen als Fourier-Transformierte von beschränkten komplexen Maßen auf  $\mathcal{E}'$  definiert und sind somit von der Form

$$f(x) = \int_{\mathcal{E}'} e^{-i\langle x', x \rangle} d\mu(x').$$

Nun sei  $V(x)$  eine Varianzfunktion mit der Eigenschaft, daß  $\varphi(x) = \exp -iV(x)/2$  eine meßbare Funktion ist. Damit definiert man Pfadintegrale durch

$$I(f) = \int_{\mathcal{E}} e^{-iV(x)/2} d\mu(x)$$

womit man eine Art Synthese des Zugangs über Fresnel-Integrale und desjenigen über Promaße erhält. Für die spezielle Wahl  $V(x) = \|x\|^2$  erhält man gerade wieder das Fresnel-Integral von  $f$ . Damit läßt sich die Feynman-Itô-Formel für eine wesentlich größere Klasse von Pfaden formulieren, wodurch ihre Anwendbarkeit stark erweitert wird. Das Problem der Interpretation von  $d\gamma/dt$  für nicht definierbare Pfade wird dabei gelöst, indem diese Ableitung als schwache Ableitung im Distributionensinn aufgefaßt wird.

<sup>98</sup>Der Begriff der Promaße geht ursprünglich auf Bourbaki zurück [72]. Siehe auch [95].

Der Vorteil dieses Zugangs liegt in der verglichen mit den Fresnel-Integralen größeren Klasse von Potentialen, für die er anwendbar ist. Nachteile sind die ebenfalls nicht für alle Potentiale vorhandene Anwendbarkeit und insbesondere die Verwendung von zunächst polygonalen Pfaden, die erst in einem Grenzübergang zu unendlich vielen Unterteilungen zum eigentlichen Pfadintegral führen. Dieser Grenzwert existiert zwar im mathematischen Sinne, aber damit handelt es sich genau besehen nicht um eine echte Funktionalintegration.

### 2.7.2 Imaginäre Resolventen

Die zweite Kategorie ist in wesentlichen Teilen das Werk von M. L. Lapidus und verwendet eine Modifikation der Produktformeln, die in Abschnitt 2.3.1.1 aufgetreten sind. Im Zentrum dieses Ansatzes stehen sogenannte *imaginäre Resolventen* von Operatoren; während die gewöhnliche Resolvente eines linearen Operators  $\hat{A}$  bekanntlich durch

$$\hat{R}(\hat{A}, z) := (\hat{A} - z \mathbf{1})^{-1}$$

definiert ist, versteht man unter dessen imaginärer Resolvente den Operator

$$\hat{J}(\hat{A}, z) := (\hat{A} + iz \mathbf{1})^{-1}.$$

Falls  $\hat{J}(\hat{A}, z)$  existiert, ist es ein linearer Operator auf dem Hilbertraum  $\mathcal{H}$ , auf welchem der Operator  $\hat{A}$  definiert ist.  $\hat{A}$  ist dabei üblicherweise ein unbeschränkter selbstadjungierter Operator, dessen Spektrum  $\sigma(\hat{A})$  ganz in  $\mathbb{R}$  enthalten ist; folglich enthält seine Resolventenmenge  $\rho(\hat{A}) = \mathfrak{U} \setminus \sigma(\hat{A})$  die gesamte imaginäre Achse  $i\mathbb{R}$  mit Ausnahme von höchstens 0. Damit existiert  $\hat{J}(\hat{A}, t)$  für alle  $t \in \mathbb{R}$ . Für solche imaginäre Resolventen läßt sich unter geeigneten Voraussetzungen eine Produktformel ähnlich derjenigen von Trotter beweisen [341]; sie lautet

$$e^{-it(\hat{A} + \hat{B})} = \lim_{n \rightarrow \infty} \left[ \left( \mathbf{1} + \frac{it}{n} \hat{A} \right)^{-1} \left( \mathbf{1} + \frac{it}{n} \hat{B} \right)^{-1} \right]^n.$$

Dabei ist die Operatorsumme  $\hat{A} + \hat{B}$  in dem Exponentialausdruck auf der rechten Seite eine Verallgemeinerung der gewöhnlichen Operatorsumme, die als Summe von geeigneten, durch  $\hat{A}$  und  $\hat{B}$  gegebenen quadratischen Formen definiert ist, was den Vorteil hat, daß die Produktformel auch dann wohldefiniert ist, wenn die gewöhnliche Operatorsumme  $\hat{A} + \hat{B}$  nicht wesentlich selbstadjungiert ist. Wenn das doch der Fall ist, geht die Produktformel für imaginäre Resolventen in die Trotter-Produktformel über.

Mit Hilfe dieser Formel konstruierte Lapidus ein modifiziertes Feynman-Integral der Form

$$\mathcal{I}(V, t) = \lim_{n \rightarrow \infty} \left[ \left( \mathbf{1} + \frac{it}{n} \hat{A} \right)^{-1} \left( \mathbf{1} + \frac{it}{n} \hat{B} \right)^{-1} \right]^n,$$

das für eine große Klasse insbesondere auch singulärer Hamilton-Operatoren definierbar ist und sich auch auf solche Hamilton-Operatoren erweitern läßt, die nicht wesentlich selbstadjungiert sind [284], [341]. Auch Verallgemeinerungen auf Vektorpotentiale und auf Pfadintegralen auf Riemannschen Mannigfaltigkeiten sind im Rahmen der

Pfadintegral-Definition über imaginäre Resolventen möglich. Bivar-Weinholtz und Lapidus zeigten darüberhinaus, wie sich dieser Zugang auf komplexwertige Potentiale erweitern läßt, womit sich Anwendungen auf dissipative, also nicht isolierte quantenmechanische Systeme ergeben [58], [284]<sup>99</sup>. Es dürfte jedoch fast überflüssig sein zu erwähnen, daß auch die Theorie der imaginären Resolventen keine allgemeine Pfadintegraldefinition für beliebige Potentiale ermöglicht.

### 2.7.3 Non-Standard-Analysis

Die dritte Kategorie, die hauptsächlich auf Loo und Nakamura zurückgeht, verwendet Hilfsmittel aus der Non-Standard-Analysis<sup>100</sup>. Ausgangspunkt ist wieder der Ausdruck (2.27) aus Abschnitt 2.3.1.2, also ein Grenzwert aus unendlich vielen Integralen der Form

$$\begin{aligned}\psi(x, t) &= e^{-it\hat{H}} \psi(x, t_0) \\ &= \lim_{n \rightarrow \infty} \left( \frac{4\pi it}{n} \right)^{-dn/2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \dots \int_{\mathbb{R}^d} e^{iS(x_0, x_1, x_2, \dots, x_n, t)} \psi(x_0, t_0) dx_1 dx_2 \dots dx_n.\end{aligned}$$

Nun wählt man anstelle von  $n \in \mathbb{N}$  ein hyperendliches  $n \in {}^*\mathbb{N}$  und betrachtet damit die Größe

$${}^*\mathcal{J} = \left( \frac{4\pi it}{n} \right)^{-dn/2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \dots \int_{\mathbb{R}^d} e^{iS(x_0, x_1, x_2, \dots, x_n, t)} \psi(x_0, t_0) dx_1 dx_2 \dots dx_n.$$

Das läßt sich ganz anschaulich als ein Integral

$${}^*\mathcal{J} = \left( \frac{4\pi it}{n} \right) \int_{{}^*\mathbb{R}^{dn}} e^{iS(\gamma_n+x, t)} {}^*\psi(\gamma_n(0) + x) {}^*d\gamma_n$$

über alle Pfade  ${}^*\gamma_n : {}^*[0, t] \longrightarrow {}^*\mathbb{R}^d$  auffassen, die in folgendem Sinn stückweise linear sind: Wir zerlegen das Intervall  ${}^*[0, t]$  mit Hilfe von hyperendlich vielen Zwischenstellen  $t_j = jt/n$ ,  $j = 0, 1, \dots, n$ ,  $n \in {}^*\mathbb{N}$  in eine hyperendliche Anzahl von Teilintervallen  ${}^*[t_j, t_{j+1}]$  und lassen nur Wege zu, die stetig und auf allen diesen Intervallen linear sind. Solche Wege sehen gewissermaßen von weitem, also mit der Brille der Standard-Analysis betrachtet, differenzierbar aus, das Mikroskop der Non-Standard-Analysis enttarnt sie jedoch als hyperfeine Polygonzüge. Damit hat so ein Integral eigentlich genau die Eigenschaften, die man sich von einem Pfadintegral wünscht und ist zusätzlich mathematisch sauber definiert – nur ist es eben ein Non-Standard-Integral und folglich nicht direkt für

<sup>99</sup>Die Beschreibung dissipativer quantenmechanischer Systeme mit komplexwertigen Potentialen wird ausführlich in [154] dargestellt.

<sup>100</sup>Die wichtigste Referenz zur Non-Standard-Analysis ist die Monographie [435] ihres Begründers A. Robinson; [6] enthält eine kompakte, präzise Einführung zusammen mit vielen Anwendungen in der mathematischen Physik. Eine didaktisch aufbereitete, gut lesbare elementare Einführung stammt von D. Laugwitz [343].

die Zwecke der Physik verwendbar. Man kann jedoch spezielle Klassen von Potentialen angeben, für die

$$\psi(x, t) = e^{-it\hat{H}} \psi(x, t_0) \approx \left( \frac{4\pi it}{n} \right) \int_{*\mathbb{Z}^{dn}} e^{i*S(\gamma_n+x),t} * \psi(\gamma_n(0) + x) * d\gamma_n$$

gilt; mit anderen Worten, man erhält in solchen Fällen eine Lösung der Schrödingergleichung

$$i \frac{\partial \psi}{\partial t} = \hat{H} \psi$$

mit Anfangsbedingung  $\psi(x, t_0) = \varphi(x)$  in der Form

$$\psi(x, t) = \text{st} \left[ \left( \frac{4\pi it}{n} \right) \int_{*\mathbb{Z}^{dn}} e^{i*S(\gamma_n+x),t} * \varphi(\gamma_n(0) + x) * d\gamma_n \right]$$

für ein geeignetes  $n \in *\mathbb{N} \setminus \mathbb{N}$ . Darüberhinaus lassen sich halbklassische Näherungslösungen der Schrödingergleichung im Rahmen der Methode der stationären Phase mit Mitteln der Non-Standard-Analysis formulieren, indem man  $\hbar$  durch ein infinitesimales  $\hbar \approx 0$  ersetzt. Ausführliche Informationen zu Pfadintegralen und Non-Standard-Analysis liefern [346], [347], [348] und [382] sowie mit zusätzlicher Berücksichtigung der Dirac-Gleichung auch [383] und [384]<sup>101</sup>.

### 2.7.4 Pfadintegrale im Phasenraum

Als vierte Kategorie seien abschließend einige Pfadintegral-Konstruktionen im Phasenraum erwähnt. In Abschnitt 2.2.5 wurde beschrieben, wie der quantenmechanische Propagator  $K(q_a, t_a; q_b, t_b) = \langle q_a, t_a | q_b, t_b \rangle$ , also die Wahrscheinlichkeit, daß ein Quantenobjekt, das sich zum Zeitpunkt  $t_a$  am Ort  $q_a$  befindet, zum späteren Zeitpunkt  $t_b$  am Ort  $q_b$  angetroffen wird, als heuristisches Pfadintegral im Phasenraum formuliert werden kann, etwa in der Gestalt

$$K(q_a, t_a; q_b, t_b) = \left( \frac{1}{2\pi} \right)^d \int \int_{\mathcal{P}} \exp \left\{ i \int_{t_a}^{t_b} [p(t) \dot{q}(t) - H(p(t), q(t), t)] dt \right\} \mathcal{D}p \mathcal{D}q; \quad (2.69)$$

dabei ist  $\mathcal{P}$  der Raum der stetigen Pfade  $(q, p)$  im Phasenraum mit  $q(t_a) = q_a$ ,  $q(t_b) = q_b$  und  $H$  die klassische Hamilton-Funktion des Systems. Natürlich ist auch diese Größe nicht definiert. Wir betrachten kurz einige mathematische Konsolidierungsversuche; für Details sei dabei auf die jeweils zitierten Arbeiten verwiesen.

Das erste Beispiel stammt von C. DeWitt-Morette, A. Maheshwari und B. Nelson; hierbei wird der Ausdruck (2.69) als *Gaußsches Pfadintegral* definiert [119]. Dieses Konzept funktioniert für die sehr allgemeine Situation eines Konfigurationsraums, der in

<sup>101</sup>An dieser Stelle seien auch die Arbeiten von A. Sloan erwähnt, der eine Non-Standard-Variante der Trotter-Produktformel hergeleitet hat [482], [483].

Form einer  $d$ -dimensionalen differenzierbaren Mannigfaltigkeit  $\mathcal{M}$  gegeben ist.  $T\mathcal{M}$  sei das Tangentialbündel über  $\mathcal{M}$ ; es hat die Struktur einer  $2d$ -dimensionalen differenzierbaren Mannigfaltigkeit. Den zugehörigen Phasenraum erhält man als Kotangentialbündel  $T^*\mathcal{M}$  über  $\mathcal{M}$ ; dieser ist folglich ebenfalls eine  $2d$ -dimensionale differenzierbare Mannigfaltigkeit. Außerdem sei  $(x, y)$  mit  $x : [t_a, t_b] \rightarrow \mathcal{M}$  und  $y : [t_a, t_b] \rightarrow \mathcal{M}^*$  ein beliebiger Pfad im Phasenraum,  $L : T\mathcal{M} \times [t_a, t_b] \rightarrow \mathbb{R}$  die Lagrange-Funktion und  $H : T^*\mathcal{M} \times [t_a, t_b] \rightarrow \mathbb{R}$  die Hamilton-Funktion des betrachteten Systems. Das Wirkungsfunktional im Phasenraum lautet dann

$$S(x, y) = \int_{t_a}^{t_b} [\langle y(t), dx(t) \rangle - H(y(t), x(t), t)] dt.$$

(2.69) nimmt die Form

$$K(q_a, t_a; q_b, t_b) = \int_{\mathcal{P}} e^{iS(x, y)} dw(x, y)$$

an; dabei ist  $w$  eine Gaußsche Prodistribution auf  $\mathcal{P}$ . Diese erhält man folgendermaßen: Mit Hilfe der Kovarianzfunktion

$$G(r, s) = \begin{pmatrix} 2(\theta(s-r)(r-t_a) + \theta(r-s)(s-t_a)) & \theta(r-s) \\ \theta(s-r) & 0 \end{pmatrix}$$

und der damit gebildeten Bilinearform

$$\mathbf{W}(\mu, \nu) = \int_{t_a}^{t_b} \int_{t_a}^{t_b} \sum_{\alpha=1}^2 \sum_{\beta=1}^2 G^{\alpha\beta}(r, s) d\mu_{\alpha}(r) d\nu_{\beta}(s)$$

definiert man die Varianz  $\mathbf{W}(\mu) := \mathbf{W}(\mu, \mu)$ . Die Prodistribution  $w$  ist definiert durch ihre Fourier-Transformierte

$$\mathcal{F}w(\mu) = N e^{-i\mathbf{W}(\mu)},$$

mit einem geeigneten Normierungsfaktor  $N$ . Ein ähnliches Konzept wurde von M. M. Mizrahi vorgeschlagen; es stellt eine Verallgemeinerung des Zugangs über Promaße dar [368], [369].

Auf kohärenten Zuständen basiert eine Konstruktion von Ingrid Daubechies und John Klauder [112]. Dazu seien  $\{|p, q\rangle \mid t^*\mathcal{M}\}$  die kohärenten Zustände im Phasenraum und  $\hat{H}$  der Hamilton-Operator des betrachteten quantenmechanischen Systems. Dieser läßt sich in der Form

$$\hat{H} = \left(\frac{1}{2\pi}\right)^d \iint |q, p\rangle h(q, p) \langle q, p| dq dp$$

darstellen; das kohärente Zustandssymbol  $h(q, p)$  wird gleichzeitig durch diese Relation definiert<sup>102</sup>. Außerdem sei  $\mathcal{C}$  der Raum der stetigen Pfade  $(q(t), p(t))$  im Phasenraum mit festen Anfangs- und Endwerten  $(q(t_a), p(t_a)) = (q_a, p_a)$ ,  $(q(t_b), p(t_b)) = (q_b, p_b)$ . Dann

<sup>102</sup>Details hierzu und allgemein zu kohärenten Zuständen findet man in [309].

lautet die Darstellung der Matrix-Elemente des Zeitentwicklungsoperators als Klauder-Daubechies-Pfadintegral

$$\begin{aligned} \langle q_b, p_b | e^{-i(t_b-t_a)\hat{H}} | q_a, p_a \rangle &= \left( \frac{1}{2\pi} \right)^d \lim_{\tau \searrow 0} e^{C(t_b-t_a)/\tau} \times \\ &\times \iint_{\mathcal{C}} \exp \left\{ i \int_{t_a}^{t_b} \left[ \frac{1}{2} (\dot{q} p - q \dot{p}) - h(q, p) \right] dt \right\} \times \\ &\times \exp \left[ -\tau \int_{t_a}^{t_b} \frac{1}{2} (\dot{q}^2 + \dot{p}^2) dt \right] \mathcal{D}q \mathcal{D}p. \end{aligned}$$

Dabei ist  $\tau$  ein Regularisierungsparameter für die Zeit und  $C$  eine Normierungskonstante. Der dritte, rein reelle Gaußsche Phasenfaktor in diesem Ausdruck ist ein Wiener-Maß im Phasenraum, durch welches das für sich allein nicht definierte rein imaginäre Gaußsche Integral für alle  $\tau > 0$  regularisiert wird. Anschaulich repräsentiert es eine Brownsche Bewegung des Systems in einer euklidischen Hintegrundgeometrie, die der gekrümmten, beispielsweise Riemannschen Geometrie der Phasenraum-Mannigfaltigkeit unterlegt ist. Interessant daran ist insbesondere auch die hierdurch beschriebene Mischung aus quantenmechanischem und stochastischem Verhalten, was unter anderem in der Quantengravitation von Bedeutung sein könnte.

Es ist auch möglich, den Begriff der unendlichdimensionalen oszillatorischen Integrale<sup>103</sup> auf den Phasenraum zu erweitern, allerdings funktioniert das problemlos nur für eine sehr spezielle Klasse von Potentialen [7]. Der Raum stetiger Pfade im Phasenraum, der hier zum Einsatz kommt, ist der Hilbertraum  $\mathcal{H}_t \times \mathcal{L}_t$ ; dieser besteht aus dem Raum  $\mathcal{H}_t$  der absolut stetigen Funktionen  $q : [0, t] \rightarrow \mathbb{F}^d$  mit  $q(0) = 0$  und  $\dot{q} \in \mathcal{L}^2([0, t]; \mathbb{F}^d)$  und dem Raum  $\mathcal{L}_t \equiv \mathcal{L}^2([0, t]; \mathbb{F}^d)$ . Der Hamilton-Operator des betrachteten Systems sei von der Form

$$\hat{H} = \hat{P}^2 + V_1(\hat{Q}) + V_2(\hat{P})$$

mit  $V_1 \in \mathcal{F}(\mathbb{F}^d)$  und  $\int_0^t V_2(p(t')) dt' \in \mathcal{F}(\mathcal{L}_t)$ , außerdem sei  $\varphi \in \mathcal{F}(\mathbb{F}^d)$ . Dann ist

$$\begin{aligned} \psi(x, t) &= \widetilde{\iint_{\mathcal{H}_t \times \mathcal{L}_t}} \exp \left\{ i \left[ \int_0^t \dot{q}(t') p(t') dt' - \frac{1}{2} \int_0^t [p(t')]^2 dt' \right] \right\} \times \\ &\times \exp \left\{ -i \int_0^t [V_1(q(t') + x) + V_2(p(t'))] dt' \right\} \varphi(q(0) + x) dq dp \end{aligned}$$

eine Lösung der Schrödingergleichung

$$\dot{\psi} = -i\hat{H}\psi$$

<sup>103</sup>Siehe Abschnitt 2.5.

mit der Anfangsbedingung

$$\psi(x, 0) = \varphi(x), \quad x \in \mathbb{R}^d.$$

Eine besondere Eigenschaft dieses Zugangs ist seine spezielle Eignung im Zusammenhang mit der Betrachtung semiklassischer Grenzwerte unter Verwendung der Methode der stationären Phase.

Damit beenden wir unsere Rundreise durch die Welt der mathematisch rigoros definierten Pfadintegrale. Der hier dargebotene Überblick ist natürlich bei weitem nicht vollständig; er sollte jedoch durchaus genügen, um sich eine Vorstellung des formalen Aufwands zu machen, der generell erforderlich ist, um Pfadintegralen eine wohldefinierte Bedeutung zu geben.

## 2.8 Erkenntnistheoretische Einordnung der Pfadintegrale

Fassen wir die Abschnitte 2.4 bis 2.7 zusammen, können wir zu folgendem Fazit gelangen: Bei der mathematisch korrekten Definition heuristischer Feynmanscher Pfadintegrale für die nichtrelativistische Quantenmechanik steht man vor der Wahl, dieses entweder allgemeingültig zu tun, wobei man jedoch auf Wiener-Integrale ausweichen muß und damit den Rahmen der Quantenmechanik genaugenommen verläßt, oder aber auf eine Formulierung für beliebige Systeme zu verzichten, da man für alle anderen Formulierungen teilweise stark einschränkende Bedingungen an die verwendeten Potentiale stellen muß. Das hat weitreichende Konsequenzen für die Interpretation dieses Zugangs aus erkenntnistheoretischer Sicht, insbesondere, wenn man sich einer philosophischen Weltanschauung verpflichtet fühlt, die am besten mit dem Begriff *Theorienrealismus* umschrieben ist. Hierbei nehmen wir an, daß die Entitäten, die in unseren physikalischen Theorien vorkommen, sich zumindest indirekt auf real existierende Objekte in unserer Außenwelt zurückführen lassen. Im Rahmen dieser Auffassung soll die Quantenmechanik als ein Theorienrahmen gesehen werden, dessen Grundprinzipien gleichzeitig grundlegende Eigenschaften der Natur darstellen. Beispiele für solche Grundprinzipien sind etwa der Welle-Teilchen-Dualismus, die Unschärferelationen oder die kanonische Quantisierung. Jeder Zugang zur Quantenmechanik, der den Anspruch erhebt, ein fundamentaler Zugang mit entsprechendem realistischem Interpretationspotential zu sein, muß daher auf solche Grundgesetzmäßigkeiten der Natur zurückführbar und insbesondere auch – wenigstens im Prinzip – allgemein anwendbar sein. Gleichzeitig macht die offenkundig mathematische Natur der physikalischen Welt – wiederum wenigstens im Prinzip – eine mathematisch absolut saubere Formulierung physikalischer Gesetze unumgänglich.

Diese Eigenschaften treffen auf Pfadintegral-Zugänge zur Standard-Quantenmechanik, gleichgültig in welcher Variante, *nicht zu*. Sie sind entweder mathematisch nicht existent, nur für spezielle Systeme anwendbar oder aber erst nach formalen Tricks überhaupt quantenmechanisch interpretierbar. Damit läßt sich die Quantisierung mit Pfadintegralen nicht als Anwendung eines fundamentalen Grundprinzips der Natur auffassen, wie es im obigen Sinn bei anderen Zugängen wie etwa der kanonischen Quantisierung im Rahmen der Hilbertraum-Quantenmechanik der Fall ist. Pfadintegrale dürfen folglich auch nicht in der Weise realistisch interpretiert werden, wie es bei vielen anderen physikalischen



Begriffen und Konzepten denkbar ist. Beispielsweise ist die Vorstellung, quantenmechanische Superpositionszustände seien Zustände, in denen ein System zwei oder mehr, eventuell sogar unendlich viele klassisch sich widersprechende Eigenschaften gleichzeitig annimmt, zulässig. Dagegen ist es abwegig zu behaupten, ein quantenmechanisches System durchlaufe mit jeweils zu jeder Zeit gleichzeitig existierenden Werten von Ort und Impuls gleichzeitig alle möglichen stetigen Pfade von einem Anfangs- zu einem Endpunkt; diese Beschreibung folgt jedoch unweigerlich, wenn man Pfadintegrale realistisch interpretiert<sup>104</sup>.

Konsequenterweise ist man gezwungen, eine philosophisch zurückhaltendere Einschätzung zu treffen: Pfadintegrale sind für die nichtrelativistische Quantenmechanik ein rein formaler Weg zur Berechnung von Dingen wie Übergangsamplituden oder Wellenfunktionen, sie repräsentieren jedoch nicht ein Grundprinzip der Natur und können folglich auch nicht als fundamentale Quantisierungsmethode aufgefaßt werden. Daß Pfadintegrale in der Standard-Quantenmechanik im Vergleich zu herkömmlichen Techniken ein untaugliches bis bestenfalls äußerst umständliches Verfahren für konkrete Berechnungen darstellen, ist hier nicht einmal von wesentlicher Bedeutung. Denn sobald man eine physikalische Interpretation der gefundenen Resultate möchte, bleibt einem nichts anderes übrig, als auf andere Zugänge zurückzugreifen.

Anders sieht die Sache in der Quantenfeldtheorie aus. Hier darf man den euklidischen Ansatz, der in Abschnitt 2.4.3.3 erwähnt wurde, sehr wohl als grundlegendes physikalisches Konzept mit tiefgreifender konzeptueller Bedeutung auffassen. Das ist unter anderem und nicht zuletzt an der Bedeutung erkennbar, welche die Technik der Funktionalintegration für die konstruktive Quantenfeldtheorie hat, auch wenn dieses Gebiet wie erwähnt derzeit bestenfalls als Forschungsprogramm und noch in keiner Weise als Theorie vorliegt. Auch Zusammenhänge, die zwischen axiomatischen Formulierungen wie etwa derjenigen von Wightmann und der Euklidischen Quantenfeldtheorie bestehen, bestärken diese realistische Sichtweise<sup>105</sup>. Wiederum ist das hier im Vergleich zu oben umgekehrt laufende technische Argument der teilweise haushohen Überlegenheit von Funktionalintegral-Methoden gegenüber gewöhnlichen operatortheoretischen Methoden bei konkreten quantenfeldtheoretischen Berechnungen nicht von wesentlichem Gewicht. Wichtiger ist die erwähnte Bedeutung der Euklidischen Quantenfeldtheorie als offensichtlich grundlegendem Bestandteil der Prinzipien der physikalischen Natur.

## 2.9 Quantenmechanik und rotierende Zeiger

Wir kommen nunmehr wieder zurück zum eigentlichen Gegenstand des vorliegenden Buches, der Schulphysik. Im Zentrum der Betrachtung steht dabei der bereits angesprochene Versuch eines anschaulichen Zugangs zur Quantenmechanik, der unter der Bezeichnung „Zeigerformalismus“ bekannt geworden ist. Die Sache hat zwar einen ganz anderen physikalischen Hintergrund, ist jedoch nicht weniger ärgerlich als die oben beschriebene Energieträgerphysik; im Gegenteil, während letztere wenigstens teilweise physikalisch

<sup>104</sup>Es sei beiläufig erwähnt, daß die Situation auch bei nichtrealistischen Auffassungen der Physik, wie zum Beispiel konstruktivistischen oder instrumentalistischen Weltanschauungen, nicht besser aussieht. Denn auch dort bleibt das Problem der fehlenden Allgemeingültigkeit natürlich unverändert bestehen.

<sup>105</sup>Für Details siehe beispielsweise [499].

auf sicherem Boden steht, beruht der Zeigerformalismus in dieser Verwendung im günstigsten Fall auf einer Mißinterpretation quantenmechanischer Konzepte. Dies ist schon an dem selbsterhobenen Anspruch erkennbar, einen Zugang zur Quantenmechanik zu liefern, der „so klassisch wie möglich erfolgt, ohne dabei physikalisch falsch zu sein“ [38]. Das ist jedoch ein Widerspruch und damit unmöglich, denn die Quantenmechanik zeichnet sich gerade dadurch aus, *nichtklassische Physik zu repräsentieren*. Folglich gibt es keinerlei Spielraum für irgendwelche klassischen Vorstellungen, und seien diese auch noch so gering ausgeprägt<sup>106</sup>. Dabei soll gar nicht in Abrede gestellt werden, daß der Zeigerformalismus an anderer Stelle, nämlich in der klassischen Wellenoptik und hier speziell bei der Bestimmung von Interferenzmustern aufgrund von Beugung an Mehrfachspalten, durchaus seine Berechtigung haben mag. Aber dort sollte er auch bleiben.

Grundlage des Zeigerformalismus in der Quantenmechanik sind die Feynmanschen Pfadintegrale. Daß man in der gymnasialen Oberstufe gerade diesen schwierigsten aller Zugänge zur Quantenmechanik wählt, der im übrigen seine praktische Bedeutung in der Quantenfeldtheorie und der statistischen Mechanik hat und noch dazu mathematisch je nach Handhabung entweder nicht definiert oder aber überaus aufwendig ist, erscheint für sich genommen schon äußerst rätselhaft. Mag man das womöglich, wenn auch aus Sicht des Autors unberechtigterweise, als mathematische Spitzfindigkeiten abtun, so bleibt das für die Belange der Schule noch weitaus bedeutendere Problem, daß sich bei den Schülerinnen und Schülern schwerwiegende Fehlinterpretationen einstellen, etwa dadurch, daß sich die ihnen zuvor im Unterricht mühevoll ausgetriebene Vorstellung, Quantenobjekte würden sich auf Bahnen bewegen, wahrscheinlich umgehend erneut einstellt. Davon abgesehen handelt es sich auch hier wieder um einen Fall der Schaffung eigener Begrifflichkeiten durch die Schulphysik, was aus bereits dargelegten Gründen nicht zu akzeptieren ist. – Bevor wir nun jedoch den gesamten Inhalt des restlichen zweiten Kapitels gleich vorwegnehmen, kehren wir zu einer geordneten Vorgehensweise zurück und organisieren dieselbe in drei Schritte. Zunächst betrachten wir kurz den Ursprung des Zeigerformalismus, der in der elementaren Wellenoptik zu finden ist, diskutieren sodann mit dessen Verwendung im Quantenmechanik-Unterricht der gymnasialen Oberstufe den zentralen Anlaß zur Kritik, welche das vorliegende Kapitel übt, und formulieren schließlich anschließend diese Kritik im Detail.

### 2.9.1 Zeigerformalismus und klassische Wellenoptik

Der Zeigerformalismus stammt ursprünglich und wenig erstaunlich von Richard Feynman und wurde von ihm in seinem Buch [167] 1985 erstmals vorgeführt<sup>107</sup>. Bevor wir darauf zu sprechen kommen, diskutieren wir das Konzept zunächst in einem anderen, eingeschränkten Kontext, in welchem es physikalisch unbedenklich eingesetzt werden kann und in Baden-Württemberg aufgrund spezifischer Lehrplaninhalte auch verbreitet eingesetzt wird, nämlich in der klassischen Wellenoptik und dort speziell bei der Bestimmung

<sup>106</sup>Der zitierte Charakterisierungsversuch ist einem populärwissenschaftlichen Quantenmechanik-Buch von T. Görnitz entlehnt [194], der damit direkt Feynmans Pfadintegrale beschreibt. In diesem Kontext ist der Spruch natürlich genauso abwegig.

<sup>107</sup>Gegenstand dieses Buches ist eine populärwissenschaftliche Darstellung der von Feynman selbst wesentlich mitentwickelten Quantenelektrodynamik, nicht jedoch der Standard-Quantenmechanik, um die es ausschließlich in der gymnasialen Oberstufe gehen kann. Dieses fast immer übersehene, aber sehr wesentliche Detail wird uns noch beschäftigen.

der Intensitätsverteilung von Interferenzmustern bei der Beugung von Licht an Spalten und Mehrfachspalten.

In der Wellenoptik ist die Intensität des Lichts bekanntlich proportional zum Betrag der Amplitude der elektrischen Feldstärke  $\vec{\mathcal{E}}$  oder der magnetischen Feldstärke  $\vec{\mathcal{B}}$  des betrachteten elektromagnetischen Feldes<sup>108</sup>. Es genügt folglich, eine der beiden Feldstärken zu betrachten, etwa die elektrische. Beschreiben wir diese als komplexe Welle

$$\vec{\mathcal{E}}(\vec{x}, t) = \hat{\mathcal{E}} e^{i(\vec{k}\vec{x} - \omega t)},$$

so erhalten wir bei Überlagerung zweier elektromagnetischen Wellen die Feldstärke

$$\vec{\mathcal{E}}(\vec{x}, t) = \vec{\mathcal{E}}_1(\vec{x}, t) + \vec{\mathcal{E}}_2(\vec{x}, t) = \hat{\mathcal{E}}_1 e^{i(\vec{k}_1\vec{x} - \omega_1 t)} + \hat{\mathcal{E}}_2 e^{i(\vec{k}_2\vec{x} - \omega_2 t)}$$

und durch Quadrieren des Feldstärkevektors weiter

$$\begin{aligned} [\vec{\mathcal{E}}(\vec{x}, t)]^2 &= \vec{\mathcal{E}}(\vec{x}, t) \vec{\mathcal{E}}^*(\vec{x}, t) \\ &= \hat{\mathcal{E}}_1^2 + \hat{\mathcal{E}}_2^2 + \hat{\mathcal{E}}_1 \hat{\mathcal{E}}_2 \{ e^{i[(\vec{k}_1 - \vec{k}_2)\vec{x} + (\omega_1 - \omega_2)t]} + e^{i[(\vec{k}_2 - \vec{k}_1)\vec{x} + (\omega_2 - \omega_1)t]} \} \\ &= \hat{\mathcal{E}}_1^2 + \hat{\mathcal{E}}_2^2 + 2 \hat{\mathcal{E}}_1 \hat{\mathcal{E}}_2 \cos [(\vec{k}_1 - \vec{k}_2)\vec{x} + (\omega_1 - \omega_2)t] \\ &= \hat{\mathcal{E}}_1^2 + \hat{\mathcal{E}}_2^2 + 2 \hat{\mathcal{E}}_1 \hat{\mathcal{E}}_2 \cos \Delta\phi, \end{aligned}$$

mit der *Phasendifferenz*  $\Delta\phi \equiv (\vec{k}_1 - \vec{k}_2)\vec{x} + (\omega_1 - \omega_2)t$ . Für die Intensität der Superposition finden wir damit

$$\mathcal{I} = \mathcal{I}_1 + \mathcal{I}_2 + 2 \sqrt{\mathcal{I}_1 \mathcal{I}_2} \cos \Delta\phi.$$

Betrachten wir die Superposition zweier identischer Wellen, so können wir  $\mathcal{I}_1 = \mathcal{I}_2$  setzen und erhalten

$$\mathcal{I} = 2 \mathcal{I}_1 (1 + \cos \Delta\phi)$$

oder unter Verwendung der Relation  $1 + \cos \alpha = 2 \cos^2(\alpha/2)$

$$\mathcal{I} = 4 \mathcal{I}_1 \cos^2(\Delta\phi/2).$$

Diese Intensitätsverteilung beschreibt das Interferenzmuster in einem idealen Zwei-Wege-Interferometer, etwa einem solchen mit einem Doppelspalt mit infinitesimal schmalen Spalten.

Für den Zeigerformalismus macht man nun davon Gebrauch, daß aufgrund der hohen Frequenzen, die in der Optik auftreten, Terme der Form  $e^{i(\vec{k}\vec{x} - \omega t)}$  in guter Näherung durch  $e^{i\phi}$  ersetzt werden können, wobei  $\phi$  der Phasenwinkel der Welle ist. Man geht also von einer zeitunabhängigen, festen Phasenbeziehung zwischen den beiden zu überlagernden Wellen aus, oder kürzer gesagt, man setzt voraus, daß diese *kohärent* sind. Da  $e^{i\phi}$  eine komplexe Zahl mit Betrag 1 und  $\phi$  proportional zur optischen Weglänge ist, kann man  $\hat{\mathcal{E}} e^{i(\vec{k}\vec{x} - \omega t)}$  als *rotierenden Vektor der Länge  $\hat{\mathcal{E}}$*  betrachten. Stellt man sich darunter einen rotierenden Zeiger vor, sieht man, was der hier diskutierten Angelegenheit ihren Namen gab.

Zur Bestimmung von Intensitätsverteilungen bei Interferenzmustern kann man somit im Prinzip folgendermaßen vorgehen: Man ordnet jedem Lichtweg einen „Zeiger“, das

<sup>108</sup>Wer sich hier für Details interessiert, findet in [487] erschöpfende Auskunft.

heißt also einen Pfeil mit willkürlich gewählter einheitlicher Länge zu, der in vorgegebener einheitlicher Richtung, etwa im Gegenuhrzeigersinn, rotiert und jeweils eine volle Umdrehung von  $360^\circ$  ausführt, wenn sich das Licht um eine Wellenlänge weiterbewegt. Je nach Länge der Lichtwege erhält man am Ort der Beobachtung für jeden Lichtweg einen Zeiger mit jeweils unterschiedlicher Position. Addiert man die Zeiger vektoriell und quadriert den Betrag des resultierenden Zeigers, erhält man die relative Intensität des Interferenzmusters an der betrachteten Stelle. Da es dabei nur um Phasendifferenzen geht, muß man diese Prozedur natürlich nicht wirklich durchführen<sup>109</sup>. Es genügt, eine willkürliche Nullage festzulegen und mit den Zeigerpositionen relativ dazu die Phasendifferenzen zu markieren, die sich aus den Gangunterschieden der beteiligten optischen Wege ergeben. Damit lassen sich insbesondere die Interferenzmuster bei mehr als zwei Spalten unvergleichlich einfacher gewinnen, als das bei analytischer Berechnung der Fall ist. In dieser Form taucht die Sache verbreitet auch in Schulbüchern auf<sup>110</sup>. Man kann allerdings sehr wohl anzweifeln, ob es irgendwelchen Sinn hat, im Physikunterricht der Oberstufe Interferenzmuster für Vierfach-, Achtfach-, Sechzehnfach- oder wer weiß was für Spalte zu bestimmen. Seit 2004 muß man zwar damit rechnen, daß das Thema in Baden-Württemberg im Abitur auftauchen könnte; insbesondere Fragestellungen in Bezug auf die genaue Verteilung von Haupt- und Nebenmaxima bei Mehrfachspalten sind inzwischen sehr in Mode gekommen. Da sich solche Aufgaben jedoch darauf beschränken, entsprechende Interferenzbilder nur anzugeben oder zuzuordnen, liegt damit kein Grund vor, mit den Zeigern einen aufwendigen neuen Formalismus zu entwickeln, der den Beweis seines Bildungswerts schuldig bleiben muß.

### 2.9.2 Zeigerformalismus und Quantenmechanik

Bevor wir uns kritisch mit dem Einsatz des soeben beschriebenen Zeigerformalismus für die Quantentheorie in der Oberstufe auseinandersetzen, müssen wir zunächst die Grundideen solcher Unterrichtskonzepte kennenlernen. Diese Grundideen gehen wiederum ausnahmslos auf Feynman zurück und sind eine direkte Anwendung des Pfadintegralformalismus in der ursprünglichen, heuristischen Form, wie sie in Abschnitt 2.2 beschrieben wurden<sup>111</sup>. Auf den ersten Blick scheint der Zeigerformalismus jedoch mit Pfadintegralen nichts zu tun zu haben. Inwieweit hier doch ein Zusammenhang besteht, wird erst klar, wenn man den Pfadintegral-Ansatz anschaulich deutet.

Dazu startet man mit der Pfadintegral-Darstellung (2.14) des Schrödinger-Propagators aus Abschnitt 2.2.2,

$$K(x_b, t_b; x_a, t_a) = \int_c e^{iS(x(t))} \mathcal{D}x(t).$$

<sup>109</sup>Den Protagonisten des Zeigerformalismus ist auch diese Mühe nicht zuviel; von J. Küblbeck stammt die Konstruktion eines Phasenrads oder Steckerlrads, wie er es selbst nennt, das man aus Pappe basteln und durch Abfahren der Lichtwege dazu verwenden kann, den „Zeigerstand“ und damit die relative Phasenlage an jedem gewünschten Ort auf den Lichtwegen zu ermitteln. Details findet man in [325], [326] und [328].

<sup>110</sup>Siehe beispielsweise [38] und [39].

<sup>111</sup>Dieser Zusammenhang dürfte vielen Anhängern des Zeigerformalismus in der Schulphysik kaum klar sein; er wird in einschlägigen Darstellungen des Gegenstands gar nicht [328] bis nur andeutungsweise [37] erwähnt. Die einzige Ausnahme ist [443]. Daß Feynman in [167] nichts explizites über Pfadintegrale schreibt, steht ihm selbst natürlich uneingeschränkt zu.

Der Schrödinger-Propagator in dieser Form ist die Wahrscheinlichkeitsamplitude der bedingten Wahrscheinlichkeit, das betrachtete Quantenobjekt im Konfigurationsraum zum Zeitpunkt  $t_b$  am Ort  $x_b$  anzutreffen, unter der Voraussetzung, daß es zum Zeitpunkt  $t_a$  am Ort  $x_a$  angetroffen wurde. Dabei ist

$$S(x(t)) = \int_{t_a}^{t_b} L(x(t), \dot{x}(t), t) dt$$

die klassische Wirkung des Systems längs des Pfads  $x(t)$  im Konfigurationsraum. Nun hat der Integrand  $e^{iS}$  eines solchen Pfadintegrals gerade wieder die Form einer komplexen Zahl mit Einheitsbetrag, wie sie bereits im vorigen Abschnitt auftrat, und kann folglich als Pfeil mit Einheitslänge in der komplexen Ebene betrachtet werden, dessen Phasenwinkel durch die Wirkung  $S$  gegeben ist. Ändert sich  $S$ , so ändert sich der Phasenwinkel, das heißt, der Pfeil rotiert um den Ursprung. Die Rotationsgeschwindigkeit wird dabei von der Lagrange-Funktion  $L$  bestimmt, da  $S$  als Integral über diese definiert ist. Im einfachsten Fall eines freien Teilchens ist  $V = 0$  und  $L$  somit konstant, sodaß die Pfeile auf allen Pfaden mit konstanter Winkelgeschwindigkeit rotieren. Im allgemeinen Fall mit  $V \neq 0$  ist die Winkelgeschwindigkeit orts- und zeitabhängig. Wandert ein solcher Pfeil längs eines Pfades  $x(t)$  von  $x_a$  nach  $x_b$ , so wird er sich gemäß dem zeitlichen Verhalten von  $S$  um den Ursprung drehen. Für jeden möglichen Weg von  $x_a$  nach  $x_b$  erhält man daher eine charakteristische Position des zugehörigen Pfeils bei  $x_b$ . Sehr salopp gesagt ist ein Pfadintegral auch nichts anderes als eine Summe. Die Integration von  $e^{iS}$  über alle möglichen Pfade von  $x_a$  nach  $x_b$  kann man sich daher stark vereinfacht als eine Vektorsumme über alle Pfeile vorstellen, wie man sie nach Durchlaufen der Pfade jeweils am Punkt  $x_b$  erhält, mit ihren jeweils unterschiedlichen Richtungen. Das anschauliche Ergebnis des Pfadintegrals ist somit wieder ein Pfeil. Entsprechend der physikalischen Bedeutung des Propagators ist das Betragsquadrat dieses resultierenden Vektors proportional zur bedingten Wahrscheinlichkeit, das betrachtete Quantenobjekt im Konfigurationsraum zum Zeitpunkt  $t_b$  am Ort  $x_b$  anzutreffen, unter der Voraussetzung, daß es zum Zeitpunkt  $t_a$  am Ort  $x_a$  angetroffen wurde.

Dieses anschauliche, aber dafür natürlich extrem vereinfachte Bild wird nun zum Anlaß genommen, den im vorigen Abschnitt beschriebenen Zeigerformalismus der klassischen Wellenoptik auf die Quantenmechanik loszulassen, indem man die rotierenden Einheitsvektoren  $e^{iS}$  in der komplexen Ebene als Zeiger auffaßt. Danach ergäbe sich folgende Vorschrift zur Bestimmung der Wahrscheinlichkeit  $\mathfrak{P}(x_b, t_b; x_a, t_a)$ , ein Quantenobjekt, das zur Zeit  $t_a$  bei  $x_a$  angetroffen wurde, zur Zeit  $t_b$  bei  $x_b$  anzutreffen:

1. Man ermittle *alle* stetigen Wege von  $x_a$  nach  $x_b$ ,
2. stelle sich vor, daß ausgehend von einer zuvor festgelegten einheitlichen Startrichtung entlang eines jeden dieser Wege ein rotierender Zeiger entlangläuft, dessen Winkelgeschwindigkeit durch  $S$  festgelegt wird,
3. bestimme sodann für jeden Weg die Zeigerposition am Ort  $x_b$  und
4. addiere alle diese Zeiger vektoriell.  $\mathfrak{P}(x_b, t_b; x_a, t_a)$  ist dann proportional zum Betragsquadrat dieses resultierenden Vektors.



Natürlich ist das wiederum nur eine grob vereinfachte Veranschaulichung und insbesondere für praktische Belange völlig ungeeignet; für konkrete Anwendungen ist man darauf angewiesen, die dabei auftretenden Pfadintegrale auch tatsächlich auszurechnen, sofern man sich dazu entschieden hat, diese Methode zu wählen.

Um die vermeintliche Einfachheit der beschriebenen anschaulichen Interpretation der Pfadintegrale wenn schon nicht für die Anwendung auf reale physikalische Systeme, so doch wenigstens auf didaktisch reduzierte Spielzeugmodelle zu retten, geht man nun noch einen Schritt weiter. Man argumentiert hierbei, daß es in guter Näherung genügt, wenige spezielle Wege auszuwählen und nur diese bei der vektoriellen Addition der zugehörigen Zeiger zu berücksichtigen, da sich alle anderen insgesamt gerade wegheben. Genauer gesagt beschränkt man sich auf Versuchsanordnungen, die charakteristische Hindernisse wie Spaltöffnungen, reflektierende Flächen, Streuzentren oder dergleichen aufweisen, und betrachtet nur geradlinige Wege von  $x_a$  zu den Hindernissen einerseits sowie von den Hindernissen zu  $x_b$  andererseits; nur die Beiträge dieser Wege werden dann aufaddiert. Die Beiträge aller anderen Wege sollen sich gegenseitig auslöschen. Das bedeutet beispielsweise bei Interferenz am Doppelspalt<sup>112</sup> die ausschließliche Berücksichtigung der beiden Wege, die  $x_a$  und  $x_b$  jeweils über die beiden Spaltöffnungen geradlinig verbinden – womit man wieder genau bei der Situation der klassischen Optik aus dem vorigen Abschnitt angekommen ist (siehe Abbildung 2.5). Noch deutlicher wird das bei der Reflexion an einer ebenen Fläche. Hier soll man entsprechend der oben beschriebenen Taktik nur diejenigen Pfade berücksichtigen, die aus geradlinigen Verbindungen vom Ort der Quelle bei  $x_a$  zur Reflexionsfläche und ebensolchen von der Reflexionsfläche zum Ort der Beobachtung bei  $x_b$  bestehen (siehe Abbildung 2.6, wo einige Repräsentanten stellvertretend für alle dieser unendlich vielen Pfade dargestellt sind). Man kann zeigen, daß nur Pfade in der Nähe von  $x_5$  bis  $x_8$  bei Anwendung des Zeigerformalismus nennenswerte Beiträge liefern, und das um so mehr, je näher die Knickstelle des Pfades bei der Mitte zwischen  $x_6$  und  $x_7$  ist. Damit hat man das klassische Reflexionsgesetz „wiederentdeckt“. In analoger Weise läßt sich beispielsweise auch das Brechungsgesetz und allgemein die gesamte geometrische Optik rekonstruieren.

Natürlich verbirgt sich hinter diesen Überlegungen nichts anderes als das Verfahren der stationären Phase für semiklassische Näherungen, wonach sich bei Systemen mit klassischem Verhalten nur die Pfade in der unmittelbaren Umgebung des klassischen Weges durch konstruktive Interferenz verstärken und alle anderen Pfade sich durch destruktive Interferenz gegenseitig wegheben<sup>113</sup>. Daß dabei folglich klassische und quantenmechanische Physik durcheinandergeworfen werden, soll uns erst im nächsten Abschnitt beschäftigen; hier halten wir zunächst nur fest, daß die beschriebene, so unausgesprochen wie stark von der klassischen Wellenoptik motivierte Reduktion der Betrachtung aller stetiger Pfade zwischen zwei Punkten auf wenige geradlinige Verbindungslinien alles ist, was schließlich im Zeigerformalismus von der Berechnung von Pfadintegralen übrigbleibt.

Das bedeutet keineswegs, daß dabei ein für Schülerinnen und Schüler einfach zu handhabendes Konzept herauskommt. Im Gegenteil, die Hauptaufgabe, welche der Zeigerformalismus seinen Anhängern aufgibt, nämlich das Ermitteln der Zeigerpositionen am Ende der betrachteten Wege, ist nur in ganz einfachen Fällen leicht zu lösen und im allgemeinen ohne Hilfsmittel unmöglich. Hier kommen beispielsweise Kübelbecks in

<sup>112</sup>mit infinitesimal schmalen Spaltöffnungen, um ganz genau zu sein

<sup>113</sup>Siehe Abschnitt 2.2.4.

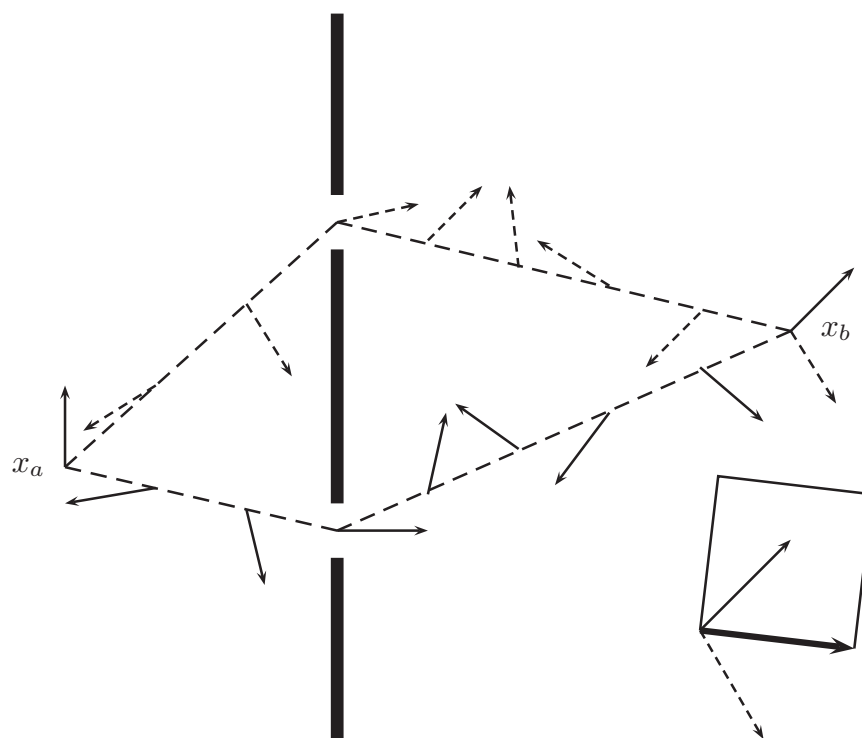


Abbildung 2.6: Rotierende Zeiger und vektorielle Zeigeraddition beim Doppelspaltexperiment. Die Fläche des Quadrats ist proportional zur Intensität des Interferenzmusters am Ort  $x_a$ .

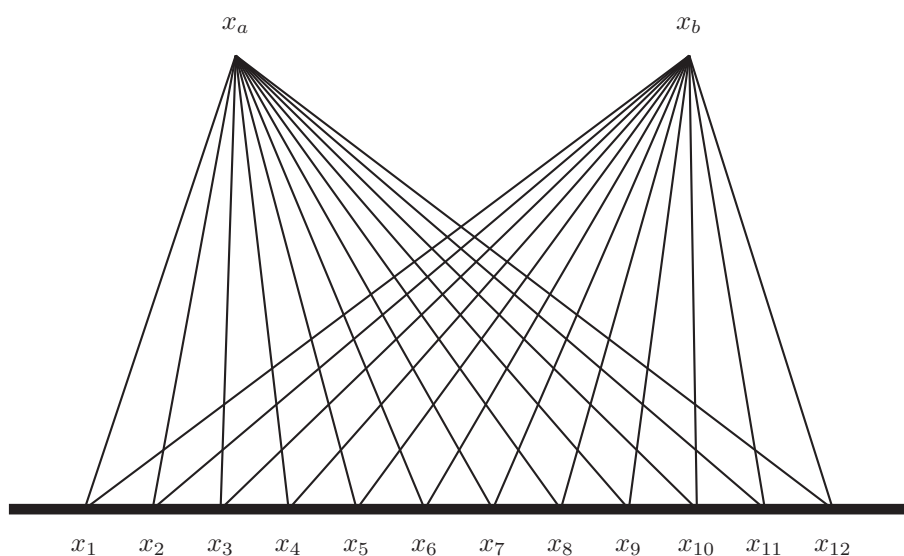


Abbildung 2.7: Zur Reflexion an einer ebenen Fläche



Anmerkung 109 bereits erwähntes Steckerlrad und insbesondere Computerprogramme zum Einsatz<sup>114</sup>, mit denen Zeigerpositionen und dazu die Beträge der Vektorsummen von Zeigern für einfache quantenmechanische Systeme berechnet werden können. Ob der teilweise beträchtliche Aufwand, der hierbei im Unterricht betrieben werden muß, zu rechtfertigen ist, darf wohl in Frage gestellt werden, worauf wir weiter unten wieder zurückkommen werden.

Im übrigen stellt sich damit unmittelbar die Frage, inwieweit man das ganze Konzept überhaupt als didaktisch vereinfachte Umsetzung des Pfadintegralformalismus betrachten kann. Zumindest bei Vorgängen, die nur eine endliche Anzahl von Streuzentren aufweisen, wie insbesondere Beugung an Mehrfachspalten und Gittern<sup>115</sup>, ist das einzig wirklich bedeutsame an den Zeigern ihre Eigenschaft, *komplexe Zahlen zu veranschaulichen*. Abgesehen davon, daß man dann die Frage stellen sollte, warum man im Physikunterricht der Oberstufe nicht gleich komplexe Zahlen anstelle eines aufwendigen Phantasieformalismus einführt<sup>116</sup>, bleibt damit der Eindruck, daß die korrekten Ergebnisse, die der Zeigerformalismus bei der Beschreibung von Interferenzeffekten bei der Beugung von Quantenobjekten an Gittern liefert, wesentlich dadurch begründet sind, daß man dabei die Beschreibung von klassischer Interferenz am Gitter eins zu eins imitiert. Insbesondere wird hier in keiner Weise deutlich, daß die Verwendung komplexer Zahlen in der Quantenmechanik unumgänglich, in der klassischen Wellenoptik dagegen ein bequemer mathematischer Trick ist. Natürlich ist auch hier in beiden Fällen die Methode der stationären Phase eine mögliche Erklärung auf ganz anderem Niveau; der Versuch, das im Unterricht des Gymnasiums zu thematisieren, wäre jedoch ziemlich Unsinn.

Die von seinen Anhängern ebenfalls als Bestandteil des Zeigerformalismus reklamierte Regel, bei ununterscheidbaren Wegen die Zeiger erst zu addieren und dann zu quadrieren, bei unterscheidbaren Wegen das jedoch in der umgekehrten Reihenfolge zu tun, ist in Wahrheit eine sehr viel fundamentalere Angelegenheit und natürlich *keine* Folgerung aus der Pfadintegral-Quantenmechanik, schon gar nicht aus der beschriebenen didaktisch reduzierten Zeigerversion. Was hier eigentlich gemeint ist, ist das Prinzip vom *Welle-Teilchen-Dualismus*, das inzwischen wieder in das Zentrum des quantenmechanischen Interesses gerückt ist. Knapp formuliert lautet es folgendermaßen:

*Je besser die Interferenzfähigkeit identischer Quantenobjekte ist, desto schlechter lassen die verschiedenen Wege, die sie nehmen könnten, voneinander unterscheiden und umgekehrt.*

Besondere Bedeutung erhält der Welle-Teilchen-Dualismus, weil er sich inzwischen als eigenständiges Grundprinzip der Quantenmechanik erwiesen hat und nicht, wie jahrzehntelang gedacht wurde, eine Konsequenz der Heisenbergschen Unschärferelation ist. Wir kommen in den Abschnitten 4.3 und 4.5.3 darauf zurück.

Eine stark verkürzte Version dieses Prinzips besagt, daß man zur Berechnung quantenmechanischer Wahrscheinlichkeitsdichten für Superpositionszustände die Komponenten-Wellenfunktionen der Superposition zuerst addieren und dann quadrieren muß, wenn die überlagerten Zustände ununterscheidbare Situationen beschreiben, und zuerst qua-

<sup>114</sup>Siehe unter anderem [37].

<sup>115</sup>Wobei stets wieder vom Grenzfall infinitesimal schmaler Spaltbreiten auszugehen ist.

<sup>116</sup>Wir kommen in Abschnitt 2.9.3.3 darauf zurück.

drieren und dann addieren muß, wenn die überlagerten Zustände unterscheidbare Situationen darstellen<sup>117</sup>. Das liefert im ersten Fall Interferenzterme und im zweiten nicht<sup>118</sup>. Da die Zeiger des gleichnamigen Formalismus genau besehen Phasen von Wellenfunktionen repräsentieren, gilt diese Additions-Quadrierungs-Regel auch für sie. Mit Pfadintegralen hat das unmittelbar gar nichts zu tun.

## 2.10 Zehn Argumente gegen den Zeigerformalismus

Der Einsatz des Zeigerformalismus bei der elementaren Einführung der Quantenmechanik, wie sie beispielsweise in Baden-Württemberg für die gymnasiale Oberstufe vorgeschrieben ist, wird mittlerweile von Schulbüchern wie Fortbildungen nahezu flächendeckend propagiert, sie wird in der Referendarsausbildung an den Seminaren häufig bevorzugt, und selbst der Bildungsplan im Ländle weist ausdrücklich darauf hin. Wir werden im vorliegenden Abschnitt ausführlich begründen, warum wir das für einen Irrweg halten. Wieder stellen wir unseren Standpunkt in drei Kategorien aufgeteilt vor, in eine physikalische, eine erkenntnistheoretische und eine fachdidaktische. In den folgenden drei Abschnitten werden wir unsere Argumente jeweils einzeln aufführen und anschließend ausführlich kommentieren.

### 2.10.1 Physikalische Aspekte

Zunächst betrachten wir die Verwendung des Zeigerformalismus im Schulunterricht im Hinblick auf dessen physikalische Ableitung aus dem Feynmanschen Pfadintegralformalismus. Die Tatsache, daß dieser aus mathematischer Sicht mehr als problematisch ist, blenden wir dabei bis zum nächsten Abschnitt vorerst aus.

#### 2.10.1.1 Pfadintegrale und nichtrelativistische Quantenmechanik

Wir beginnen mit einem ganz konkreten Problem, das bereits mehrfach erwähnt wurde und die Bedeutung der Pfadintegrale in der Standard-Quantenmechanik auf Null sinken läßt.

***Argument 1: Der Pfadintegral-Formalismus ist für praktische Anwendungen der nichtrelativistischen Quantenmechanik, insbesondere solche, die auch in der Schule von Interesse sein könnten, ziemlich unbrauchbar.***

<sup>117</sup>Vergleiche Abschnitt 4.5.4.3. Kenner der Materie bemerken natürlich sofort, daß hier die Unterscheidung zwischen reinen und gemischten Zuständen und deren Beschreibung durch statistische Operatoren dahintersteckt.

<sup>118</sup>Die Existenz interferenzfähiger Superpositionen in der Quantenmechanik ist *der* entscheidende Unterschied zur klassischen Physik. Auch dieser Sachverhalt wird in der Schulphysik verbreitet falsch dargestellt, wenn etwa über Interferenz von Wahrscheinlichkeiten [240] oder Interferenz der verschiedenen Möglichkeiten [38], [39] gesprochen wird. Beides ist Unsinn, denn Wahrscheinlichkeiten sind positive reelle Zahlen und können nicht interferieren, und das, was da interferiert, sind keineswegs nur Möglichkeiten, sondern sehr reale Sachverhalte. Richtige Formulierungen dagegen sind *Interferenz von Wahrscheinlichkeitsamplituden* oder besser noch *Interferenz von Zuständen*.

Auf den ersten Blick scheint das ein Aspekt zu sein, der für die Schulphysik von untergeordneter Bedeutung ist, da man selbst die einfachsten Anwendungen der Quantenmechanik, mit Ausnahme höchstens des Potentialtopfs mit unendlich hohen Wänden, im Physikunterricht der Oberstufe gar nicht analytisch behandeln kann. Das sollte uns jedoch nicht von dem bereits in Kapitel 1 bemühten Grundprinzip abbringen, als Basis sämtlicher Begriffsbildungen der Schulphysik ausschließlich etablierte Begriffe der Hochschulphysik zuzulassen und irgendwelche Eigenzüchtungen konsequent zu vermeiden. Im gymnasialen Unterricht verwendete Konzepte müssen stets „aufwärtskompatibel“ sein, das heißt, man sollte sie in der wissenschaftlichen Physik in geeignet frisierter Form wiederfinden können. Nur so kann man einerseits denjenigen Schülerinnen und Schülern, die nach dem Abitur nichts mehr mit Physik zu tun haben, wenigstens einen minimalen, aber dennoch realistischen Eindruck dieses Faches vermitteln und verliert andererseits nicht die Glaubwürdigkeit vor den anderen, denjenigen nämlich, die in irgendeiner Form damit weitermachen. Das sollte stets Ziel eines jeden Physikunterrichts sein.

*Diesem Anspruch genügt der Zeigerformalismus in keiner Weise.* Für sich genommen ist er mehr als nur ohne Bedeutung in der Hochschulphysik; es gibt ihn dort schlicht und einfach überhaupt nicht, was sehr deutlich daran erkennbar ist, daß das Konzept Hochschulphysikerinnen und Hochschulphysikern im allgemeinen völlig unbekannt ist, wenn sie nicht gerade ausgerechnet Spezialisten für Physikdidaktik sind. Das ist auch kein Wunder, denn man kann mit dem Zeigerformalismus für ernsthafte Berechnungen nichts, aber auch gar nichts anfangen. Und auch wenn man ihm das – wie wir noch sehen werden sehr wacklige – Zugeständnis macht, in Gestalt der Feynmanschen Pfadintegrale in der Hochschulphysik ein alltagstaugliches Gegenstück zu besitzen, sieht die Sache nicht wirklich besser aus. Denn die Alltagstauglichkeit dieses Gegenstücks weist die beschriebenen erheblichen Einschränkungen auf; von besonderer Bedeutung ist hier die bereits mehrfach erwähnte Tatsache der Untauglichkeit der Pfadintegrale für die nichtrelativistische Quantenmechanik und damit gerade auch für solche konkrete Anwendungen, die in der Schulphysik von Interesse sind. Die Quantenfeldtheorie und die statistische Mechanik, welche die hauptsächlichen Gebiete sinnvoller Anwendung von Pfadintegralen darstellen, sind sehr weit von der Physik des Gymnasiums entfernt.

### 2.10.1.2 Klassische Näherung und Quantenmechanik

Nach dieser rein fachwissenschaftlichen Angelegenheit kommen wir zu einem Punkt, der die Schulphysik direkt betrifft und nicht nur deren begriffliche Hintergründe.

***Argument 2: Die didaktische Reduktion, welche den Pfadintegral-Formalismus für die Schulphysik überhaupt erst in Frage kommen läßt, führt zu einem chaotischen Durcheinander aus klassischen, halbklassischen und quantenmechanischen Betrachtungen.***

Mit dem zweiten Argument berücksichtigen wir den Schwierigkeitsgrad, den der Pfadintegralformalismus schon in seiner heuristischen Form aufweist. Um Pfadintegrale im Gymnasium überhaupt in irgendeiner Form andeutungsweise verwenden zu können, ist eine sehr weitgehende Reduktion und damit Vereinfachung des Gegenstands erforderlich, wie sie mit dem Zeigerformalismus in der beschriebenen Form versucht wird. Das

beginnt bei der unproblematischen Veranschaulichung komplexer Phasenfaktoren durch rotierende Pfeile, den namensgebenden Zeigern, geht aber erheblich darüber hinaus. Die Vorstellung, entlang jedes stetigen Weges vom Anfangs- zum Endpunkt der zeitlichen Entwicklung des betrachteten Systems einen solchen Zeiger entlangrotieren zu lassen und diese überabzählbar unendlich vielen Zeiger an den Enden der Wege vektoriell aufzuintegrieren, ist einerseits völlig korrekt und liefert eine gewisse anschauliche Vorstellung von Pfadintegralen, andererseits jedoch keinerlei Hilfe für konkrete Berechnungen. Und da die Zielsetzung der Zeiger-Fachdidaktik natürlich darin besteht, die Schülerinnen und Schüler selbst Ergebnisse erzielen zu lassen, ist mit diesem Bild nichts anzufangen. Folglich argumentiert man, daß nur ganz wenige spezielle Wege tatsächlich nennenswerte Beiträge zur Vektorsumme der Pfeile liefern, während alle anderen Wege Pfeile liefern, die in alle möglichen Richtungen weisen und sich somit gegenseitig so gut wie aufheben. Diese wenigen speziellen Wege sind wie beschrieben geradlinige Verbindungslinien, die nur an für das System charakteristischen Hindernissen abgelenkt sind; dafür lassen sich die von ihnen gelieferten Zeigerpositionen tatsächlich ermitteln und vektoriell addieren, wenn auch nur in den allereinfachsten Fällen von Hand und ansonsten meist nur mit Hilfe von Computerprogrammen.

Daß man dabei im wesentlichen wieder bei der klassischen Optik oder dazu analogen Situationen landet, ist natürlich kein Wunder, denn die beschriebene Auswahl der Pfade orientiert sich wie erwähnt an der Methode der stationären Phase. Damit argumentiert man jedoch bei der Erklärung typisch quantenmechanischen Verhaltens mit einem Verfahren, das zur Beschreibung *semiklassischer Systeme* konstruiert ist, also erklären soll, warum in einer quantenmechanischen Welt klassisches Verhalten auftritt. Auch wenn der Zeigerformalismus in seiner gymnasialen Form zu im Rahmen der zugehörigen vereinfachten Betrachtungsweise fachlich korrekten Resultaten führt, geht dabei eine begrifflich saubere Trennung zwischen klassischen, halbklassischen und quantenmechanischen Beschreibungen verloren. Das führt beispielsweise für den Fall der Beugung von materiellen Quantenobjekten am Doppelspalt dazu, den Schülerinnen und Schülern bei der Erklärung der Vorgehensweise zunächst eine klassische Näherung vorzuführen, um danach daraus rein quantenmechanisches Verhalten abzuleiten. Wenn man mit unterschiedlichen Zugängen zur Quantenmechanik bereits vertraut ist, kommt man damit klar; für diejenigen, die in einem Physikkurs der Oberstufe des Gymnasiums davon hören, ist es jedoch der Erstkontakt, mit der Konsequenz, daß das Bild der Teilchenbahnen wörtlich genommen und anschaulich interpretiert wird. Damit geht gerade der erste Eindruck einer unvermeidlichen Unanschaulichkeit der Quantenmechanik verloren, der für einen korrekten Umgang mit derselben und ganz besonders für das Vermeiden typischer Fehlinterpretationen von fundamentaler Bedeutung ist.

### 2.10.1.3 Anschaulichkeit und Unanschaulichkeit

Wir bleiben gleich bei diesem Aspekt und kümmern uns damit um einen angeblich besonders gewichtigen Vorteil, der den rotierenden Zeigern von seinen Anhängern nachgesagt wird und in Wirklichkeit ein Nachteil ist.

***Argument 3: Der Zeigerformalismus gaukelt Anschaulichkeit vor, wo es keine gibt.***

Hierbei handelt es sich inhaltlich um ein didaktisches Argument, die weitere Begründung ist jedoch rein physikalischer Natur. Die Attraktivität, welche dem Zeigerformalismus anzuhafte scheint, dürfte wohl zu erheblichen Teilen auf die vermeintliche Einfachheit zurückzuführen sein, die in einschlägigen Darstellungen des Gegenstands vermittelt werden sollen. Von einer systematisch unanschaulichen, sehr formalen und mathematisch aufwendigen Theorie bleibt dabei gerade etwas Geometrie in der Ebene übrig, nämlich das Verschieben, Drehen und Addieren zweidimensionaler Vektoren, und auch das nicht einmal rechnerisch, sondern nur zeichnerisch. Das wirkt alles wunderbar leicht zugänglich und anschaulich vorstellbar, womit das Publikum jedoch konsequent über die unausweichliche Unanschaulichkeit der Quantenmechanik hinweggetäuscht wird. Das geschieht natürlich nicht absichtlich, sondern als Folge einer unbrauchbaren didaktischen Taktik, denn die Ursache des falschen Eindrucks ist der auch andernorts<sup>119</sup> anzutreffende, aber stets zum Scheitern verurteilte Versuch, *klassische Physik in die Quantenmechanik herüberzuretten*<sup>120</sup>.

Oberflächliche Lektüre der grundlegenden Arbeiten von Wentzel, Dirac und Feynman, welche zum Pfadintegralformalismus in seiner heuristischen Form geführt haben<sup>121</sup>, verleitet womöglich zur Auffassung, dieser stelle einen klassischen Zugang zu zentralen Resultaten der Quantenmechanik dar. Das ist selbstverständlich völlig abwegig; richtig ist gerade einmal, daß zunächst Dirac und dann Feynman versuchten, die Bedeutung der *klassischen Wirkungsfunktion* in der Quantenmechanik auszuloten. Man kann es noch etwas drastischer formulieren: Feynmans Ansatz war es, die Quantenmechanik weitestgehend *mit klassischen Begriffen* zu formulieren. Das heißt jedoch *nicht*, sie „weitgehend klassisch“ zu formulieren, wie immer wieder behauptet wird. Im Gegenteil, die nichtklassischen Aspekte sind in vollem Umfang vorhanden, sie sind aber *sehr gut getarnt*. Hierin liegt insbesondere für Ungeübte die begriffliche Gefahr des Ansatzes. Er stellt vermeintlich eine verlockende Möglichkeit dar, doch wieder mit klassischer Physik auszukommen, ein Eindruck, der jedoch auf das äußerste täuscht. Denn was dabei herauskam, war schließlich ein Formalismus, der bei genauerer Betrachtung noch viel bizarrer, seltsamer und damit deutlich weiter von der klassischen Physik entfernt erscheint als die Standardversionen der Quantenmechanik.

Aus der Sicht der angewandten Physik sind Pfadintegrale insbesondere in der relativistischen Quantenfeldtheorie gelegentlich in der Lage, ansonsten seitenlange Rechnungen auf das Format einer kleineren Wandtafel zu reduzieren. In der nichtrelativistischen Quantenmechanik dagegen ist die Sache sehr genau besehen ein zwar formal genialer, praktisch jedoch unbrauchbarer Rechenrick. Beides zusammen sollte eigentlich völlig ausreichen, Pfadintegrale weder mit anschaulichen Vorstellungen in Verbindung zu bringen, noch sie für ein einfaches Hilfsmittel bei der Betrachtung elementarer quantenmechanischer Fragestellungen zu halten – wenn man mit der Angelegenheit bereits vertraut ist. Für Schülerinnen und Schüler, die ja in ihrer Mehrzahl keine weiterführenden Themen mehr kennenlernen, bleibt dagegen der Eindruck, daß Quantenmechanik nichts

<sup>119</sup>Hier wären beispielsweise die vielfältigen, aber allesamt untauglichen Ansätze Bohmscher Quantenmechanik zu nennen. Sehr fundierte Kritik dazu findet man bei [114], [148], [188], [189], [210], [390], [391] und [495].

<sup>120</sup>Man erinnere sich an die zu Beginn des Abschnitts 2.9 erwähnte Selbstbeschreibung des Zeigerformalismus, ein Zugang zur Quantenmechanik zu sein, der „so klassisch wie möglich ist, ohne dabei physikalisch falsch zu sein“, welche ihre eigene Unmöglichkeit besser beschreibt als viele Worte.

<sup>121</sup>Siehe dazu die Zitate zu Beginn von Abschnitt 2.2.



weiter als ein schön anschauliches Hantieren mit einfachen mathematischen Objekten ist<sup>122</sup>. Erfahrungsgemäß hilft es hier wenig, zu betonen, daß es sich beim Zeigerformalismus wie bei jeder physikalischen Beschreibung der Natur um ein Modell und nicht ein direktes Abbild der Wirklichkeit handle, selbst wenn das noch so oft wiederholt wird. Die durch die Verwendung des Zeigerformalismus im Physikunterricht sich unmittelbar aufdrängende Vorstellung, daß ein Großteil der Abiturientinnen und Abiturienten das Gymnasium mit der Überzeugung verläßt, Quantenmechanik sei im wesentlichen die Betrachtung rotierender Zeiger und weiter nichts, ist wohl gleichermaßen realistisch wie deprimierend.

#### 2.10.1.4 Quantenmechanik und Quantenelektrodynamik

Zum Abschluß dieser ersten Kategorie kümmern wir uns wieder einmal um einen Fall mangelnder Präzision bei der Verwendung physikalischer Fachbegriffe.

***Argument 4: Mit dem Zeigerformalismus werden elementare Quantenmechanik und Quantenelektrodynamik durcheinandergebracht.***

Dieses Argument sollte eigentlich überflüssig sein, denn man traut sich kaum zu erwähnen, daß die Quantenelektrodynamik *kein Teilgebiet* der Standard-Quantenmechanik ist, so selbstverständlich ist diese Aussage. Und dennoch, während Feynman in seiner populärwissenschaftlichen Darstellung des Zeigerformalismus [167] nicht nur durch den Buchtitel, sondern auch inhaltlich keinen Zweifel daran läßt, worum es geht, scheint das in der gymnasialen Physikdidaktik längst nicht so klar zu sein. Interessanterweise wird im Physik-Bildungsplan von 2001 ausdrücklich auf die „didaktische Reduktion der Quantenelektrodynamik von Richard Feynman“ verwiesen [2]. Abgesehen von der Frage, was ein Begriff wie *Quantenelektrodynamik* überhaupt in einem Lehrplan für Gymnasien zu suchen hat, fällt dabei die fast kommentarlose Empfehlung eines Rezepts auf, das offensichtlich einer ganz anderen Sparte der Physik als diejenige, um die es eigentlich geht, zugeordnet wird. Dadurch wird der Eindruck erweckt, die Quantenelektrodynamik sei gewissermaßen eine etwas vornehmere oder leicht verbesserte Version der Standard-Quantenmechanik; überflüssig zu betonen, daß das kompletter Unsinn ist.

Im Gegensatz zur Standard-Quantenmechanik, die insofern hybrid ist, als sie Teilchen quantenmechanisch, Felder jedoch klassisch beschreibt, ist die Quantenelektrodynamik die relativistische Quantentheorie des elektromagnetischen Feldes, das heißt, sie hat Wechselwirkungen des quantisierten elektromagnetischen Feldes mit geladenen Quantenobjekten zum Gegenstand. Dabei stehen Streuprozesse im Mittelpunkt der Betrachtung, deren Streuquerschnitte störungstheoretisch, aber zum Teil mit sagenhafter Genauigkeit berechnet werden können. Dabei ist es überflüssig zu betonen, daß so etwas nichts mit Schulphysik zu tun hat. Kommen dabei Pfadintegrale zum Einsatz, so sind es solche aus der relativistischen Quantenfeldtheorie<sup>123</sup>. Damit ist eine Berufung auf die Quantenek-

<sup>122</sup>Wobei noch gar nicht berücksichtigt ist, was bereits angedeutet wurde, nämlich daß sich dieses auf den ersten Blick schön einfache Hantieren schnell als äußerst aufwendig und mühsam erweist. Darauf kommen wir in Abschnitt 2.10.3.3 zurück.

<sup>123</sup>Feynman berücksichtigt das in seiner schon mehrfach zitierten populären Darstellung [167], indem er genuin quantenelektrodynamische Begriffe nur im Zusammenhang mit Prozessen verwendet, die auch dorthin gehören, also beispielsweise Streuprozesse von Elektronen und Photonen.

trodynamik bei der Rechtfertigung des Zeigerformalismus in der Schule nicht gerechtfertigt. Das führt höchstens zu einer Begriffsverwirrung zwischen elementarer nichtrelativistischer und relativistischer Quantenmechanik.

## 2.10.2 Erkenntnistheoretische Aspekte

Wir kommen nun zu Argumenten, die mit philosophischen Grundlagen der Quantenmechanik und damit mit deren Interpretation zu tun haben. Diese können zwar nicht direkt im Unterricht thematisiert werden, letzterer sollte jedoch stets auf der Grundlage einer soliden wissenschaftlichen und damit auch philosophischen Basis erfolgen<sup>124</sup>.

### 2.10.2.1 Der Status der Quantenmechanik als grundlegende Naturbeschreibung

Die Quantenmechanik erhebt für sich den Anspruch, eine universelle und damit die Natur auf grundlegende Weise beschreibende Theorie zu sein. Das führt unmittelbar auf den folgenden Aspekt.

*Argument 5: Der Pfadintegral-Zugang zur Quantenmechanik darf nicht als metaphysisch-grundlegendes Naturprinzip aufgefaßt werden, wie es bei der kanonischen Quantisierung der Fall ist.*

Dieser Punkt wurde in Abschnitt 2.8 bereits aus allgemeiner Sicht diskutiert, ohne daß dabei auf schulische Belange Bezug genommen wurde. Sieht es auf den ersten Blick so als, als seien solche Dinge weit oberhalb jedes vernünftigen Niveaus der Physik am Gymnasium, so hat man auf den zweiten Blick gewichtige Gründe, genau darüber nachzudenken. Gerade weil man sich beim Unterrichten der Quantenmechanik in der Schule auf die allerelementarsten Aspekte beschränken muß, ist es unbedingt erforderlich, hier auf grundlegende Prinzipien der betrachteten Theorie aufzubauen; nur dann sind die dabei unvermeidlichen Vereinfachungen wissenschaftlich unbedenklich.

*Der Pfadintegral-Zugang zur elementaren Quantenmechanik ist kein solches grundlegendes Naturprinzip*, ganz im Gegensatz zum Standard-Zugang über Zustandsvektoren und selbstadjungierte Operatoren in Hilberträumen sowie die kanonische Quantisierung, der als fundamentale Grundlage der Quantenmechanik vom Rang eines axiomatischen Systems aufzufassen ist<sup>125</sup>. Natürlich ist auch das nicht im entferntesten in der Schule zu stemmen, nicht einmal in der formal-robust-unsauberen Version der theoretischen und schon gar nicht in der subtil ausgefeilten, formal strengen Form der mathematischen Physik. Hier ist jedoch eine Elementarisierung nicht nur problemlos möglich, sie steht vor allem dank ihrer Herkunft auf einer in jeder Hinsicht soliden Basis.

Beim Pfadintegral-Zugang sieht das völlig anders aus. Selbst wenn man von allen anderen hier beschriebenen Nachteilen seiner Simplifizierung in Form des Zeigerformalismus absieht, bleibt das schwerwiegende Problem, daß hier ein wenn auch geniales, so doch rein heuristisches Verfahren vom Rang eines Rechentricks zum Ausgangspunkt der ersten Einführung des Themas für Lernanfänger gemacht wird. Wir haben damit eine zu

<sup>124</sup>Wir kommen in Abschnitt 4.2 darauf zurück.

<sup>125</sup>Ausführlicher dazu in Abschnitt 4.2.2.



derjenigen beim Karlsruher Physikkurs völlig analoge Situation; auch dort wird so getan, als sei ein Verfahren, das in Wahrheit nur ein rechentechnischer Behelf ist, nämlich die phänomenologische Gleichgewichtsthermodynamik, das fundamentale Grundprinzip der Natur. In beiden Fällen kommen Modelle zum Einsatz, die dem Anspruch der Allgemeingültigkeit, die ihre Anhänger für sie erheben, nicht gerecht werden können.

Möchte man den Verfechtern dieser Konzepte nicht Unkenntnis vorwerfen, so bleibt nur der Schluß, daß sie einer fragwürdigen erkenntnistheoretischen Haltung anhängen, nämlich derjenigen, daß alle Modelle gleich gut sind, solange sie nur korrekte Vorhersagen für Meßergebnisse zu machen imstande sind. Inwieweit das überhaupt tatsächlich der Fall ist, lassen wir hier einmal dahingestellt und betonen stattdessen, daß wir eine solche instrumentalistische oder, um es etwas pointierter zu formulieren, positivistische Haltung nicht teilen. Das käme der Aussage gleich, daß Metaphysik an sich sinnlos ist. Zwar ist auch das eine Art von Metaphysik, aber die primitivste denkbare überhaupt. Keine philosophische Theorie kommt ohne einen philosophischen Unterbau aus, und der unmittelbar physikalisch relevante Teil desselben wird stets durch ein System von Axiomen repräsentiert, aus denen die Theorie ableitbar ist. Auch wenn das im Detail selbstverständlich in der Schule nicht im entferntesten realisierbar ist, sollte man dennoch versuchen, wenigstens eine elementare vereinfachte Version davon zu entwickeln. Das kann jedoch im Fall der Quantenmechanik nur auf Basis des Standard-Zugangs und keineswegs mit dem Pfadintegral-Zugang geschehen. Wie so etwas aussieht, ist in Abschnitt 4.2 beschrieben.

### 2.10.2.2 Die mathematischen Defizite der Feynmanschen Pfadintegrale

Das soeben aufgeführte Manko des Zeigerformalismus ist nicht das einzige mit fundamental-philosophischem Hintergrund. Wir kommen nun zu einem weiteren mit noch tieferschürfender Bedeutung, was schon daran zu erkennen ist, daß man die Abschnitte 2.3 bis 2.7 als Vorarbeit dafür auffassen kann.

***Argument 6: Der Zeigerformalismus basiert mit den Feynmanschen Pfadintegralen auf einem Zugang, der mathematisch nicht definierte Objekte verwendet.***

Auch hier würde man möglicherweise zunächst jeden Bezug zum Physikunterricht des Gymnasiums in Frage stellen. Das ist eine Fehleinschätzung, zu deren Begründung wir etwas ausholen. Es wurde oben ausführlich dargelegt, daß das Konzept der heuristischen Pfadintegrale mathematisch nicht durchführbar ist und man beim Versuch, einen wohldefinierten Ersatz zu finden, tief in die Trickkiste der Mathematik greifen mußte und außerdem nur partiell erfolgreich war; entweder man verläßt das physikalische Szenario, von dem man gestartet ist (Wiener-Integrale), oder man erhält Konstruktionen, die nur bei speziellen Potentialen, nicht aber im Allgemeinen funktionieren (alle anderen bis jetzt bekannten Zugänge). So oder so führt die rigorose Betrachtung des Gegenstands zur Erkenntnis, daß Pfadintegrale, möchte man sie nicht als nur heuristisch, keineswegs jedoch in Strenge definierbare Objekte stehenlassen, generell ein höchst anspruchsvolles Spezialgebiet der mathematischen Physik darstellen.

Das hat unmittelbar die Folge, daß der Zeigerformalismus, sobald man ihm mehr zumuten will als eine Veranschaulichung komplexer Zahlen oder ein nicht sauber begründbares Behelfsmodell der Schulphysik zu sein, mit einer mathematisch nicht definierten oder besser gesagt gar nicht definierbaren Rechenmethode gerechtfertigt werden

muß. Natürlich merken das die Schülerinnen und Schüler nicht (und die Lehrerinnen und Lehrer im allgemeinen wohl auch nicht), aber unabhängig davon haben wir es hier mit derselben Schwierigkeit wie beim gerade eben diskutierten Argument zu tun. Ganz genau wie der schulische Physikunterricht stets auf einer hieb- und stichfesten naturphilosophischen und erkenntnistheoretischen Basis stehen sollte, muß seine Basis auch aus mathematischer Sicht immer unerschütterlich sein. Nur so kann dem Charakter der Physik als einer mathematischen Naturwissenschaft Rechnung getragen werden, insbesondere auch, weil die formale Rigorosität dieses Fachs natürlich in seinen didaktisch bearbeiteten Formen in unterschiedlichem Maß nicht einzuhalten ist. Umso wichtiger ist es dafür, bei dieser Bearbeitung auf formal unangreifbarer Grundlage zu stehen. *Das ist beim Zeigerformalismus definitiv nicht der Fall*, wie wir ausführlich gezeigt haben. Die Situation erinnert an den Versuch, einem Kunden die Bedienung einer komplizierten modernen Waschmaschine anhand der Bedienungsanleitung eines viel einfacheren Vorgängermodells zu erklären, das zudem nicht richtig funktioniert. Selbst wenn der Kunde beim einfachen alltäglichen Gebrauch der neuen Waschmaschine wohl meist einigermaßen klarkommt, ist so ein Vorgehen nicht akzeptabel. Wie fast jeder Vergleich hinkt allerdings auch dieser, und zwar insofern, als man natürlich bei der Waschmaschine ohne weiteres die richtige Gebrauchsanleitung verwenden könnte, bei den Pfadintegralen dieser Weg jedoch für die Schule nicht in Frage kommt. Aus den Pfadintegralen der mathematischen Physik läßt sich keine vernünftige Didaktik für die Schulphysik basteln, nicht nur, weil sie viel zu kompliziert sind – man überfliege erneut die Abschnitte 2.3 bis 2.7 – sondern vor allen Dingen auch, weil sie wie beschrieben nur für die Quantenfeldtheorie ein wirklich fundamentales Werkzeug der Physik darstellen.

### 2.10.3 Didaktische Aspekte

Mit der dritten Kategorie kommen wir nun zu einigen Mängeln, die wir direkt mit Blick auf den Physikunterricht am Zeigerformalismus auszusetzen haben. Seine Anhänger beanspruchen, damit ein umfassendes Konzept für die Quantenmechanik in der Oberstufe gefunden zu haben, was ganz besonders aus didaktischer Sicht zu dessen kritischer Hinterfragung verleiten sollte.

#### 2.10.3.1 Pfadintegrale für die Schule

Wir beginnen mit einem Problem, das Kennern der fachwissenschaftlichen Hintergründe wohl zuallererst auffallen dürfte.

***Argument 7: Ausgerechnet der schwierigste und problematischste aller Zugänge zur Quantenmechanik soll als Grundlage eines didaktischen Konzepts für die Schulphysik herhalten.***

Der hier beschriebene Sachverhalt mutet für sich genommen schon äußerst merkwürdig an; da das jedoch so ohne weiteres nicht als sachliche Diskussionsgrundlage akzeptiert werden dürfte, sind weitere Kommentare angebracht. Wir blenden dazu für den Moment einmal alle bisher aufgeführten Argumente aus und konzentrieren uns auf den Gegenstand, der die Motivation zum Abfassen des vorliegenden Kapitels bildet, das heißt das Unterrichten der Quantenmechanik in der gymnasialen Oberstufe. Man sieht sich dabei

mit der Aufgabe konfrontiert, ein sehr schwieriges und zudem äußerst unanschauliches Theoriengebäude auf ein für die Schule erträgliches Maß zu reduzieren. Als besonders schwierig erweist sich das unter anderem, weil einerseits angesichts der Zielgruppe extrem elementarisiertes Vorgehen erforderlich ist, andererseits die tiefgreifende Bedeutung und insbesondere die notwendige Unanschaulichkeit des Gegenstands auch im Gymnasium deutlich hervortreten sollte.

Nun findet man bei unvoreingenommener Betrachtungsweise zwei grundlegend unterschiedliche formale Zugänge zur Quantenmechanik in ihrer modernen Form, einen über Operatoralgebren und einen über Funktionalintegration<sup>126</sup>. Der erste findet in stark vereinfachter Form weitverbreitete Anwendungen auf die unterschiedlichsten konkreten physikalischen Probleme<sup>127</sup> und läßt sich direkt weiter elementarisieren, und zwar in einer Weise, daß der Formalismus *sowohl* einfach genug für die Oberstufe des Gymnasiums ist, *als auch* den prinzipiell unanschaulichen Charakter behält, welcher der Quantenmechanik unvermeidlicherweise anhaften muß. Der andere basiert auf einem raffinierten, aus der unendlichdimensionalen Analysis stammenden Trick, der jedoch technisch außerordentlich schwierig und in noch viel stärkerem Maße mathematisch windig ist, außerdem eine Anschaulichkeit simuliert, die in umgekehrt proportionalem Verhältnis zur tatsächlichen Anschaulichkeit sowohl seiner selbst als auch der Quantenmechanik insgesamt steht, und zudem auch noch für praktische Belange, soweit sie für die Schule auch nur ein wenig interessant sind, völlig ohne Bedeutung ist. Wenn man sich zum Ziel setzt, im Gymnasium das Fach Physik in einer Weise zu unterrichten, die zumindest ein wenig als vereinfachtes Modell der tatsächlichen wissenschaftlichen Beschäftigung mit diesem Fach jenseits der Schule durchgehen kann, sollte die Auswahl zwischen diesen beiden Zugängen eigentlich nicht schwer fallen.

### 2.10.3.2 Klassische Physik durch die Hintertür

Bisher haben wir uns – auch beim vorigen Argument, obwohl es zur dritten Kategorie zu rechnen ist – im wesentlichen auf einer didaktisch-transzendentalen Ebene bewegt, das heißt, wir haben hauptsächlich über Voraussetzungen zum Betreiben von Fachdidaktik und weniger über Fachdidaktik selbst geredet. Das ändern wir jetzt, indem wir einige Aspekte des Zeigerformalismus anführen, die sich als unmittelbar nachteilig für den schulischen Physikunterricht und seine Zielsetzungen erweisen werden. Der erste davon betrifft die eigenen Vorstellungen, die Schülerinnen und Schüler von der Quantenmechanik bilden.

***Argument 8: Der Zeigerformalismus provoziert Fehlvorstellungen und Fehlinterpretationen, auch solche, die meist zuvor mühsam ausgeräumt worden sind.***

Eines der wesentlichen Ziele eines jeden Quantenmechanik-Unterrichts sollte es sein, die fundamentalen Unterschiede zur klassischen Physik zumindest ansatzweise herauszuarbeiten. Dazu gehört beispielsweise die Aufgabe des Bahnbegriffs als universelle, stets

<sup>126</sup>Ist man doch etwas voreingenommen oder hat man das vorliegende Buch bis hierher gelesen, weiß man, daß man nur in der Quantenfeldtheorie, nicht aber in der nichtrelativistischen Standard-Quantenmechanik tatsächlich zwei ernstzunehmende Möglichkeiten hat, aber wir wollten ja vorübergehend alle anderen Argumente ausblenden.

<sup>127</sup>In dieser Form wird er üblicherweise auch in den Kursvorlesungen der Universitäten unterrichtet.

mögliche Beschreibung des zeitlichen Verhaltens physikalischer Systeme, oder auch die Erkenntnis, daß Quantenobjekte häufig nicht als räumlich lokalisierte Gebilde aufgefaßt werden dürfen. Der Weg dorthin ist anstrengend, und das gewünschte Ziel dürfte wohl selten wirklich flächendeckend erreicht werden, übrigens an den Gymnasien so wenig wie an den Universitäten. Dennoch sollte man es sich stets setzen, und selbst wenn man nur etwas Zweifel an den gefestigten klassischen Vorstellungen der Schülerinnen und Schüler sät und diese wenigstens realisieren, daß physikalische Erkenntnisse ihrem „gesunden Menschenverstand“ häufig widersprechen, hat man schon einiges erreicht.

Diesem Ziel läuft die Verwendung des Zeigerformalismus diametral zuwider. Für einen Anfänger sieht es so aus, als ob man es hier mit klassischen, im Idealfall punktförmigen Partikeln und wohldefinierten Bahnen zu tun hat, wenn auch mehr davon auftauchen als bisher gewohnt. In der Tat ist ständig von möglichen Bahnen die Rede, die auch noch ausgiebig in Skizzen von Versuchsanordnungen eingezeichnet werden<sup>128</sup>. Protagonisten desselben werden nicht müde zu betonen, daß es sich dabei um ein Modell und nicht um ein Abbild der Wirklichkeit handelt, womit Kenner der Materie natürlich auch keinerlei Probleme haben. Schülerinnen und Schüler werden jedoch das vermeintliche Angebot, zu vertrauten Begriffen der klassischen Physik zurückzukehren, häufig bereitwillig bis dankbar annehmen, insbesondere, wenn man sie zuvor ausgiebig mit viel (berechtigtem) Gerede von deren Verbot strapaziert hat. Damit gehen mühsam aufgebaute Zweifel an der klassischen Physik und in günstigen Fällen erreichte Überzeugungen, daß dieselbe zumindest im mikroskopischen Bereich nicht gültig ist, wieder verloren. Insbesondere konterkariert das die Aussage, daß Quantenobjekte normalerweise die Eigenschaft, einen Ort und einen Impuls aufzuweisen und damit eine Bahn zu beschreiben, nicht besitzen, und läuft auch der Erkenntnis zuwider, daß man sich Quantenobjekte normalerweise nicht als räumlich lokalisierte Materiepakete vorstellen darf. Man ist begrifflich unversehens wieder in bedenklicher Nähe des Bohrschen Atommodells gelandet, auch wenn man eigentlich etwas ganz anderes erzählt zu haben glaubt.

Noch dramatischer wird die Sache, wenn man der inzwischen ebenfalls verbreiteten Gepflogenheit folgt, im Schulunterricht *Feynman-Diagramme* zu verwenden. Das hat nicht unmittelbar etwas mit dem Zeigerformalismus zu tun, tritt aber häufig Hand in Hand mit diesem auf und ist ein weiterer untauglicher Versuch, Methoden aus der Quantenelektrodynamik für die Schulphysik verwendbar zu machen. Feynman-Diagramme sind graphische Darstellungen der einzelnen Ordnungen der störungstheoretischen Entwicklungen von Übergangsamplituden bei Streuprozessen; sie stehen damit jeweils für komplizierte Integrale im Minkowski-Raum. Die Übersetzungsregeln von Diagrammen zu Integralen sind unter der Bezeichnung *Feynman-Regeln* bekanntgeworden<sup>129</sup>. Feynman-Diagramme haben nichts, aber auch gar nichts im Physikunterricht des Gymnasiums verloren, und das umso mehr, als sie dort so gut wie immer in unzulässiger Weise anschaulich als räumliche Darstellung der Wege der gestreuten Teilchen und dabei auftretenden Austauscheteilchen gedeutet werden, wie man sie von Abbildungen in populärwissenschaftlichen Berichten über Teilchenbeschleuniger kennt. Sie sind jedoch nichts so wenig wie das.

<sup>128</sup>Im Bewußtsein dieser Problematik wird inzwischen teilweise zu Kunstbegriffen wie „Rechenpfade“ und dergleichen gegriffen, ein weiteres absurdes Beispiel einer Begriffsneuschöpfung durch die Schulphysik.

<sup>129</sup>Feynman beschreibt solche Diagramme auf populärwissenschaftliche Weise in [167]. Eine besonders gut lesbare Einführung in dieses Gebiet auf akademischem Niveau ist [452].

Wenn man schon irgendwelche Streuprozesse mit einlaufenden und auslaufenden Teilchen diskutieren möchte, kann man diese zwar fraglos graphisch veranschaulichen, man darf dann jedoch solche Darstellungen keineswegs als Feynman-Diagramme bezeichnen. Umgekehrt sollte man, bevor man sich mit denselben beschäftigt, zumindest eigenhändig einige Streuquerschnitte berechnet haben.

### 2.10.3.3 Aufwand und Nutzen des Zeigerformalismus

Als nächstes kommen wir zu einem ganz praktischen Kritikpunkt, der direkt mit Unterrichtsplanungen und erst in zweiter Linie mit inhaltlichen Dingen zu tun hat.

*Argument 9: Mit dem Zeigerformalismus wird ein weiterer aufwendiger Formalismus in die Schulphysik eingeführt, der keine neuen Erkenntnisse liefert, in den meisten Fällen ohne Computereinsatz nicht handhabbar und für weiterführende Einsätze völlig untauglich ist.*

Die Oberstufe des Gymnasiums ist nicht nur in Physik, sondern auch in Mathematik und Chemie vollgestopft mit für schulische Verhältnisse teilweise sehr abstrakten formalen Systemen, die aus wissenschaftlicher Sicht zwar überaus elementar sind, für Schülerinnen und Schüler jedoch schwierige Herausforderungen darstellen. Entsprechend sollte man bei jeder Einführung irgendwelcher neuer abstrakter Sachverhalte gründlich nachdenken, ob das wirklich erforderlich und die Anstrengungen wert ist. Das trifft in besonderem Maße auf den Zeigerformalismus zu. Mit ihm gelangt ein Verfahren in den Physikunterricht, das von Null an komplett neu entwickelt werden muß und schon dazu etliche Unterrichtsstunden und viel Mühe erfordert, andererseits aber nicht gerade den Anspruch erheben darf, ein grundlegendes Naturprinzip zu sein. Ein solcher Aufwand läßt sich nur rechtfertigen, wenn es sich anderweitig irgendwie lohnt. Davon kann jedoch keine Rede sein.

Zunächst liefert der Zeigerformalismus keinerlei Ergebnisse, die sich auf der Ebene der Schulphysik mit ausreichender Genauigkeit nicht auch anderweitig erzielen ließen. Beugung am Doppelspalt beispielsweise stellt das vielleicht verbreitetste Anwendungsfeld des Zeigerformalismus dar; das Interferenzbild läßt sich aber genauso gut und völlig ausreichend in der herkömmlichen Weise über das Huygenssche Prinzip und die Betrachtung des Gangunterschieds in Abhängigkeit vom Ablenkungswinkel  $\alpha$  hinter dem Doppelspalt beschreiben. Beim Gitter, wo man mit dem Zeigerformalismus schon ziemliche Kopfstände machen muß, funktioniert es ganz genauso. In beiden Fällen führt das auf die wohlbekannt Formel

$$\sin \alpha = \frac{n\lambda}{g}, \quad n \in \mathbb{N},$$

für die Ablenkungswinkel  $\alpha$ , unter denen beim beobachteten Interferenzmuster die Intensitätsmaxima  $n$ -ter Ordnung erscheinen;  $\lambda$  ist dabei die Wellenlänge der verwendeten Photonen beziehungsweise de Broglie-Wellenlänge der verwendeten materiellen Teilchen und  $g$  der Spaltabstand beziehungsweise die Gitterkonstante. Der Vorteil des Zeigerformalismus, bei Drei-, Vier- oder Mehrfachspalt-Interferenz einfacher zum richtigen Interferenzmuster mit Haupt- und Nebenmaxima zu führen, ist keiner, da solche Details für die Quantenmechanik in der Oberstufe völlig irrelevant sind.



Die Beschreibung von Interferenzmustern bei Beugung am Gitter ist im übrigen nicht das einzige Beispiel, bei dem der Zeigerformalismus ungleich aufwendiger zu handhaben ist als das herkömmliche Verfahren. In der Tat läßt sich das Verschieben von rotierenden Zeigern und anschließende vektorielle Addieren und Quadrieren nur in den allereinfachsten Fällen direkt von Hand durchführen; im allgemeinen ist man auf die Verwendung von Computerprogrammen angewiesen, was den unterrichtlichen Aufwand weiter erhöht und die Durchsichtigkeit des Verfahrens entsprechend einschränkt.

Umgekehrt gibt es viele elementare quantenmechanische Situationen, bei denen der Zeigerformalismus nur eingeschränkt oder gar nicht funktioniert. Die wichtigsten Beispiele hierfür sind stationäre Zustände. Auf diese ist der Zeigerformalismus *überhaupt nicht anwendbar*, sofern man ihn wie üblich als stark elementarisierten Pfadintegralformalismus betrachtet, denn Pfadintegrale stellen Integraldarstellungen des quantenmechanischen Zeitentwicklungsoperators dar; sie dienen zur Berechnung von Übergangsamplituden, und diese verschwinden bei stationären Zuständen stets, solange man nicht irgendwelche zeitabhängigen Störungen mitberücksichtigt<sup>130</sup>. Wenn hier Zeiger zum Einsatz kommen, sind sie folglich *nur noch* Hilfsmittel zur Vermeidung komplexer Zahlen, wo diese notwendig sind und haben damit nichts mehr mit dem Zeigerformalismus im eigentlichen Sinn zu tun<sup>131</sup>.

Man sieht das sehr deutlich am ersten (und wohl häufig auch einzigen) konkreten quantenmechanischen System, das traditionellerweise in der Schule betrachtet wird, ein Teilchen im Potentialtopf mit unendlich hohen Wänden. Um die eindimensionale Variante für den Zeigerformalismus zugänglich zu machen, wird nun wie folgt argumentiert: Als „mögliche Pfade“ des Teilchens identifiziert man den „Weg“ einer nach rechts und einer nach links laufenden ebenen Welle  $e^{i\omega t}$  zwischen den Wänden. Da man diese beiden Wellen nun zu einer stehenden Welle im Potentialtopf überlagern möchte, aber keine komplexen Zahlen zur Verfügung hat, veranschaulicht man die komplexen Phasen der Wellen wiederum durch rotierende Zeiger entlang der beiden Pfade und addiert die Zeiger der beiden Wellen jeweils an jeder Stelle des Potentialtopfs<sup>132</sup>. Die resultierenden Zeiger liefern natürlich gerade die gewünschte stehende Welle. Es werden insbesondere keine resultierenden Zeiger an den Endpunkten irgendwelcher Pfade, entlang derer sie rotieren, im Sinne aufgesammelter Phasen aufaddiert. Daß das Verfahren folglich nicht das Geringste mit dem Zeigerformalismus im Sinne einer didaktischen Reduktion der Pfadintegrale zu tun hat, bräuchte eigentlich gar nicht erwähnt zu werden. Es handelt sich hierbei ganz einfach um die Verwendung komplexer Zahlen in der Schule, ohne daß man das zugibt. Stationäre Zustände im dreidimensionalen, und dazu gehören mit dem dreidimensionalen unendlich tiefen Potentialtopf und dem Wasserstoffatom Systeme, die im Schulunterricht gemeinhin ebenfalls angesprochen werden, lassen sich mit rotierenden Zeigern überhaupt nicht beschreiben, nicht einmal, wenn man sie nur zur Umgehung komplexer Zahlen einsetzt.

<sup>130</sup>Dieser Sachverhalt ist schuld daran, daß das übliche quantenmechanische Atommodell die Stabilität der Atome genau genommen gar nicht erklärt, sondern vielmehr ein Zuviel an Stabilität liefert. Erst quantenelektrodynamische Betrachtungen und damit die Einbeziehung zeitabhängiger Prozesse liefert eine adäquate Beschreibung der Physik der Atomhülle.

<sup>131</sup>Einschlägige Darstellungen der vermeintlichen Beschreibung stationärer Zustände mit dem Zeigerformalismus wecken gewisse Zweifel, ob die Autoren über diesen Sachverhalt informiert sind.

<sup>132</sup>Auch hier drängt sich Computereinsatz auf, möchte man die Klasse nicht ausgiebig mit dem Zeichnen und zeichnerischen Addieren von Vektorpfeilchen beschäftigen.

Die hier beschriebene Herleitung ist im übrigen nur eine heuristische Plausibilisierung der resultierenden stehenden Welle und alles andere als ein mathematisch oder physikalisch zwingendes Verfahren; außerdem ist es bei ihrer alleinigen Verwendung völlig unklar, um was für Wellen es sich dabei überhaupt handelt. Man muß folglich den Begriff der quantenmechanischen Wellenfunktion zusätzlich thematisieren. Damit wird ein weiteres schwerwiegendes Manko des Zeigerformalismus deutlich: *Er stellt keinerlei Zugang zur Schrödingergleichung als fundamentale Bestimmungsgleichung für Wellenfunktionen dar*<sup>133</sup>. Man kann zwar im Rahmen des Pfadintegral-Formalismus Wellenfunktionen eines quantenmechanischen Systems definieren und zeigen, daß diese Lösungen der zugehörigen Schrödingergleichung sind<sup>134</sup>, das übersteigt die Möglichkeiten der Schulphysik und insbesondere des Zeigerformalismus jedoch bei weitem. Ähnliches gilt für die Heisenbergsche Unschärferelation, der zwar mit Pfadintegralen und entsprechendem Aufwand, nicht aber mit rotierenden Zeigern beizukommen ist. Wir müssen daher zum Schluß kommen, daß mit dem Zeigerformalismus wesentliche elementare Aspekte der Quantenmechanik, insbesondere auch solche mit Relevanz in der Schulphysik, nicht zugänglich sind.

#### 2.10.3.4 Blick über den Horizont

Nachdem wir nun gezeigt haben, daß die Einführung des Zeigerformalismus in der Schule abgesehen von vielen anderen Nachteilen einen Aufwand darstellt, der sich in keiner Weise lohnt, beschäftigen wir uns abschließend mit den Perspektiven, die er im Hinblick auf die Physik als Fachwissenschaft vermittelt.

***Argument 10: Der Zeigerformalismus vermittelt einen völlig irreführenden Eindruck von der praktischen Anwendung der Quantenmechanik, der nichts mit der Realität zu tun hat.***

Es wurde bereits erwähnt, daß der Zugang zur Quantenmechanik über rotierende Zeiger für die Physik als Fachwissenschaft ohne Bedeutung ist. Das mag etwa für das Elementarmagnete-Modell ebenso der Fall sein, aber letzteres ist an den Universitäten immerhin bekannt und hat im dortigen Unterricht in entsprechend subtilerer Ausführung auch seinen Platz. Darüberhinaus erhebt es an keiner Stelle den Anspruch, den Schülerinnen und Schülern den eigenen aktiven Umgang mit einem modernen physikalischen Theoriegebäude zu ermöglichen, wie das von den Anhängern der rotierenden Zeiger für dieselben verbreitet getan wird. Daraus ergibt sich ein Kritikpunkt, der den allgemeinen und den fachlich bildenden Aspekt des Physikunterrichts am Gymnasium gleichermaßen betrifft: Wer die Quantenmechanik am Gymnasium mit Hilfe des Zeigerformalismus kennengelernt hat, nimmt im Hinblick auf die Formulierung zentraler Aussagen wie auch die wissenschaftliche Tätigkeit in dieser Sparte der Physik einen Eindruck mit, der nicht das Geringste mit der Wirklichkeit zu tun hat. Bleibt damit das Bild der Quantenmechanik bei denjenigen, die nach dem Abitur nichts mehr damit zu tun haben, unkorrigiert falsch, was schon schlimm genug ist, so erleben die anderen, die, in welcher Form auch immer, mit der Physik weitermachen, beim ersten Wiedersehen mit der Quantenmechanik an der

<sup>133</sup>Es sollte bereits hier betont werden, daß Wellenfunktionen und die Schrödingergleichung nicht nur, weil sie im Lehrplan stehen, für den gymnasialen Physikunterricht unentbehrlich sind. Auch hier verweisen wir für Details auf Kapitel 4.

<sup>134</sup>Siehe Abschnitt 2.2.3.



Hochschule ihr blaues Wunder, da sie einerseits selbst nichts wiedererkennen und andererseits eine Sprache sprechen, die ihre künftigen Lehrer nicht verstehen oder günstigenfalls ziemlich merkwürdig finden.

In diesem Licht erscheint die Behauptung, mit dem Zeigerformalismus vermittele man „Lerninhalte, von denen später nichts mehr zurückgenommen werden muß“ [37], als völlig unhaltbar. Der Zugang über rotierende Zeiger liefert nichts, worauf später aufgebaut werden könnte, so daß insbesondere die methodischen Lerninhalte vielmehr komplett wieder zurückgenommen werden müssen, will man sich weitergehend mit der Quantenmechanik beschäftigen.

Angesichts der Verbreitung, auf die der Zeigerformalismus als Zugang zur Quantenmechanik im Physikunterricht der gymnasialen Oberstufe mittlerweile nicht nur in Lehrbüchern, sondern vor allen Dingen auch in der Referendarsausbildung an den Seminaren und in Fortbildungen verweisen kann, erscheint ein energisches Entgegenreten unerlässlich. Wir denken, ausführlich gezeigt zu haben, daß dieses Konzept für eine genügend elementarisierte und dennoch fachlich korrekte Darstellung quantenmechanischer Sachverhalte in der Schule *ungeeignet ist*, aus fachwissenschaftlichen, philosophischen wie auch fachdidaktischen Gründen. Es sollte weder an der Notwendigkeit der Berücksichtigung der Quantenmechanik im schulischen Physikunterricht noch an der dringenden Erfordernis einer Modernisierung traditioneller Wege der Physikdidaktik dahin gezweifelt werden; unsere Auffassung, wie diese Modernisierung aussehen sollte, weicht jedoch stark von der Auffassung derer ab, die im Moment die Meinungshoheit darüber zu haben scheinen. Wir plädieren nachdrücklich dafür, beim Unterrichten der Quantenmechanik in der Oberstufe auf rotierende Zeiger zu verzichten. Anders als beim Karlsruher Physikkurs ist das jedoch explizit nicht mit der Aufforderung verbunden, stattdessen bei traditionellen Wegen zu bleiben, denn während sich dort solche Zugänge über Jahrzehnte bewährt haben, ist hier wie gesagt ein Umdenken notwendig, sowohl aufgrund von sehr weitgehenden Fehlinterpretationen und schwerwiegenden fachlichen Fehlern bei herkömmlichen Unterrichtsgängen<sup>135</sup> als auch von bedeutenden theoretischen und experimentellen Neuentdeckungen der jüngsten Vergangenheit, die noch nicht Eingang in die Schul- und Hochschulbuchliteratur gefunden haben. Dieses Umdenken sollte aber eben nicht über den Weg des Zeigerformalismus erfolgen. Wie ein Unterrichtsgang zur Quantenmechanik in der Oberstufe nach unserer Auffassung aussehen sollte, ist im übernächsten Kapitel dargestellt.

---

<sup>135</sup>Siehe dazu Kapitel 4.

# Kapitel 3

## Entropie in Klasse 10

Kritik sollte niemals nur destruktiv, sondern immer auch konstruktiv sein; deshalb lassen wir es nach ausführlichem Gemecker an aktuellen Konzepten des gymnasialen Physikunterrichts natürlich nicht einfach dabei bewenden, sondern stellen auch zur Diskussion, wie man es unserer Meinung nach anders und vor allem besser machen kann. Wir beschreiben im vorliegenden und im nächsten Kapitel zwei mehrfach erprobte Unterrichtsgänge, die sich mit den beschriebenen Themenfeldern beschäftigen, eine Einheit zur Wärmelehre für die neunte oder zehnte Klasse<sup>1</sup> sowie eine zur Quantenmechanik in der Kursstufe. Die Einheiten unterscheiden sich drastisch von dem, was zur Zeit in Schulbüchern, Aus- und Fortbildung propagiert wird, sie tragen damit aber genau jenen Anforderungen Rechnung, die wir in den beiden vorigen Kapiteln aufgestellt haben und denen die aktuellen Produkte der Physikdidaktik nicht gerecht werden.

Wir stellen in diesem Kapitel eine Unterrichtseinheit zur Einführung der Entropie in der Mittelstufe vor, welche die in Kapitel 1 aufgeführten Defizite des Karlsruher Physikurses konsequent vermeidet. Begriffe werden klar definiert, und das Konzept der unterschiedlichen Energieformen und deren Umwandlungen steht immer im Vordergrund. Sehr wesentlich ist dabei, daß wir mit der statistischen Interpretation der Entropie beginnen, natürlich nur auf qualitativer Ebene. Sie allein ermöglicht eine einigermaßen anschauliche Vorstellung dieses Begriffs. Dadurch wird insbesondere und von Beginn an verhindert, daß sich bei den Schülerinnen und Schülern Vorstellungen der Entropie als irgendeiner Substanz oder sonst etwas geheimnisvollem einschleichen. Jeder anfänglich rein thermodynamische Zugang zur Entropie kann aufgrund seiner phänomenologischen Natur nicht erklären, was man sich unter diesem Begriff vorzustellen hat und provoziert damit unweigerlich genau solche Fehlinterpretationen. Erst anschließend und auf die qualitative statistische Begriffserklärung aufbauend kommen wir zu einer elementaren thermodynamischen Definition der Entropie, die von da an im Mittelpunkt steht, wobei jedoch von Zeit zu Zeit auch wieder auf die statistische Betrachtungsweise zurückgegriffen wird. Das Gelernte wird anschließend durch einige praktische Beispiele illustriert. Dabei wird jedoch konsequent darauf verzichtet, die Berechnung irgendwelcher Entropiedifferenzen als Selbstzweck vorzunehmen; solche Aufgabenstellungen müssen stets als Hilfsmittel zur Beschreibung anderer, konkreter Sachverhalte deutlich gemacht werden.

---

<sup>1</sup>Die neuen Bildungspläne für Gymnasien in Baden-Württemberg geben Inhalte nur noch in Zwei-Jahres-Blöcken vor, so daß man hier Spielraum hat. Ob das sinnvoll ist, darf wohl bezweifelt werden, insbesondere weil dadurch ein Schulwechsel unter Umständen außerordentlich erschwert werden kann.

Der eigentliche zentrale Inhalt der Unterrichtseinheit ist der zweite Hauptsatz der Thermodynamik, und entsprechend ist sie aufgebaut. Dabei werden mehrere Formulierungen des zweiten Hauptsatzes entwickelt, zunächst aus statistischer, danach aus thermodynamischer Sicht. Als optionale Anteile tauchen dabei der Carnotsche Wirkungsgrad sowie über das eigentliche Thema hinausgehend auch der dritte Hauptsatz der Thermodynamik auf. Abschließend werden Anwendungen auf technische und natürliche thermodynamische Systeme diskutiert.

### 3.1 Vorbemerkungen

Das Vorkommen des Begriffs der Entropie im Pflichtprogramm des gymnasialen Physikunterrichts ist derzeit als Tatsache hinzunehmen; wir kommen dennoch nicht daran vorbei, auf damit unvermeidlich einhergehende Probleme hinzuweisen. Diese treten unabhängig von der gewählten Unterrichtsphilosophie auf und sind allein auf die fachliche Nichttrivialität des Gegenstands zurückzuführen. Will man diesen vernünftig in den Unterricht einbringen, ist man gut beraten, vorher über solche Aspekte nachzudenken. Über die grundsätzlichen begrifflichen Schwierigkeiten, die eine abstrakte statistische Größe wie die Entropie mit sich bringt, wurde bereits ausführlich geredet. Hier soll es wesentlich um praktische Belange, das heißt um die Messung und Berechnung dieser Größe gehen, denn damit hat man es nicht zuletzt im schulischen Physikunterricht ausgiebig zu tun.

Die Messung der Entropie als absolute Größe ist überhaupt nicht und die Messung von Entropiedifferenzen nur mit beträchtlichem Aufwand möglich. Das liegt an ihrer Eigenschaft, keine Erhaltungsgröße zu sein, woraus folgt, daß beim Versuch, einen Entropiestrom durch ein wie auch immer aufgebautes „Entropiemeßgerät“ zu leiten, im allgemeinen zusätzliche Entropie entsteht. Man kann das umgehen, indem man gerade den Prozeß der Entropieerzeugung zur Messung der dabei auftretenden Entropiedifferenz nutzt, etwa wenn man ein System zunächst abkühlt und dann adiabatisch mit einer Energiequelle bekannter Leistung wieder auf die Ausgangstemperatur erwärmt, aber das ist eine indirekte Methode, da man dabei die erbrachte Leistung sowie die Zeit mißt, während der das geschieht, um daraus und aus der abgegebenen Wärme die Entropiedifferenz zu berechnen. Das ist eigentlich trivial, für den Physikunterricht der neunten oder zehnten Klasse jedoch nicht zu leisten, worauf wir sogleich zu sprechen kommen.

Bei solchen und ähnlichen Prozessen kommt die thermodynamische Entropiedefinition

$$dS = \frac{dQ}{T}$$

zum Einsatz. Ist die Leistung  $P$  bekannt, mit welcher während dem Zeitintervall  $\Delta t = t_2 - t_1$  Wärme übertragen wird, so erhält man die Entropiedifferenz

$$\Delta S = \int_{Q_1}^{Q_2} \frac{dQ}{T} = P \int_{t_1}^{t_2} \frac{dt}{T}.$$

Diese Integration kann nur umgangen werden, wenn die dabei auftretende Temperaturdifferenz klein gegen die absolute Temperatur des Systems ist. In diesem Fall schreibt

man näherungsweise

$$\Delta S \approx \frac{\Delta Q}{T} = P \frac{\Delta t}{T},$$

wobei man für  $T$  beispielsweise die mittlere Temperatur einsetzen kann<sup>2</sup>; nur dieser Spezialfall kann in einem Unterrichtsgang wie dem hier beschriebenen betrachtet werden. Das schränkt natürlich die Möglichkeiten für sinnvolle quantitative Aufgaben stark ein.

Nach unserer Ansicht ist das ein weiterer Anlaß, auf eine thermodynamische Betrachtung der Entropie im gymnasialen Physikunterricht komplett zu verzichten, und sich damit zu begnügen, eine einfache statistische Definition dieses Begriffs zur Grundlage eines entsprechenden Unterrichtsganges zu machen. Thermodynamische Aspekte der Entropie sollten eigentlich im Chemieunterricht verbleiben, wo sie besser aufgehoben sind, etwa bei der Fragestellung, ob chemische Reaktionen von selbst ablaufen oder nicht. Da der Bildungsplan eine solche Selbstbeschränkung jedoch nicht zuläßt, führen wir im vorliegenden Unterrichtsgang die Entropie trotz der erwähnten Bedenken und unter Verwendung der beschriebenen Näherung dennoch ein. Der Bildungswert dieser Thematik darf gleichwohl bezweifelt werden; alle wesentlichen Erkenntnisse, die man dabei den Schülerinnen und Schülern nahebringen möchte, etwa im Zusammenhang mit der vieldiskutierten „Energieentwertung“ und der Frage, warum es kein Perpetuum Mobile zweiter Art gibt, kann man genauso gut auch ohne den thermodynamischen Entropiebegriff erreichen.

Vernünftigerweise sollte die Einführung der Entropie im achtjährigen Gymnasium nach unserer Auffassung in der zehnten Klasse erfolgen. Das entspricht vom Ablauf her dem Zeitraum, in dem sie – wenn überhaupt – auch bisher erfolgte, nur eben um ein Jahr nach vorne verschoben, und ist schon aus diesem Grund früh genug. Allerdings ist wohl zu befürchten, daß die Entropie in Anlehnung an die erwähnten Umsetzungsbeispiele des Kultusministeriums verbreitet bereits in der neunten Klasse auftauchen wird. Unser Unterrichtsvorschlag ist folglich so aufgebaut, daß er gleichermaßen für die neunte und die zehnte Klasse geeignet ist.

Der hier beschriebene Unterrichtsgang ist recht umfangreich. Die Lehrerin oder der Lehrer muß jedoch nicht unbedingt den kompletten Stoff durchnehmen, insbesondere dann, wenn die Einheit bereits in der neunten Klasse auf dem Plan steht. Abschnitt 3.3 und Abschnitt 3.4 bis einschließlich 3.4.2 stellen bereits eine komplette Einführung in die Thematik dar, die etwa den Anforderungen des Lehrplans in Baden-Württemberg genügt. Der Rest ist optional, kann aber bei genügend zur Verfügung stehenden Stunden und vor allem beim Unterrichten der Einheit in der zehnten Klasse problemlos hinzugefügt werden.

## 3.2 Notwendige Vorkenntnisse

Da Physik im achtjährigen Gymnasium von der siebten Klasse an unterrichtet wird, ist die neunte Klasse bereits das dritte Lernjahr. Das bedeutet jedoch nicht, daß man für die Einführung der Entropie davon ausgehen kann, schon sehr viel mehr Grundlagen zur Verfügung zu haben als bisher. Das Gegenteil ist der Fall, denn die Wärmelehre wurde im

---

<sup>2</sup>Diese Relation ist natürlich exakt, wenn bei dem dabei beschriebenen Prozeß die Temperatur konstant gehalten wird.

neunjährigen Gymnasium üblicherweise auf drei Klassenstufen verteilt unterrichtet, auf die achte, neunte und elfte Klasse. Wenn die Entropie überhaupt zur Sprache kam, dann höchstens in der elften Klasse, so daß nun von einer um zwei Jahre früheren Einführung dieses Begriffes im Unterricht ausgegangen werden muß. Entsprechend setzen wir nur gerade das voraus, was auch in der bisherigen neunten Klasse schon vorausgesetzt werden konnte. Die Schülerinnen und Schüler sollten<sup>3</sup>

- den Begriff der Temperatur und des absoluten Nullpunkts kennen,
- diese sowie die Aggregatzustände, deren Übergänge und den Begriff der thermischen Ausdehnung im Rahmen des Teilchenmodells elementar erklären können,
- etwas von thermischem Gleichgewicht gehört haben (Nullter Hauptsatz der Thermodynamik),
- über Arbeit, Energie und den Energieerhaltungssatz (erster Hauptsatz der Thermodynamik) sowie damit zusammenhängend über Energieformen und deren Umwandlungen Bescheid wissen und
- mit den Begriffen der inneren Energie und der Wärme etwas anfangen und diese unterscheiden können.

Die Begriffe der Arbeit und der Wärme werden hier auch deswegen besonders angesprochen, da wie bereits erwähnt oft vehement gefordert wird, sie ganz wegzulassen. Wir halten sie jedoch für unverzichtbar. Ähnliches gilt für Teilchenmodelle, die häufig in die zehnte Klasse oder auf noch spätere Zeitpunkte verbannt werden. Sie können gar nicht früh genug eingeführt werden. Schon im Naturphänomene-Unterricht der fünften und sechsten Klasse sollten Elektronen, Atome und Moleküle zur Sprache kommen, nicht zuletzt deswegen, weil Schülerinnen und Schüler üblicherweise diese Begriffe bereits mitbringen.

## 3.3 Der zweite Hauptsatz der Thermodynamik

### 3.3.1 Entropie als Maß des inneren Durcheinanders

Unsere Unterrichtseinheit zur Einführung des Begriffs der Entropie beginnen wir mit einem Versuch, nämlich mit der Diffusion eines Tropfens Tinte in Wasser. Dabei macht man folgende Beobachtungen:

- Die Tinte *verteilt sich allmählich im Wasser*,
- das geschieht *von ganz allein* und
- der umgekehrte Prozeß tritt erfahrungsgemäß *nicht* auf.

Auf den ersten Blick ist klar was hier passiert; aufgrund der ungeordneten Bewegung der Wassermoleküle werden die Tintenteilchen herumgeschubst<sup>4</sup> und verteilen sich. Warum

<sup>3</sup>Wir verwenden hier mit Bedacht den Konjunktiv, um den äußerst lästigen, stets im Indikativ verfaßten Formulierungen der einschlägigen Bildungspläne, mit denen behauptet wird, was die Schülerinnen und Schüler angeblich alles *können, wissen und machen* anstelle von *können, wissen und machen sollen*, auch stilistisch entgegenzutreten.

<sup>4</sup>Falls die *Brownsche Bewegung* behandelt wurde, kann man jetzt darauf zurückgreifen.

dabei jedoch schließlich eine *völlig gleichmäßige Verteilung auf den ganzen Behälter* eintritt, bedarf der weiteren Erklärung.

Nun kann man die Klasse fragen, wieviele Möglichkeiten des Verhaltens der Tintenteilchen es jeweils für den Zustand

$$Z_1 = \text{„Tintentropfen in einer Ecke des Behälters“}$$

und für den Zustand

$$Z_2 = \text{„Tinte gleichmäßig über den gesamten Behälter verteilt“}$$

gibt<sup>5</sup>. Man kommt schnell zum Ergebnis, daß auch eine kleine Menge Tinte sehr viele Tintenteilchen enthält und es folglich für  $Z_1$  sehr sehr viele, für  $Z_2$  jedoch noch *unvorstellbar viel mehr* Möglichkeiten gibt. Damit ist die Sache klar: Da das Herumschubsen der Tintenteilchen und folglich deren Verhalten rein zufällig erfolgt, und Zustände des Systems um so wahrscheinlicher sind, je mehr unterschiedliche Möglichkeiten des Verhaltens der Teilchen diese Zustände realisieren, entwickelt sich das System so, daß sich die Tintenteilchen so gut wie möglich verteilen.

Folgende Modellrechnung verdeutlicht das sehr eindrucksvoll: Wir nehmen an, 1,66 Mol, also  $10^{24}$  Moleküle eines idealen Gases befinden sich in einem Würfel mit einem Kubikmeter Volumen. Wir denken uns diesen Würfel in 1000 kleine Würfel von je einem Liter unterteilt und betrachten die Möglichkeiten, wie sich die Gasmenge auf diese Würfel verteilen können. Dabei gilt generell:

*M Moleküle haben in N Würfeln  $N^M$  Möglichkeiten, sich unterschiedlich zu verteilen.*

Anfangs sei die gesamte Gasmenge in dem Würfel links unten. Dann ist  $N = 1$  und es gibt genau diese eine Möglichkeit. Bei zwei Würfeln sind es bereits  $2^{10^{24}} \approx 10^{3 \cdot 10^{23}}$  Möglichkeiten und bei drei  $3^{10^{24}} \approx 10^{5 \cdot 10^{23}}$ , also  $10^{2 \cdot 10^{23}}$  mal so viele wie bei zwei. Bei 999 Würfeln sind es  $999^{10^{24}} \approx 10^{2,9996 \cdot 10^{24}}$  und bei allen tausend Würfeln sind es  $1000^{10^{24}} = 10^{3 \cdot 10^{24}}$  Möglichkeiten. Damit sind es bei 1000 Würfeln ungefähr  $4 \cdot 10^{20}$  so viele Möglichkeiten wie bei 999. Daraus folgt zweierlei:

- Eine Rückkehr aller Moleküle in die Ausgangslage im Würfel links unten ist angesichts ihrer ungeordneten Bewegungen im Prinzip jederzeit möglich. Dazu müßten sie aber unter den  $10^{3 \cdot 10^{24}}$  Möglichkeiten die eine richtige erwischen. Angesichts der Zufälligkeit der Bewegungen ist das natürlich extrem unwahrscheinlich.
- Die bei weitem wahrscheinlichste Anordnung ist die Gleichverteilung auf alle 1000 Würfel, da schon eine Verteilung auf nur 999 davon, also auf 99,9 % des Gesamtvolumens des großen Würfels, nur noch eine Anzahl von Möglichkeiten zuläßt, die gerade einmal 0,000 000 000 000 000 25 % der Anzahl der Möglichkeiten der Verteilung auf 1000 Würfel beträgt.

Dieses Gedankenexperiment unterbietet die Realität natürlich immer noch beträchtlich; die Unterteilung des Würfels in 1000 Teilwürfel ist eine sehr grobe Näherung, die umso

<sup>5</sup>Es erscheint im Interesse der Begriffsökonomie fraglich, ob es sinnvoll ist, hier von *Mikrozuständen* und *Makrozuständen* zu sprechen.



besser wird, je größer die Anzahl der Teilwürfel gewählt wird. Zusätzlich muß man auch die Geschwindigkeiten der Moleküle und genaugenommen auch noch sonstige Variablen wie Rotations- Schwingungs- und elektronische Anregungen berücksichtigen, wenn man alle möglichen Zustände des Systems ermitteln will. Zusammengenommen folgt daraus, daß die Zahl der Möglichkeiten bei der Verteilung der Moleküle noch viel, viel größer ist als oben abgeschätzt und die Wahrscheinlichkeit des Eintretens irgendeines anderen Zustands als der Gleichverteilung entsprechend noch viel geringer. Wir folgern daraus:

- *Ein sich selbst überlassenes System erfährt stets Zustandsänderungen, bei denen die Anzahl der unterschiedlichen Möglichkeiten des Verhaltens der Teilchen zunimmt.*
- Die Wahrscheinlichkeit einer zufälligen Rückkehr in den Ausgangszustand ist unvorstellbar gering.
- Die Wahrscheinlichkeit der schließlichen Entwicklung zu einem völlig ungeordneten Zustand (Gleichverteilung) ist praktisch gleich 1.

Die Vorstellung eines ungeordneten Gewimmels von Teilchen legt noch eine weitere Formulierung desselben Sachverhalts nahe:

*Zustandsänderungen erfolgen ohne äußeren Einfluß stets so, daß das innere Durcheinander des Systems zunimmt.*

Denn je größer das innere Durcheinander eines Systems, desto mehr Möglichkeiten gibt es für das Verhalten der das System konstituierenden Teilchen<sup>6</sup>.

Die beiden kursiv gedruckten Aussagen sind genaugenommen bereits spezielle Formen des zweiten Hauptsatzes der Thermodynamik. An dieser Stelle kann nun der zentrale Begriff der Unterrichtseinheit erstmals und qualitativ zur Sprache kommen:

*Die Entropie ist ein Maß für das innere Durcheinander eines Systems.*

Wenn im Mathematikunterricht bereits der Logarithmus eingeführt wurde, kann man hinzufügen, daß die Entropie proportional zum Logarithmus der Anzahl der Möglichkeiten für das Verhalten der Teilchen ist, die es beim vorliegenden Zustand des Systems gibt (Boltzmannsches Prinzip). Davon wird hier jedoch im weiteren Verlauf kein Gebrauch gemacht.

---

<sup>6</sup>Wir weisen an dieser Stelle erneut auf den in bereits in Abschnitt 1.2.3 erwähnten Sachverhalt hin, daß der hier verwendete Begriff des inneren Durcheinanders eines Systems nichts mit fehlenden Ordnungsstrukturen oder Symmetrien zu tun hat. Ein näherungsweise ideales Gas hat beispielsweise eine sehr viel höhere Symmetrie als ein amorpher Kristall aus der gleichen Menge desselben Stoffes, dennoch ist die Entropie des Kristalls viel niedriger als diejenige des Gases. Genaugenommen ist das innere Durcheinander eines Systems im obigen Sinne gerade dadurch definiert, daß es als die Anzahl der Mikrozustände des Systems aufgefaßt wird, die dessen vorliegenden Makrozustand repräsentieren, womit wir einfach die im Physikunterricht verwendete Sprechweise wissenschaftlich präzisiert haben. Wenn Schülerinnen oder Schüler im Unterricht Fragen zu diesem Aspekt stellen, sollte man unbedingt ausführlich darüber reden. Genaueres dazu steht in Abschnitt 1.2.3.



### 3.3.2 Entropiedifferenzen

Es dürfte klar sein, daß man im allgemeinen keine Chance hat, konkrete quantitative Aussagen über die Anzahl der Möglichkeiten zu machen, die es für das Verhalten der Teilchen eines Systems gibt, um irgendeinen Zustand desselben zu realisieren. Wir brauchen daher eine andere Entropie-Definition, eine, die nicht von mikroskopischen, sondern von makroskopischen Eigenschaften physikalischer Systeme Gebrauch macht. Um zu einer solchen Definition zu gelangen, betrachten wir einen weiteren Versuch; bei diesem vergleichen wir Diffusionsvorgänge bei unterschiedlichen Temperaturen. Diesmal beobachtet man folgendes:

- Das innere Durcheinander eines Systems nimmt umso stärker zu, je mehr Wärme zugeführt wird und
- bei gleicher Wärmezufuhr nimmt es umso weniger zu, je höher die Temperatur des Systems ist.

Die Vermutung liegt nahe: *Das Maß des inneren Durcheinanders eines Systems ist proportional zur zugeführten Wärmemenge  $\Delta Q$  und indirekt proportional zur absoluten Temperatur  $T$ .* Wir definieren folglich:

$$\Delta S = \frac{\Delta Q}{T} \quad \text{heißt Entropiedifferenz.}$$

Dabei ist von entscheidender Bedeutung, im Zähler  $\Delta Q$  und nicht etwa  $\Delta E$  zu schreiben, denn es ist ausschließlich der *Eintritt von thermischer Energie*, also eben gerade *Wärmezufuhr*, die für ein Ansteigen des internen ungeordneten Gewimmels sorgt. Wird an dem System *Arbeit verrichtet*, etwa indem man es anhebt, mechanischer oder elektrischer Spannung aussetzt oder dergleichen, verstärkt sich allein dadurch dessen inneres Durcheinander *nicht*; das geschieht nur, wenn dabei zusätzlich dessen Temperatur erhöht wird, also wieder Wärme übertragen wird. Daraus folgt:

*Nur der Übergang von Wärme führt zur Erhöhung der Entropie eines Systems, nicht aber das Verrichten von Arbeit.*

Da Wärmezufuhr zu Temperaturerhöhung führt, erhalten wir einen weiteren Merksatz:

*Je höher die Temperatur eines Systems ist, desto größer ist der Wert von dessen Entropie.*

Kühlt ein Körper ab, so sinkt seine Entropie, wird er erwärmt, so nimmt sie zu. Das entspricht dem wohlbekanntem Sachverhalt, daß die ungeordnete Bewegung der Teilchen, aus denen ein System aufgebaut ist, und damit auch sein inneres Durcheinander um so stärker sind, je höher dessen Temperatur ist. Da hier leicht Mißverständnisse auftreten können, wiederholen wir den Zusammenhang zwischen Entropie und Temperatur kurz im Überblick: Die *Entropieänderung* bei Wärmezufuhr ist um so größer, je *tiefer* die Temperatur ist, die *Entropie* selbst ist um so größer, je *höher* die Temperatur ist. Das können wir quantitativ formulieren:

*Wird einem System mit der Temperatur  $T$  die Wärmemenge  $\Delta Q$  zugeführt, so wächst seine Entropie um den Betrag  $\Delta S = \Delta Q/T$ . Wird dem System die Wärmemenge entzogen, so wird der selbe Entropiebetrag abgegeben.*

Die hier beschriebene Relation für die Entropiedifferenz gilt natürlich nur, wenn die Temperatur konstant gehalten wird. Da das normalerweise nicht der Fall ist, haben wir es in konkreten Fällen meist mit einer Näherung zu tun, die nur zulässig ist, wenn die dabei auftretende Temperaturdifferenz  $\Delta T$  *klein im Vergleich zur absoluten Temperatur  $T$  ist*. Trifft das nicht zu, muß man über  $dQ/T$  integrieren, so daß diese allgemeinere Situation im Unterricht der neunten oder zehnten Klasse nicht behandelt werden kann. Diese Einschränkung sollte dort daher auch unbedingt erwähnt werden.

Zwei zusätzliche Bemerkungen sind hier angebracht, auch wenn sie im Unterricht nicht speziell thematisiert werden müssen beziehungsweise können. Zum einen können in der Thermodynamik wie in der klassischen statistischen Mechanik nur Aussagen über Entropiedifferenzen gemacht werden, nicht aber über absolute Werte der Entropie. Wir können absolute Entropiewerte nur vergleichen, sowohl bei ein und demselben System vor und nach Zustandsänderungen, als auch von unterschiedlichen Systemen. Das ändert sich erst in der Quantenmechanik, wo absolute Werte der Entropie zwar nicht naturgegeben festgelegt sind, aber immerhin mit guten physikalischen Gründen festgelegt werden können. Die Grundlage dafür liefert der dritte Hauptsatz der Thermodynamik, der bei genauer Betrachtung rein quantenmechanischer Natur ist. Zum anderen muß man sich bei der Einführung der thermodynamischen Entropie auf die hier geschilderte heuristische Überlegung beschränken; der formale Nachweis der Äquivalenz der beiden Entropiedefinitionen ist mit den Mitteln der Schulphysik nicht möglich<sup>7</sup>.

### 3.3.3 Der zweite Hauptsatz

Damit sind wir soweit, eine präzise Formulierung des zweiten Hauptsatzes vorzunehmen.

*Zweiter Hauptsatz der Thermodynamik: Die Entropie eines abgeschlossenen Systems kann zwar zunehmen, aber nicht abnehmen.*

Das ist nichts anderes als eine Zusammenfassung der beiden vorigen Abschnitte in einem Satz.

Der zweite Hauptsatz der Thermodynamik erklärt, warum gewisse Vorgänge, die gemäß dem Energieerhaltungssatz eigentlich möglich sein müßten, trotzdem bisher noch niemals beobachtet wurden. Er legt damit eine Zeitrichtung fest, nämlich diejenige in Richtung der Zunahme der Entropie. Beispielsweise verstößt es in keiner Weise gegen den Energieerhaltungssatz, wenn ein Gegenstand seine Umgebung abkühlt, die dadurch gewonnene thermische Energie in Bewegungsenergie und weiter in Lageenergie umwandelt und von selbst hochspringt, oder wenn sich an einer bestimmten Stelle in einem Eimer mit Wasser eine bestimmte Wassermenge unter Wärmeabgabe an das umgebende Wasser abkühlt und spontan einen Eiswürfel bildet. Das selbe gilt ganz generell für Vorgänge, die sich „rückwärts“ abspielen. In all diesen Fällen würde die Entropie von System und Umgebung *abgesenkt*, was dem zweiten Hauptsatz widerspricht. Dieser bestimmt

<sup>7</sup>Wie das geht wurde in Abschnitt 1.2.4.2 gezeigt.

offensichtlich in über die Bedeutung des Energiesatzes hinausgehendem Maße, welche physikalischen Vorgänge in der Realität auftreten können und welche nicht. Vorgänge, die in diesem Sinne nur in eine Richtung ablaufen können, nennt man *irreversible Vorgänge*. Der zweite Hauptsatz sagt folglich, welche Prozesse reversibel und welche irreversibel sind.

Man darf jedoch nicht vergessen, daß es sich hierbei nicht um ein Naturgesetz von vergleichbarer Fundamentalität wie beispielsweise dem Energierhaltungssatz oder dem Trägheitssatz handelt, sondern eher um einen Erfahrungssatz. Das ist in dem Sinne zu verstehen, daß Vorgänge, die vom zweiten Hauptsatz ausgeschlossen werden, nicht im strengen Sinne unmöglich sind, sondern die Wahrscheinlichkeit ihres Eintretens so unvorstellbar gering ist, daß so schnell nicht damit zu rechnen ist, daß sie sich ereignen. Das ist im übrigen ein Standardbeispiel für eine Aussage, die unter Verwendung der thermodynamischen Entropie unverständlich bleibt und postuliert werden muß, mit der statistischen Entropiedefinition dagegen geradezu selbstverständlich erscheint.

## Aufgaben

1. Eine Schule hat 40 Klassenzimmer und wird von 1200 Schülerinnen und Schülern besucht.
  - a) Eines der Klassenzimmer wird gerade renoviert. Wieviele Möglichkeiten haben die Schülerinnen und Schüler, sich auf die restlichen Klassenzimmer zu verteilen?
  - b) Wieviele Möglichkeiten haben sie, wenn sie sich wieder auf alle Klassenzimmer verteilen dürfen?
2. Ein auf dem Boden liegender Stein springt spontan in die Höhe.
  - a) Erkläre unter alleiniger Berücksichtigung des ersten Hauptsatzes der Thermodynamik (Energieerhaltungssatz), wie das gehen könnte.
  - b) Erkläre unter zusätzlicher Berücksichtigung des zweiten Hauptsatzes der Thermodynamik, warum so etwas normalerweise nicht geschieht.
3. Ein Kubikmeter Wasser soll bei einer Umgebungstemperatur von 300 K um 10 K erwärmt werden.
  - a) Berechne die Zunahme der Entropie, wenn die Anfangstemperatur des Wassers 290 K beträgt.
  - b) Berechne die Zunahme der Entropie, wenn die Anfangstemperatur des Wassers 310 K beträgt.
  - c) In beiden Fällen nimmt die Entropie zu. Warum läuft der erste Prozeß spontan ab, der zweite jedoch nicht?

## 3.4 Nutzbare und verlorene Arbeit

Wir können die bisherigen Resultate nun zur Erklärung eines Sachverhalts heranziehen, der für technische Zivilisationen wohl das größte Problem überhaupt darstellt. Es geht um die wohlbekanntete Tatsache, daß trotz des Energieerhaltungssatzes, nach dem Energie weder erzeugt noch vernichtet werden kann, die *Bereitstellung* der für solche Zivilisationen lebenswichtigen Energie große Schwierigkeiten bereiten kann – jedenfalls für unsere eigene. Es wird sich zeigen, daß die Ursache hierfür natürlich im zweiten Hauptsatz liegt, genauer gesagt in dessen Aussage, daß für die Entropie *kein Erhaltungssatz gilt*. Offensichtlich sind Energieumwandlungen nicht beliebig möglich, und die Entropie ist dafür zuständig, welche gehen und welche nicht.

### 3.4.1 Entropieübertragung und Entropieerzeugung

Wir untersuchen nun genauer, wie sich die Entropie von Systemen verändern kann. Wenn wir das zunächst mit der entsprechenden Situation bei der Energie vergleichen, so stellen wir fest: Energie kann von einem System auf ein anderes übergehen (häufig ändert sie dabei ihre Form), aber ihre Menge bleibt insgesamt gleich. Auch Entropie kann von einem System auf ein anderes übergehen; das ist jedoch nicht die einzige Möglichkeit zur Veränderung ihrer Menge. Dazu betrachten wir zwei Versuche.

Beim Versuch 1, einem klassischen, auf Gay-Lussac zurückgehenden Experiment<sup>8</sup>, geht es um Kontraktion beziehungsweise Expansion zweier Gasreservoirs, die in thermischem Kontakt stehen, also Wärme austauschen können. Dazu befinden sich die beiden Gasportionen in einem beidseitig durch verschiebbare Stempel verschlossenen Kolben und sind durch eine wärmeleitfähige Wand verbunden. Die eine wird komprimiert, die andere in gleichem Maße expandiert, und währenddessen wird beiderseits die Temperatur gemessen (siehe Abbildung 3.1). Dabei gilt folgendes:

- Wenn der Vorgang sehr langsam erfolgt, ändert sich die Temperatur nicht<sup>9</sup>.
- Die Temperatur des komprimierten Gases müßte sich eigentlich erhöhen und die des expandierten entsprechend absinken, es wird also Wärme übertragen.
- Da das Volumen des komprimierten Gases sinkt und das des expandierten steigt, sinkt die Anzahl der Möglichkeiten der Teilchenzustände in ersterem und steigt in letzterem.

Beides führt zum Schluß, daß *die Entropie im komprimierten Gas im gleichen Maße sinkt wie sie im expandierten Gas zunimmt und die Gesamtentropie somit gleichbleibt*.

Versuch 2 besteht im *Erwärmen eines Wasserreservoirs mit einem Tauchsieder*. Dem System „Behälter + Wasser + Tauchsieder“ wird elektrische Energie zugeführt, *es wird also Arbeit daran verrichtet*. Wie wir gesehen haben, *erhöht sich dadurch die Entropie*

<sup>8</sup>Genaugenommen handelt es sich hierbei um eine Modifikation des Original-Experiments von Gay-Lussac, das in Abschnitt 1.2.4.2 beschrieben wurde. Diese Variante wird jedoch in der Literatur unter dem selben Namen geführt.

<sup>9</sup>Das ist der Prototyp einer *reversiblen* Zustandsänderung, während die im folgenden betrachteten Vorgänge Beispiele für *irreversible* Zustandsänderungen darstellen.

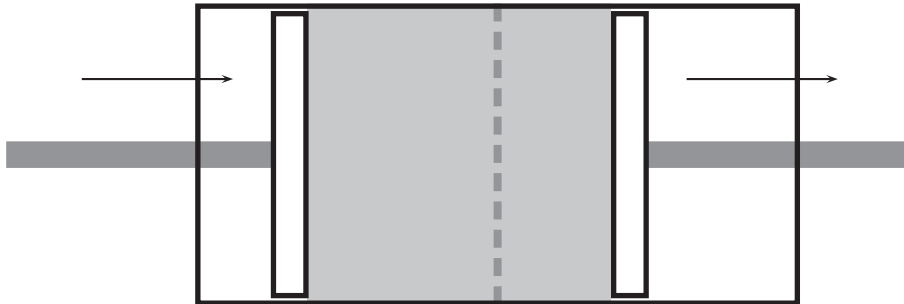


Abbildung 3.1: Zu Versuch 1 (Experiment von Gay-Lussac)

des Systems nicht. Gleichzeitig wird die elektrische Energie in dem System in thermische Energie umgewandelt, es wird ihm also *Wärme zugeführt*. Seine Entropie erhöht sich dadurch, ohne daß von außen welche dazugekommen wäre, es ist also *Entropie neu entstanden*. Für die Entropie gibt es offensichtlich keinen Erhaltungssatz.

Warum ist nun im zweiten Versuch zusätzliche Entropie dazugekommen? Man sieht sofort, daß der Unterschied in der in Versuch 1 fehlenden, in Versuch 2 jedoch vorhandenen *Temperaturdifferenz* zwischen den jeweils miteinander in Kontakt stehenden Systemen liegt. Genauer gesagt geht die Wärme in Versuch 2 von einem Teilsystem mit *höherer Temperatur* (dem Tauchsieder) auf Teilsysteme mit *tieferer Temperatur* (dem Wasser, dem Behälter und genaugenommen auch der Umgebung) über – so wie sie es stets freiwillig tut. Es handelt sich um ein typisches Beispiel eines Übergangs von einem Zustand, für den es weniger Möglichkeiten für das Verhalten der beteiligten Teilchen gibt, zu einem solchen mit mehr Möglichkeiten für die Teilchen und damit eines Übergang, der von selbst ablaufen kann, wie wir weiter oben gesehen haben. Gleichzeitig ist damit der Ursprung der neu entstandenen Entropie im Rahmen der statistischen Deutung, also auf mikroskopische Weise erklärt. *Offensichtlich entsteht Entropie immer dort neu, wo Wärme von einem System mit höherer Temperatur auf eines mit tieferer Temperatur übergeht*. Wir folgern also aus diesen beiden Versuchen:

- Entropie kann *übertragen* und *erzeugt* werden.
- *Nur der Übergang von Wärme* führt zur Übertragung von Entropie von einem System auf ein anderes, *nicht aber das Verrichten von Arbeit*.
- Wird einem System mit der Temperatur  $T$  die Wärme  $\Delta Q$  zugeführt, so wächst seine Entropie um  $\Delta S = \Delta Q/T$ . Erfolgt diese Wärmezufuhr über ein Temperaturgefälle hinweg, so wird diese Entropie *neu erzeugt*, was der intuitiven Vorstellung entspricht, daß hierbei die innere Unordnung des Systems zunimmt.

Das kann man quantitativ formulieren:

Wenn ein Teilsystem mit höherer Temperatur  $T_1$  und ein Teilsystem mit tieferer Temperatur  $T_2$  in thermischem Kontakt sind und dabei die Wärmemenge  $\Delta Q$  übertragen wird, dann wird im Gesamtsystem die Entropie

$$\Delta S = \frac{\Delta Q}{T_2} - \frac{\Delta Q}{T_1}$$

erzeugt, das heißt, die Entropie des Gesamtsystems nimmt zu.

Man erkennt daran sofort, daß Wärmeübertragung von einem kälteren auf ein heißeres System *Entropieerniedrigung* bedeutet. Damit erhalten wir eine alternative Formulierung des zweiten Hauptsatzes:

*Wärme kann nicht von selbst von einem System mit tieferer Temperatur auf ein System mit höherer Temperatur übergehen.*

Wir betrachten ein quantitatives Beispiel. Eine Heizung werde bei einer Heizkörpertemperatur von  $60^\circ\text{C}$  betrieben und halte die Raumtemperatur auf konstant  $20^\circ\text{C}$ . Für je  $1000\text{ J}$  Wärme, die die Heizung an die Raumluft abgibt, überträgt sie an dieselbe einen Entropiebetrag von

$$\Delta S_H = \frac{\Delta Q}{T_H} = \frac{1000\text{ J}}{333\text{ K}} = 3,00 \frac{\text{J}}{\text{K}}.$$

Gleichzeitig beträgt die Entropiezunahme im Raum

$$\Delta S_R = \frac{\Delta Q}{T_R} = \frac{1000\text{ J}}{293\text{ K}} = 3,41 \frac{\text{J}}{\text{K}}.$$

Netto hat man damit für je  $1000\text{ J}$  einen Entropiezunahme von

$$\Delta S = \Delta S_H - \Delta S_R = 0,41 \frac{\text{J}}{\text{K}}.$$

Die Entropie im gesamten System „Heizkörper + Raum“ nimmt also tatsächlich zu, es entsteht neue Entropie.

### 3.4.2 Verlorene Arbeit

Offensichtlich sorgen sämtliche Prozesse, bei denen Wärme im Spiel ist, für eine Zunahme der Entropie. Das geschieht vorzugsweise dadurch, daß sich Temperaturendifferenzen ganz oder teilweise ausgleichen, daß also Wärme selbständig von heißeren auf kältere Systeme übergeht. Man ahnt schon, wo hier ein Problem auftauchen muß: Häufig ist es erforderlich, dafür zu sorgen, daß Wärme den umgekehrten Weg geht, also Entropieerniedrigung zu erzwingen; da die Entropie insgesamt aber niemals abnehmen kann, muß sie dafür woanders überproportional zunehmen. Exemplarisch kann man das der Klasse an folgendem Versuch demonstrieren: Man betrachtet ein schweres Fadenpendel und stellt dabei fest, daß

- die Schwingungen des Pendels gedämpft sind und
- am Aufhängepunkt und damit auch in der Umgebung die Temperatur ansteigt.



Fragt man, was hier geschieht, so erkennen die Schülerinnen und Schüler sofort, daß aufgrund der Reibung im Aufhängepunkt (und genaugenommen auch ein wenig aufgrund des Luftwiderstandes) ein Teil der Lage- und Bewegungsenergie des Pendels als Wärme an das Pendel und an die Umgebung übertragen wird. Damit wird die Entropie des Pendels und der Umgebung erhöht. Wenn das Pendel ungedämpft schwingen soll, muß man ständig Energie hineinpumpen.

Warum ist das so? Warum kann sich das Pendel die erforderliche Energie nicht selbst aus der Umgebung zurückholen? Dazu müßte es diese jedoch abkühlen und folglich, da es selbst Umgebungstemperatur besitzt, die Entropie absenken, was den zweiten Hauptsatz verletzt.

Man könnte an dieser Stelle fragen, ob es mit einem tiefgekühlten Pendel funktionieren würde, oder allgemeiner gesagt, wenn die Temperatur des Pendels unterhalb der Umgebungstemperatur liegt. Aber das hilft auch nichts, denn dann könnte das Pendel zwar völlig legal unter Entropiezunahme von der Umgebung Energie abzapfen, aber nur in Form von Wärme. Diese müßte dann erst noch in mechanische Energie umgewandelt werden, wozu sich das Pendel abkühlen und damit seine Entropie absenken müßte. Wieder haben wir einen Verstoß gegen den zweiten Hauptsatz. Die als Wärme abgegebene und für die Entropieerhöhung verantwortliche Energie kann nicht zurückgewonnen werden und ist damit nicht mehr nutzbar. Dies können wir wieder allgemein formulieren:

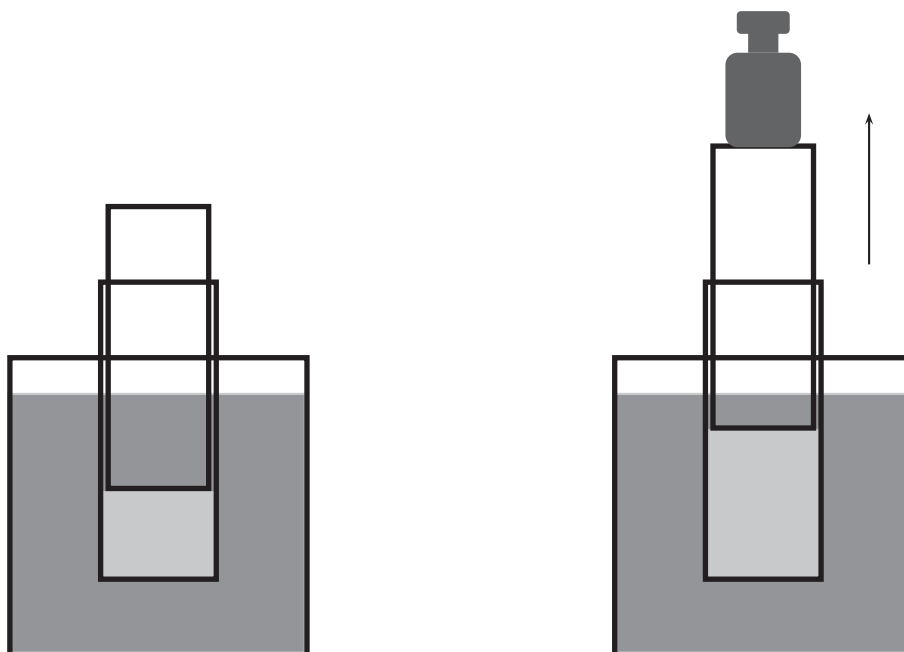
*Immer dann, wenn in einem abgeschlossenen System Entropie erzeugt wird, geht dabei ein Teil der Energie für die weitere Nutzung verloren, das heißt sie kann nicht für das Verrichten von Arbeit zurückgewonnen werden.*

Die nächste Frage an die Klasse drängt sich nun geradezu auf: Wohin hat sich die verlorene Energie verdrückt? Verschwunden sein kann sie nicht, das würde gegen den Energieerhaltungssatz verstoßen. Die Erklärung findet sich in der statistischen Deutung der Entropie. *Die Energie, die als Wärme an die kältere Umgebung abgegeben wird, verkrümelt sich auf die unvorstellbar vielen Möglichkeiten der die Umgebung bildenden Teilchen, wild durcheinander zu schwirren.* Wollte man sie zurückgewinnen, müßte man die Teilchen dazu bringen, sich geordnet zu verhalten, wofür es viel weniger Möglichkeiten gibt. Man hat sozusagen keine Chance, diese Möglichkeiten aus der Riesenanzahl der gesamten Möglichkeiten für das Verhalten der Teilchen herauszuklauben. Womit wir die nächste Formulierungsmöglichkeit für den zweiten Hauptsatz gefunden haben:

*Wärme kann nicht von selbst von einem Reservoir mit tieferer auf ein Reservoir mit höherer Temperatur übergehen.*

Idealerweise sollte von den Schülerinnen und Schülern jetzt der Vorschlag kommen, es mit einem Temperaturgefälle zu versuchen; dann müßte man doch die dabei freiwillig übertretende thermische Energie zum Verrichten von Arbeit veranlassen können. Das ist in der Tat möglich und wird in der Realität bekanntlich in unterschiedlichster Weise durchgeführt. Zur Veranschaulichung betrachten wir im nächsten Versuch exemplarisch die Umwandlung von thermischer Energie in mechanische Energie im Detail: Ein Rundkolben, der luftdicht mit einem senkrecht darüberstehenden Kolbenprober verbunden ist, wird nacheinander in zwei große Bechergläser gebracht, wobei das erste Wasser mit





Wasser mit Umgebungstemperatur

Heißes Wasser

Abbildung 3.2: Thermische Energie wird in mechanische Energie umgewandelt

Umgebungstemperatur und das zweite heißes Wasser enthält. Im ersten Fall passiert nichts, im zweiten bewegt sich der Kolben des Kolbenprobers nach oben und kann dabei ein Wägestück anheben (siehe Abbildung 3.2). Dabei sollte die Klasse zwei wesentliche Aspekte und vielleicht auch gleich noch eine Anwendungsmöglichkeit herausfinden:

- Ein Teil der thermischen Energie des Wassers wird *in mechanische Energie umgewandelt*, ein weiterer Teil *geht als Wärme an die Umgebung*.
- Der Vorgang funktioniert nur, *wenn die Temperatur des Wassers höher als diejenige der Umgebung und damit der Luft im Rundkolben und im Kolbenprober ist*.
- Wenn man das Wasser abkühlt, anschließend neu erhitzt und diesen Vorgang periodisch wiederholt, hat man eine einfache *Wärmekraftmaschine* gebaut, also eine Maschine, die ein heißes Wärmereservoir abkühlt und die abgegebene Wärme teilweise in mechanische Arbeit umwandelt und teilweise an ein kälteres Wärmereservoir weitergibt.

Insondane ist es dabei *nicht möglich, die gesamte der Temperaturdifferenz entsprechenden thermischen Energie in eine andere Energieform umzuwandeln*.

Die Konsequenzen aus alledem sind erheblich und betreffen Vorgänge in der Natur gleichermaßen wie solche technischer Art. Sie stellen eine fundamentale Einschränkung der möglichen Energieumwandlungen dar. Wir können das als eine weitere Form des zweiten Hauptsatzes auffassen:

*Es gibt keinen Prozeß, der thermische Energie vollständig in eine andere Energieform umwandeln kann.*

Die Begründung wurde oben in der Sprache des Teilchenmodells bereits gegeben; man kann diesen Sachverhalt jedoch auch makroskopisch verstehen. Soll thermische Energie in eine andere Energieform umgewandelt, also zum Verrichten von Arbeit genutzt werden, so muß sie aus dem Reservoir, von dem sie geliefert wird, heraus, also als *Wärme* übertragen werden. Das sollte natürlich von selbst passieren, was durch das erwähnte Temperaturgefälle und das zweite Reservoir, das eine tiefere Temperatur hat als das erste, gewährleistet wird. Sofern sich dieses nicht am absoluten Nullpunkt befindet, wird es stets nur einen von der Temperaturdifferenz abhängenden Betrag an Wärme aufnehmen können, der kleiner ist als der eigentlich im ersten Reservoir zur Verfügung stehende Betrag an thermischer Energie. Der absolute Nullpunkt ist bekanntlich von keinem System erreichbar, das heißt ein Mindestbetrag an nicht umwandelbarer thermischer Energie ist schon aufgrund der nicht verschwindenden Temperatur des zweiten Reservoirs unvermeidlich<sup>10</sup>. Nur der verbleibende Rest der abgegebenen Wärmemenge kann theoretisch in eine nichtthermische Energieform umgewandelt werden; praktisch jedoch nur ein Teil davon, denn der Rest wird stets in das zweite Reservoir übergehen, um dessen Temperatur zu erhöhen, bis die beiden Reservoirs die selbe Temperatur besitzen. Danach geht es nur weiter, wenn man das Temperaturgefälle wiederherstellt. In der Realität ist es meistens noch weniger, da stets auch ein Teil der nichtthermischen Energie (üblicherweise unerwünscht) zur Temperaturerhöhung des zweiten Reservoirs beiträgt und darüberhinaus meistens auch noch (ebenfalls unerwünscht) thermische Energie in Form von Wärme in die Umgebung abwandert. Oder um es in obiger Sprache zu formulieren: Die Umwandlung thermischer Energie in andere Energieformen ist stets mit einer mehr oder weniger großen, aber niemals verschwindenden Entropieproduktion und damit mit einer Temperaturerhöhung verbunden, so daß diese Umwandlung nie vollständig stattfinden kann.

Hierin liegt der Ursprung des umgangssprachlich verbreiteten, aber physikalisch gesehen inkorrekten Ausdrucks „Energieverbrauch“. Was sich dahinter verbirgt, ist folgendes: Energie kann als Erhaltungsgröße natürlich nicht verbraucht werden; ihre Verfügbarkeit im Sinne des Verrichtens von Arbeit bleibt jedoch nicht erhalten, sondern verschwindet ganz oder teilweise, wo immer Wärme über ein Temperaturgefälle wandert<sup>11</sup>. Obwohl Energie nicht verschwinden kann, muß sie für alle Vorgänge, die nicht von selbst ablaufen, also lokal mit Entropieerniedrigung verbunden sind, immer wieder neu verfügbar gemacht werden. Energiesparen bedeutet daher insbesondere, Entropie erzeugende Prozesse so weit wie möglich einzuschränken.

Zur Veranschaulichung kann man etwa Wärmekraftwerke mit Wasserkraftwerken vergleichen. Bei ersteren *wird thermische Energie in elektrische Energie umgewandelt*, was prinzipiell nicht vollständig erfolgen kann. Es entsteht zwangsläufig Abwärme, die durch verbesserte Technologie zwar minimiert, aber nicht vermieden werden kann. Bei letzte-

---

<sup>10</sup>Vergleiche Abschnitt 3.4.

<sup>11</sup>Es ist populär geworden, diesen Vorgang als *Energieentwertung* zu bezeichnen, womit man einen auf Poincaré zurückgehenden und später in der Fachsprache verschwundenen Begriff reaktiviert. Wir verzichten auf dessen Verwendung, denn dadurch wird ein zusätzlicher Fachausdruck vermieden, der nichts weiter erklärt; außerdem besteht andernfalls die Gefahr, daß der Wert der Energie mit deren Betrag verwechselt wird. Der physikalisch korrekte Begriff dafür lautet übrigens *dissipierte Energie*.

ren wird mechanische Energie in elektrische Energie umgewandelt. Zumindest im Idealfall wird dabei keine Entropie erzeugt, die Umwandlung kann vollständig erfolgen. In Wirklichkeit ist das nicht ganz der Fall, denn auch hier wird ein Teil der mechanischen Energie in Wärme umgewandelt (Reibung, Joulsche Wärme). Wie weit das gehen kann, zeigen Wasserfälle, bei denen praktisch die gesamte Lageenergie des herunterstürzenden Wassers zur Erhöhung der Umgebungstemperatur und damit zur Entropieproduktion herhalten muß.

Damit kann man es eigentlich bewenden lassen, denn den Anforderungen des Bildungsplanes wird dadurch voll und ganz Rechnung getragen. Wir sind nun in der Lage, die allgemeine Tendenz der Natur zu erklären, Energie unbrauchbar zu machen und können darüberhinaus eine ganze Menge praktischer Anwendungsbeispiele diskutieren. Abschnitt 3.5 beschreibt zwei Beispiele dafür; in den Übungsaufgaben findet man weitere. Möchte man formal und quantitativ weiter in die Tiefe gehen, kann man das Phänomen des Verlusts arbeitsfähiger Energie sehr viel detaillierter beleuchten. Man kann dann beispielsweise auch den Betrag der Energie berechnen, der in konkreten Situationen für das Verrichten von Arbeit verloren geht. Das führt natürlich zu einem deutlichen Anstieg des Schwierigkeitsgrades im Vergleich zu den bis hierher durchgenommenen Inhalten und sollte in jedem Fall erst in der zehnten Klasse stattfinden.

### 3.4.3 Der Carnotsche Wirkungsgrad

Um die angedeutete Vertiefung zu realisieren, betrachten wir erneut den im vorigen Abschnitt beschriebenen Versuch mit dem Rundkolben im wassergefüllten Becherglas und die drei erwähnten Folgerungen daraus. Die zweite Feststellung ist die entscheidende, wie wir gleich sehen werden. Die Temperatur des heißen Wassers sei  $T_1$ , diejenige der Umgebung  $T_2$ . Das heiße Wasser gibt die Wärmemenge  $\Delta Q_1$  an den Rundkolben und den Kolbenprober ab; ein Teil davon wird in mechanische Arbeit  $\Delta W$  umgewandelt, der Rest geht als Wärme  $\Delta Q_2$  an die Luft im Rundkolben und im Kolbenprober (sowie genaugenommen an diesen selbst und an die Umgebung). Das bedeutet einerseits eine *Entropieabgabe*

$$\Delta S_1 = \frac{\Delta Q_1}{T_1}$$

vom heißen Wasser und andererseits eine *Entropieaufnahme*

$$\Delta S_2 = \frac{\Delta Q_2}{T_2}$$

durch den Rundkolben und den Kolbenprober (und den Rest der Welt). Für die *gesamte Entropieänderung* folgt daraus und aus dem zweiten Hauptsatz

$$\Delta S = \frac{\Delta Q_2}{T_2} - \frac{\Delta Q_1}{T_1} \geq 0.$$

Setzt man die aus dem Energieerhaltungssatz folgende Relation  $\Delta Q_2 = \Delta Q_1 - \Delta W$  ein, erhält man

$$\Delta S = \frac{\Delta Q_1 - \Delta W}{T_1} - \frac{\Delta Q_1}{T_2} \geq 0.$$

Rechts ist offensichtlich nur dann ein Gleichheitszeichen erreichbar, wenn  $\Delta Q_1 = 0$  ist und damit keine thermische Energie übertragen wird; das entspricht genau dem, was weiter oben über die Erzeugung von Entropie gesagt wurde.

Da man sich dafür interessiert, wieviel von der übertragenen Wärme nutzbar ist und wieviel verloren geht, definiert man den *Wirkungsgrad*  $\eta$  als das Verhältnis aus nutzbarer und hineingesteckter Energie, also  $\eta = \Delta W / \Delta Q$ . Dabei gilt natürlich  $\eta \leq 1$ , aber  $\eta$  unterliegt einer schärferen Einschränkung. Setzt man nämlich  $\Delta W = \eta \Delta Q_1$  oben ein, findet man

$$\Delta S = \Delta Q_1 \left( \frac{1}{T_2} - \frac{1}{T_1} - \frac{\eta}{T_2} \right) \geq 0.$$

$\Delta Q_1$  ist stets positiv, also muß es die Klammer auch sein; das liefert

$$\eta \leq \frac{T_1 - T_2}{T_1} = 1 - \frac{T_2}{T_1}$$

wobei für reale Prozesse nur „<“, nicht aber „=“ eintreten kann.

Definieren wir nun den *Carnotschen Wirkungsgrad*  $\eta_C = 1 - T_2/T_1$  als den maximal theoretisch bei der Umwandlung von thermischer Energie in mechanische Arbeit möglichen Wirkungsgrad, so erhalten wir die nächste Formulierung des zweiten Hauptsatzes:

*Es gibt keinen realen Prozeß der Umwandlung von thermischer Energie, der den Carnotschen Wirkungsgrad erreicht oder überschreitet.*

Der umgekehrte Prozeß ist ganz offensichtlich unmöglich: Man kann keine Arbeit verrichten, indem man einem Wärmereservoir durch Abkühlung Energie entzieht, ohne gleichzeitig ein zweites Reservoir mit tieferer Temperatur zur Verfügung zu haben. Denn das würde eine Absenkung der Entropie bedeuten, ohne parallel dazu woanders Entropie zu erzeugen. Auch das liefert eine Variante des zweiten Hauptsatzes<sup>12</sup>:

*Es unmöglich, eine Maschine zu konstruieren, die Arbeit verrichtet, indem sie als Energiequelle ausschließlich ein Wärmereservoir abkühlt.*

Eine solche hypothetische Maschine nennt man ein Perpetuum Mobile zweiter Art, im Gegensatz zu einem Perpetuum Mobile erster Art, das den Energieerhaltungssatz austrickst. Als Beispiel kann man sich die Frage stellen, warum man einen Tank mit 3000 l Wasser der Temperatur 30° C nicht zum Kochen verwenden kann, obwohl doch sehr viel thermische Energie darin gespeichert ist. Oder wir stellen fest, daß ein Ozeandampfer nicht dadurch angetrieben werden kann, daß er durch Abkühlen des Meerwassers dieses riesige thermische Energiereservoir anzapft. In solchen Fällen betrachte man etwa den Carnotschen Wirkungsgrad  $\eta = 1 - T_2/T_1$ ; soll die einem Reservoir durch Abkühlung entzogene Energie zum Heizen, also zum Aufrechterhalten einer höheren Temperatur in einem zweiten Reservoir verwendet werden, wird der Wirkungsgrad negativ, was offensichtlich Unsinn ist.

<sup>12</sup>Dies Form ist strengenommen nicht vollständig äquivalent zu den anderen, da es in der Quantenmechanik Gegenbeispiele gibt.

Wir kommen nun zur angekündigten Berechnung des Betrags der bei einem Entropie erzeugenden Prozeß nicht mehr nutzbaren Energie. Dazu betrachten wir ein System mit der Temperatur  $T_1$ , dem die Wärmemenge  $Q$  zur Verrichtung von mechanischer Arbeit entnommen wird und verwenden die aus dem zweiten Hauptsatz in obiger Formulierung folgenden Relation für die von der Temperatur abhängige *maximal nutzbare Energie*  $E_{\max}(T)$ , die man beim Übergang der Wärmemenge  $Q$  von dem betrachteten System auf eines mit der Temperatur  $T_0 < T_1$  gewinnen kann,

$$E_{\max}(T_1) = Q \left( 1 - \frac{T_0}{T_1} \right).$$

Das zweite System dient zur Aufnahme der zwangsläufig entstehenden Abwärme  $Q - E_{\max}$ . Typischerweise handelt es sich hier um die Umgebung des ersten Systems, weswegen wir  $T_0$  als *Umgebungstemperatur* bezeichnen. Wenn nun  $Q$  vom ersten System auf ein weiteres System mit der Temperatur  $T_2 < T_1$  übertragen wird, um damit Arbeit zu verrichten, beträgt die maximal nutzbare Energie nur noch

$$E_{\max}(T_2) = Q \left( 1 - \frac{T_0}{T_2} \right) < E_{\max}(T_1).$$

Die verwertbare Energie sinkt bei diesem Vorgang also um

$$\begin{aligned} \Delta E &= E_{\max}(T_1) - E_{\max}(T_2) = Q \left( 1 - \frac{T_0}{T_1} \right) - Q \left( 1 - \frac{T_0}{T_2} \right) \\ &= \frac{QT_0}{T_2} - \frac{QT_0}{T_1} = T_0 \left( \frac{Q}{T_2} - \frac{Q}{T_1} \right) = T_0 \Delta S, \end{aligned}$$

die nicht mehr verwertbare Energie ist also das Produkt aus Umgebungstemperatur und Entropiezuwachs. Vorausgesetzt werden muß dabei natürlich, daß nicht irgendwie anderweitig Energie zugeführt oder abgezweigt wird. Wir haben also folgenden Merksatz:

*Wird bei einem ungestörten Prozeß bei der Umgebungstemperatur  $T_0$  die Entropie  $\Delta S$  erzeugt, so geht dabei Energie vom Betrag*

$$\Delta E = T_0 \Delta S$$

*für die Verrichtung von Arbeit verloren.*

Das Attribut „ungestört“ ist hier besonders wichtig. Die Größe  $\Delta E$  ist derjenige Anteil der an die Umgebung abgegebene Wärme, der nach Maßgabe des zweiten Hauptsatzes *maximal in mechanische Arbeit umgewandelt werden könnte*, sich jedoch in die Umgebung verflüchtigt, wenn man den betrachteten Prozeß einfach so passieren läßt, und der, wenn er erst einmal dort ist, nicht mehr zurückgewonnen werden kann.

An dieser Stelle sollte unbedingt erwähnt werden, daß es sich hierbei um ein *Naturgesetz* handelt<sup>13</sup>, das in völliger Allgemeinheit gilt, und nicht um eine technische Aussage. Der hier beschriebene Vorgang ist selbstverständlich von fundamentaler Bedeutung für

<sup>13</sup>Wenn auch um eines von nur statistischer Art; vergleiche dazu Abschnitt 1.2.4.3.

jegliche technische Energieumwandlung, er spielt sich jedoch auch jederzeit in der Natur ab, und zwar immer dann, wenn Wärme über ein Temperaturgefälle von selbst von einem System in ein anderes übertritt und damit unumkehrbare Prozesse ablaufen.

Wir erwähnen aufgrund der häufigen Verwendung bei technischen Anwendungen eine weitere Möglichkeit, die Effizienz einer Wärmekraftmaschine auszudrücken. Zusätzlich zum Wirkungsgrad  $\eta = \Delta W / \Delta Q_{\text{auf}}$ , dem Verhältnis der aus dem wärmeren Reservoir aufgenommenen Wärmemenge und der nutzbaren mechanischen Arbeit, kann man auch die *Leistungszahl*  $\varepsilon_L = \Delta Q_{\text{ab}} / \Delta W$ , also das Verhältnis aus an das kältere Reservoir abgegebene Wärmemenge zur nutzbaren mechanischen Arbeit definieren. Für die Leistungszahl gilt  $\varepsilon_L > 1$ , und der zweite Hauptsatz der Thermodynamik schreibt sich als  $\varepsilon_L \leq 1/\eta_C$ .

### 3.4.4 Beispiele

Zunächst kehren wir zu der in Abschnitt 3.3.1 betrachteten Heizung zurück. Dort war ein Temperaturgefälle von  $T_1 = 333 \text{ K}$  auf  $T_2 = 293 \text{ K}$  gegeben. Dabei berechneten wir eine Entropieproduktion von  $\Delta S = 0,41 \text{ J/K}$  je  $1000 \text{ J}$  übertragener Wärme. Hätte man stattdessen eine ideale Wärmekraftmaschine eingesetzt, so hätte diese die mechanische Arbeit

$$\Delta W = T_2 \Delta S = 293 \text{ K} \cdot 0,41 \frac{\text{J}}{\text{K}} = 120,13 \text{ J}$$

verrichten können. Dieser Energiebetrag verdünnt sich jedoch durch den Vorgang des Heizens im Gewimmel der Luftmoleküle im beheizten Raum und kann von dort nicht mehr zurückgewonnen werden. Im Gegenteil, da die Wände wärmeleitend sind, geht er als Wärme unter zusätzlicher Entropieproduktion weiter nach draußen, da davon ausgegangen werden muß, daß dort die Temperatur tiefer als im Haus ist – sonst würde man nicht heizen.

Als weiteres Beispiel betrachten wir ein Kraftwerk, daß thermische Energie durch eine Turbine in mechanische Energie und diese anschließend durch einen Generator in elektrische Energie umwandelt. Wir nehmen an daß im Heizkessel eine Temperatur von  $900^\circ \text{ C}$  herrscht und im Kühlwasser  $12^\circ \text{ C}$ . Der Wirkungsgrad der Anlage betrage  $39\%$ . Damit berechnen wir den Verlust an verwertbarer Energie. Je  $1000 \text{ J}$  Arbeit, die der vom Kraftwerk bereitgestellte elektrische Strom verrichtet, müssen vom Heizkessel zunächst

$$Q = \frac{W}{\eta} = \frac{1000 \text{ J}}{0,39} = 2564,1 \text{ J}$$

an Wärme geliefert werden; damit gehen  $Q_a = 1564,1 \text{ J}$  Abwärme in das Kühlwasser. Folglich wird die Entropie

$$\Delta S = \frac{Q_a}{T_0} - \frac{Q}{T} = \frac{1564,1 \text{ J}}{285 \text{ K}} - \frac{2564,1 \text{ J}}{1173 \text{ K}} = 3,3 \frac{\text{J}}{\text{K}}$$

erzeugt, so daß je  $1000 \text{ J}$  der Energiebetrag

$$\Delta E = T_0 \Delta S = 285 \text{ K} \cdot 3,3 \frac{\text{J}}{\text{K}} = 940,5 \text{ J}$$

verloren geht. Mit anderen Worten: Ein (theoretisches) ideales Wärmekraftwerk würde  $1940,5 \text{ J}$  anstelle von nur  $1000 \text{ J}$  pro angelieferten  $2564,1 \text{ J}$  Wärme in elektrische Arbeit umwandeln.



Wir kommen nun noch zu einem Beispiel, daß in freier Natur zu finden ist. Beim Ausbruch eines Supervulkans in der La Guerita-Caldera in Colorado ergossen sich vor 27,8 Millionen Jahren 5000 Kubikilometer Lava in die Umgebung. Das sind ungefähr 15 Billionen Tonnen; es war womöglich der größte Vulkanausbruch der Erdgeschichte. Bei einer Umgebungstemperatur von 20° C und einer Lava-Temperatur von typischerweise 1200° C wurden bei deren Abkühlung ungefähr  $1,416 \cdot 10^{21}$  J oder 1,416 Milliarden Terajoule Wärme abgegeben. Damit entstand Entropie vom Betrag

$$\Delta S = \frac{Q}{T_0} - \frac{Q}{T_1} = \frac{1,416 \cdot 10^{21} \text{ J}}{293 \text{ K}} - \frac{1,416 \cdot 10^{21} \text{ J}}{1473 \text{ K}} = 3,871 \cdot 10^{18} \frac{\text{J}}{\text{K}}.$$

Folglich verflüchtigten sich dabei

$$\Delta E = T_0 \Delta S = 293 \text{ K} \cdot 3,871 \cdot 10^{18} \frac{\text{J}}{\text{K}} = 1,134 \cdot 10^{24} \text{ J}$$

an theoretisch nutzbarer Energie unwiederbringlich in die Umgebung.

Die Tatsache, daß sich solche Vorgänge pausenlos in Natur und Technik abspielen, zeigt, daß eine wesentliche Aufgabe dieser beiden jeweils *Entropieerniedrigung* ist. Da der zweite Hauptsatz der Thermodynamik jedoch nicht ausgetrickst werden kann, muß dabei stets woanders *Entropie erzeugt werden*, und zwar so, daß sich die Entropie insgesamt erhöht. Folglich sind alle natürlichen und zivilisatorischen Vorgänge auf das Vorhandensein von Temperaturdifferenzen und auf Energieressourcen angewiesen. Genau das ist mit dem umgangssprachlichen Ausdruck des Energiebedarfs gemeint.

## Aufgaben

1. Vergleiche die jeweilige Entropiezunahme
  - a) beim Schmelzen von 10 kg Eis bei einer Temperatur von 0°C und beim Erhöhen der Temperatur von 10 kg Wasser von 0°C auf 1°C,
  - b) beim Verdampfen von 10 kg Wasser bei 100°C und Normaldruck und bei Erhöhen der Temperatur von 10 kg Wasserdampf von 100°C auf 101°C bei Normaldruck.
  - c) Was für physikalische Folgerungen kann man aus diesen Resultaten ziehen?
2. Ein Wärmekraftwerk soll die Temperaturdifferenz des Meerwassers zwischen Oberfläche und Meeresboden ausnutzen. Berechne den maximalen Wirkungsgrad, wenn die Wassertemperatur oben 25°C und unten 4°C beträgt. Was folgt daraus für die mechanische Arbeit, die ein solches Kraftwerk verrichten kann?
3. Ein Wärmereservoir der Temperatur 1000 K und ein Wärmereservoir der Temperatur 300 K sind über eine Wärmeleitung verbunden. Dabei wird eine Wärmemenge von 1000 J übertragen. Welcher Teil dieser Wärmemenge könnte maximal in mechanische Arbeit umgewandelt werden?
4. Eine Carnotsche Wärmekraftmaschine arbeite mit einem Wirkungsgrad von 0,4 und besitze ein unteres Wärmereservoir der Temperatur 10°C. Um wieviel K muß die Temperatur des oberen Wärmereservoirs erhöht werden, wenn der Wirkungsgrad auf 0,7 gesteigert werden soll?



### 3.5 Der dritte Hauptsatz der Thermodynamik

Es ist nun sehr einfach möglich, über den Begriff des Wirkungsgrades auch zum dritten Hauptsatz zu gelangen, zumindest qualitativ. Das stellt jedoch ebenfalls eine Vertiefung dar, die sicherlich nicht unbedingt erforderlich ist. Sie bietet jedoch eine einfache Möglichkeit, zu sehr tiefschürfenden physikalischen Sachverhalten vorzudringen.

Die Definition des Wirkungsgrades zeigt unmittelbar, wie man die Ausbeute an mechanischer Arbeit bei Wärmekraftmaschinen jeder Art erhöhen kann. Da hierzu der zweite Summand im Wirkungsgrad möglichst klein sein sollte, muß die Temperatur des ersten Wärmereservoirs möglichst hoch und diejenige des zweiten möglichst tief sein, oder anders gesagt, der Wirkungsgrad ist umso größer, je größer die Temperaturdifferenz der beiden Reservoirs ist. Sie zeigt aber auch, daß es hier eine grundsätzliche Grenze gibt. Wenn  $T_2 = 0$  ist, *sich das zweite Wärmereservoir also am absoluten Nullpunkt befindet*, gilt  $\eta = 1$ .

Nun ist der absolute Nullpunkt aber aus quantenmechanischen Gründen nicht exakt erreichbar; die Temperatur eines Systems kann nur asymptotisch gegen den Wert  $T = 0$  gehen. Man kann das den Schülerinnen und Schülern auf dieselbe Weise plausibel machen, wie man auch den absoluten Nullpunkt eingeführt hat: Temperaturabsenkung bedeutet Verminderung der ungeordneten Bewegung der Teilchen eines Systems. Irgendwann hört diese Bewegung ganz auf, und man erreicht eine Temperatur, die nicht weiter abgesenkt werden kann, nämlich den absoluten Nullpunkt<sup>14</sup>. Da die quantenmechanische Natur der Teilchen es aber verhindert, daß sie völlig zur Ruhe kommen<sup>15</sup>, ist dieser absolute Nullpunkt nicht exakt, sondern nur näherungsweise erreichbar.

Hieraus folgt zweierlei. Zunächst ist damit klar, *daß ein Wirkungsgrad von Eins auch theoretisch nicht erreichbar ist*. Dazu bräuchte man zusätzlich zum heißen Wärmereservoir ein zweites am absoluten Nullpunkt, was wie gesehen nicht möglich ist. Außerdem bedeutet es, *daß am absoluten Nullpunkt die Entropie nicht mehr weiter abgesenkt werden kann*. Dazu wäre ein Wärmereservoir von noch tieferer Temperatur notwendig. Das ist jedoch auch aus statistischer Sicht vollkommen klar, denn wenn das ungeordnete Gewimmel der Teilchen eines Systems aufgehört hat, kann dessen inneres Durcheinander nicht weiter abnehmen.

Eine genauere Analyse liefert folgendes Ergebnis, das als Nernstsches Wärmetheorem oder auch dritter Hauptsatz bekannt geworden ist:

*Dritter Hauptsatz der Thermodynamik: Die Entropie physikalischer Systeme nimmt am absoluten Nullpunkt einen konstanten Wert an.*

Wie bereits erwähnt, lassen sich in der Thermodynamik nur Entropiedifferenzen angeben, so daß der konstante Wert der Entropie am absoluten Nullpunkt nicht bestimmt werden kann. Klar ist nur, daß es sich um einen nicht unterschreitbaren Wert handeln muß. Nach einem Vorschlag von Planck ist es üblich, ihn Null zu setzen, ein Wert, der sich im Rahmen der Quantenmechanik auch begründen läßt<sup>16</sup>.

<sup>14</sup>Sehr häufig bringen Schülerinnen und Schüler diesen Begriff bereits in den Unterricht mit.

<sup>15</sup>Hier darf man ohne weiteres auch über Begriffe wie *Nullpunktsenergie* und *Nullpunktschwankungen* reden.

<sup>16</sup>Daran ist erneut erkennbar, was schon mehrfach gesagt wurde, nämlich daß der dritte Hauptsatz der Thermodynamik genau gesehen quantenmechanischer Natur ist. Im Übrigen gilt er nicht in völliger

## Aufgaben

1. Zeige unter Verwendung der Leistungszahl einer als Kältemaschine arbeitenden idealen Wärmekraftmaschine, daß der absolute Nullpunkt in endlich vielen Arbeitsschritten niemals erreicht werden kann.
2. Erläutere den dritten Hauptsatz der Thermodynamik mit Hilfe der statistischen Interpretation der Entropie.

## 3.6 Entropie in Technik und Natur

### 3.6.1 Wärmepumpen

Wärmepumpen sind Maschinen, die unter Aufwenden von Arbeit *Wärme von einem Reservoir mit niedrigerer Temperatur zu einem solchen mit höherer Temperatur pumpen*. Sie stellen damit die Umkehrung der oben diskutierten Wärmekraftmaschinen dar. Der zweite Hauptsatz wird dabei nur scheinbar verletzt, denn zwar bedeutet Wärmeübergang von einem kälteren zu einem heißeren System Entropieabnahme in ersterem, aber gleichzeitig erfolgt eine Entropiezunahme in letzterem, die das wieder ausgleicht; in der Realität wird darüberhinaus die dem zweiten System zugeführte Wärme ganz oder teilweise an die Umgebung abgegeben, so daß auch hier netto Entropie erzeugt wird. Anders ausgedrückt: Eine Wärmepumpe senkt die Entropie in einen Reservoir und erhöht sie dafür im anderen und in der Umgebung.

Wir betrachten die Funktionsweise dieser Technologie am Beispiel der verbreitetsten Variante, den *Kompressions-Wärmepumpen*, die den gewünschten Effekt mit Hilfe eines Kältemittels erreichen (siehe Abbildung 3.3). Das ist eine Substanz, deren Siedepunkt bei

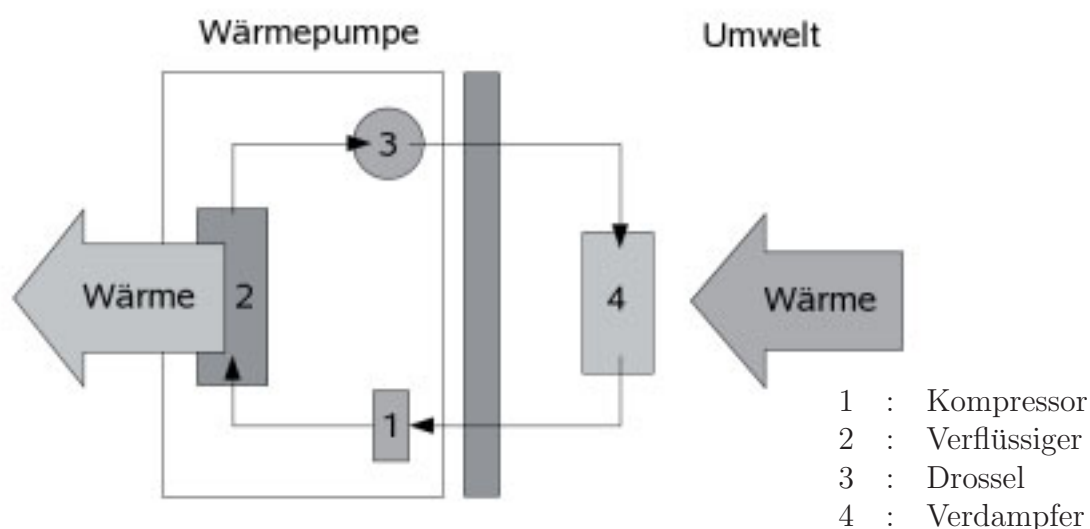


Abbildung 3.3: Schematischer Aufbau einer Kompressions-Wärmepumpe

Allgemeinheit. Genaueres steht in Abschnitt 1.2.4.4.

hohem Druck über der Temperatur des heißeren Reservoirs und bei niedrigem Druck unter der Temperatur des kälteren Reservoirs liegen muß. Das zunächst gasförmige Kältemittel wird in einem Kompressor verdichtet, wobei sich seine Temperatur erhöht. Im Verflüssiger wird Wärme abgegeben, das Kältemittel wird dabei kondensiert und fließt anschließend durch eine Drossel, wobei es expandiert, weiter in den Verdampfer, wo es verdampft und dabei seiner Umgebung Wärme entzieht. Danach beginnt der Prozeß von Neuem, so daß ein Kreislauf stattfindet.

Wärmepumpen haben vielfältige technische Anwendungen. Sie können unter anderem zum Heizen von Räumen verwendet werden. Wir betrachten dazu ein Beispiel. Mit einer Wärmepumpe soll bei einer Grundwassertemperatur von  $10^\circ\text{C}$  das Heizwasser einer Fußbodenheizung auf  $30^\circ\text{C}$  gehalten werden. Ist  $\Delta E_{\text{el}}$  die von der Wärmepumpe pro Zeiteinheit aufgewendete elektrische Energie und  $\Delta Q$  die dadurch aus dem Grundwasser in das Heizwasser transportierte Wärme, dann gilt für die Entropieabnahme des Grundwassers

$$\Delta S_G = \frac{\Delta Q}{T_G}$$

und für die Entropiezunahme der Heizung

$$\Delta S_H = \frac{\Delta E_{\text{el}} + \Delta Q}{T_H}.$$

Für die gesamte Entropiezunahme gilt daher und aufgrund des zweiten Hauptsatzes

$$\Delta S = \frac{\Delta E_{\text{el}} + \Delta Q}{T_H} - \frac{\Delta Q}{T_G} \geq 0.$$

Das kann man nach  $\Delta Q$  auflösen und erhält

$$\Delta Q \leq \Delta E_{\text{el}} \frac{T_G}{T_H - T_G},$$

mit obigen Temperaturwerten also

$$\Delta Q \leq 14,2 \Delta E_{\text{el}}.$$

Man erhält also bei diesem Beispiel im Idealfall 14,2 mal soviel thermische Energie, wie man erreicht hätte, wenn man die gleiche elektrische Energie im Hausinnern direkt in thermische Energie umgewandelt hätte, etwa durch eine gewöhnliche Elektroheizung. Das ist natürlich kein realistischer Wert, aber das dreifache ist mit Wärmepumpen durchaus zu holen<sup>17</sup>.

<sup>17</sup>Es ist üblich, die Leistung von Wärmepumpen durch eine Leistungszahl anzugeben, dem sogenannten *Coefficient of Performance*, abgekürzt COP. Dieser ist definiert als Quotient aus abgegebener Wärmemenge und aufgewendeter elektrischer Energie,

$$\text{COP} = \frac{Q}{E_{\text{el}}}.$$

Er ist wie jede Leistungszahl und wie obige Rechnung zeigt, durch den zweiten Hauptsatz beschränkt gemäß

$$\text{COP} \leq \frac{1}{\eta_c}.$$

Der COP wird gelegentlich als „Wirkungsgrad“ bezeichnet, was jedoch irreführend ist, da es sich dabei formal gerade um den Kehrwert des thermodynamischen Wirkungsgrades handelt.

Kühlschränke sind eine weitere verbreitete Anwendung dieser Technologie. Ein Kühlschrank ist eine Wärmepumpe, die dem Kühlschrankinneren Wärme entzieht und diese an der Gehäuserückseite an die Umgebung abgibt. Er senkt somit die Entropie in seinem Innern ab und erhöht sie dafür in der Umgebung um so mehr. Die Funktionsweise von Kühlschränken läßt sich durch die folgende bei mündlichen Prüfungen beliebte Frage veranschaulichen: Ist es sinnvoll, zur Kühlung eines Raumes die Kühlschranktür offen stehen zu lassen? Die Antwort lautet natürlich „nein“, denn der Kühlschrank transportiert Wärme nur aus seinem Inneren in den Raum und nicht etwa aus diesem hinaus. Außerdem machen sich der elektrische Widerstand der Wärmepumpe sowie mechanische Reibung ihrer beweglichen Teile als zusätzliche Erwärmung der Kühlschrankrückseite bemerkbar. Es wird in dem Raum deshalb keineswegs kälter, sondern im Gegenteil sogar deutlich wärmer.

### 3.6.2 Entropiebilanzen der Natur

Schon bei ihren ersten Begegnungen mit der Thermodynamik kommen die Schülerinnen und Schüler selbst darauf, daß die Vorstellung eines abgeschlossenen Systems eine Idealisierung und das einzig wirklich abgeschlossene System das Universum ist. In diesem Sinne formulierte Clausius den ersten und den zweiten Hauptsatz wie folgt: *Die Energie des Universums ist konstant. Die Entropie strebt einem Maximum zu.*

Die gesamte Biosphäre mit ihren vielfältigen Lebewesen scheint dem zweiten Hauptsatz zu widersprechen. Es finden hier laufend Prozesse statt, die Entropie erzeugen, dennoch versinken biologische Systeme nicht im Chaos und kochen auch nicht über. Sie scheinen irgendwie die Entropie absenken zu können. Da das aber global nicht geht, bleibt ihnen nichts anderes übrig, als die überschüssige Entropie an die Umgebung abzugeben. Lebewesen sind entropieexportierende Systeme und damit auf Wechselwirkung mit ihrer Umgebung angewiesen. Das ist am allgemeinen Ressourcenverbrauch wahrnehmbar und wird durch die pausenlose Umwandlung irgendwelcher Energieformen in thermische Energie verursacht.

Bei Menschen und Tieren läuft dieser Prozeß im wesentlichen über die Umwandlung von durch die Nahrung bereitgestellter chemischer Energie in thermische und mechanische Energie. Die Tatsache, daß es nicht nur mechanische Energie ist, führt zur Erhöhung der Entropie innerhalb der Organismen. Diese fast ausschließlich neu entstandene Entropie wird an die Umgebung abgegeben, was sich in der Abgabe von Wärme, aber auch in der Produktion von Abfällen und sonstiger Unordnung bemerkbar macht. Bei Pflanzen ist es etwas komplizierter. Sie nehmen Lichtenergie auf und wandeln diese zunächst durch Glucoseproduktion per Photosynthese in chemisch gespeicherte Energie um. Glucosemoleküle haben jedoch aufgrund ihrer im Vergleich zu den Ausgangsstoffen komplexeren Struktur eine geringere Entropie, und da Entropie nicht abnehmen kann, muß sie auch hier an die Umgebung abgegeben werden. Die Abgabe der überschüssigen Entropie bewerkstelligen die grünen Pflanzen durch Aufnahme von entropiearmem flüssigem Wasser und Abgabe desselben im entropiereichen gasförmigen Zustand an die Atmosphäre.

Auch die Erde als ganzes ist ein entropieexportierendes System. Sie nimmt von der Sonne entropiearmes, eher kurzwelliges, insbesondere sichtbares Licht auf und gibt entropiereiches infrarotes Licht an den Weltraum ab. Dieser Prozeß wird umso stärker behindert, je mehr Kohlendioxid in der Atmosphäre enthalten ist, da infrarotes Licht von

diesem absorbiert wird. Das ist die Ursache des Treibhauseffektes, denn wenn die auf der Erdoberfläche erzeugte Entropie nicht in ausreichendem Maße abgegeben werden kann, dann steigt sie in der Atmosphäre an, und das bedeutet ein Ansteigen der Temperatur. Das einzige wirksame Gegenmittel ist auf Dauer ein Zurückfahren unseres aufwendigen Lebensstils auf ein vernünftigeres Niveau.

Mit diesen etwas pathetischen Worten beschließen wir die Vorstellung unseres Entropie-Kurses. Wie angekündigt bleiben wir dabei stets auf der traditionellen Ebene der Energieformen und Energieumwandlungen, außerdem wird der Begriff der Entropie aus der allgemeineren, statistischen Perspektive eingeführt und die thermodynamische Perspektive erst anschließend nachgeliefert. Damit werden sämtliche Probleme, die der Karlsruher Physikkurs mit sich bringt, von Beginn an vermieden.

## Aufgaben

- 1) Betrachte ein Auto der Masse 1 t und der Geschwindigkeit 10 m/s, das bei einer Umgebungstemperatur von 300 K vollständig abgebremst wird, und einen thermodynamischen Prozeß, bei dem eine Wärmemenge von 40 kJ von einem Reservoir der Temperatur 400 K auf eines mit 300 K über eine Wärmeleitung zwischen den beiden Reservoirs übertragen wird.
  - a) Vergleiche die dabei jeweils als nutzbare mechanische Arbeit verlorengelassene Energie.
  - b) Was kann man über die jeweilige Entropiezunahme sagen?
- 2) Eine Wärmepumpe arbeitet mit der Umgebungstemperatur von 0°C und erwärmt unter Verwendung eines Diesellagers mit dem Wirkungsgrad 0,4 das Heizwasser permanent auf 40°C. Um das wievielfache ist die Wärmepumpe effektiver als eine direkte Dieselheizung mit einem konventionellen Brenner?
- 3) Wir betrachten eine Kühlanlage, deren Aufgabe es ist, bei einer Umgebungstemperatur von  $T_u$  die Temperatur in einem Kühlraum auf dem Wert  $T_k$  zu halten.
  - a) Warum ist das mit einer Entropieabnahme im Kühlraum verbunden?
  - b) Leite einen Ausdruck für die Arbeit her, welche die Kühlanlage verrichten muß.
  - c) Warum sollte man nie tiefer kühlen als unbedingt erforderlich?
  - d) Leite eine Beziehung zwischen der Leistungsziffer einer idealen Kältemaschine und dem Carnotschen Wirkungsgrad her.

# Kapitel 4

## Quantenmechanik in der Oberstufe

Als zweite Unterrichtseinheit stellen wir eine Einführung in die Quantenmechanik für die gymnasiale Oberstufe vor. Anders als die üblichen Konzepte stellt unser Kurs den Welle-Teilchen-Dualismus in den Mittelpunkt und setzt sich nicht nur, aber insbesondere zum Ziel, den grundlegenden Unterschied zwischen klassischem und quantenmechanischem Verhalten besonders deutlich zu machen.

### 4.1 Vorbemerkungen

Der hier dargestellte Lehrgang orientiert sich an zwei didaktischen Grundüberzeugungen:

1. Es werden nur in der Physik wohletablierte Begrifflichkeiten verwendet und jegliche „Eigenzüchtungen“ der Schulphysik konsequent vermieden.
2. Die Quantenmechanik wird axiomatisch aufgezo-

gen. Während die erste weiter oben bereits ausführlich diskutiert wurde, sollte die zweite näher begründet werden. Im einzelnen gibt es drei gewichtige Argumente für eine solche Vorgehensweise.

- Es ist eine der wesentlichen Eigenschaften der Quantenmechanik, etwas grundlegend neues, nicht aus der klassischen Physik ableitbares darzustellen. Diesem Sachverhalt wird in der Schule genauso wie in der harten wissenschaftlichen Realität nur eine axiomatische Grundlegung und ein darauf basierender deduktiver Aufbau gerecht. Irgendwelche induktiven Zugänge, Plausibelmachen mit Hilfe klassischer Vorstellungen oder sonstige angebliche Herleitungen, wo es strenggenommen keine gibt, führen bei den Schülerinnen und Schülern zwangsläufig zu Fehlvorstellungen, die ihnen schwer wieder abzugewöhnen sind. Hierzu gehören insbesondere solche Dinge wie die „Herleitung“ der Schrödingergleichung aus dem Energiesatz, der dreidimensionale Potentialtopf als Behelfsmethode zur Berechnung der Energieniveaus des Wasserstoffatoms, oder, am schlimmsten von allen, das Bohrsche Atommodell. Die Schrödingergleichung kann man nicht herleiten, sie muß postuliert werden, andernfalls hantiert man hier genau wie beim Potentialtopf-„Modell“ der Schulbuchliteratur mit potentieller und kinetischer Energie von Systemen herum, die sich in Zuständen befinden, in denen sie so etwas überhaupt nicht besitzen. Und



der Schaden, den die Elektronenbahnen des Bohrschen Atommodells anrichten, ist wissenschaftlich dokumentiert und kaum zu reparieren.

- Eine grundlegende Erfahrung, an der keiner vorbeikommt, der sich in diese Thematik einarbeitet, ist die Erkenntnis, daß man Quantenmechanik nicht verstehen, sondern *nur akzeptieren* kann – sofern man mit „Verstehen“ das sich Vorstellen eines Sachverhalts in klassischen Bildern meint. Die Notwendigkeit, sich an etwas völlig neues zu gewöhnen, kann nur durch *vom Himmel fallende physikalische Grundgesetze* deutlich genug hervortreten.
- Das *Postulieren von physikalischen Prinzipien* ist exemplarisch für die Vorgehensweise der theoretischen und mathematischen Physik. Weitere Beispiele neben der Quantenmechanik sind alle Eichtheorien, bei denen das Eichprinzip selbst sowie die jeweils verwendeten Lagrange-Dichten axiomatisch vorausgesetzt werden. Abgesehen davon, daß sich diese Verfahrensweise hervorragend bewährt, was für sich allein schon ein überzeugendes Argument zu ihrer Rechtfertigung wäre, liegt ein wesentlicher Grund dafür, axiomatisch vorzugehen, offensichtlich in der Natur selbst. Diese scheint uns bei der Konstruktion physikalischer, und damit in mathematischer Sprache abgefaßter Theorien zu zwingen, *ganz spezielle mathematische Strukturen zu verwenden* und keineswegs etwa irgendwelche völlig willkürlichen. Physikalische Theorien legen darüberhinaus sehr häufig ein ganz erstaunliches Eigenleben an den Tag in dem Sinne, daß Strukturen, die zunächst scheinbar aus rein mathematischen Gründen und teilweise sogar ungewollt oder unwillkommen in ihnen auftauchen, später in neuentdeckten physikalischen Objekten oder Phänomenen wiedergefunden werden<sup>1</sup>. Die Kunst scheint darin zu liegen, diese Strukturen und die ihnen vorgeordneten Axiome zu *erraten*.

Damit favorisieren wir für die Quantenmechanik einen völlig anderen Zugang als sonstige Bereiche der Physik, doch das läßt sich selbst wieder zu einer Grundregel der Physikdidaktik erheben. Generell sollte nämlich folgendes Prinzip eingehalten werden: In dem Maße, wie die Einführung klassisch-physikalischer Sachverhalte operational erfolgen sollte, ist bei der Einführung quantenmechanischer Sachverhalte axiomatisches Vorgehen angebracht.

Der Aufbau des Kurses orientiert sich im wesentlichen, aber nicht vollständig, an der üblichen Reihenfolge in den einschlägigen Schulbüchern. Der hauptsächliche Unterschied betrifft die zeitliche Abfolge der typischen einführenden Schulversuche in Abschnitt 4.5.2. Während üblicherweise zuerst der Photoeffekt und anschließend die Elektronenbeugungsröhre behandelt wird, wählen wir, einem Vorschlag von H. Fischler folgend [51], [170], [172], die umgekehrte Reihenfolge. Das rechtfertigt sich dadurch, daß die Vorstellung von Photonen als teilchenähnliche Konstituenten elektromagnetischer Felder für die meisten Schülerinnen und Schüler nichts Neues darstellt; die Beobachtung wellenähnlichen Verhaltens bei etwas, das man sich bisher nur als Teilchen, mit anderen Worten als winzige Kügelchen vorgestellt hat, zieht einen viel größeren Überraschungseffekt nach sich und leitet genau das wesentliche Lernziel jeder Unterrichtseinheit über Quantenmechanik ein,

---

<sup>1</sup>Prominente Beispiele hierfür sind der Maxwellsche Verschiebungsstrom oder die kosmologische Konstante.



nämlich festgefügte physikalische Weltbilder zum Einsturz zu bringen. Das bewährt sich seit Jahren bestens.

Entgegen der immer wieder zu hörenden Forderung, der Welle-Teilchen-Dualismus müsse überwunden werden, was immer das auch heißen soll, setzen wir denselben hier in Abschnitt 4.5.3 an den Anfang und bringen ihn damit in die Position eines nicht weiter hinterfragbaren Grundprinzips der Quantenmechanik. Denn seit den 90er Jahren weiß man, daß es sich dabei genau um ein solches handelt, und nicht etwa, wie man zuvor jahrzehntelang dachte, um eine Konsequenz der Unschärferelation. Die besonderen nichtklassischen Eigenschaften von Quantenobjekten lassen sich in keiner anderen Weise vergleichbar eindrucksvoll und gleichzeitig elementar vermitteln. Entgegen meiner sonstigen Überzeugung, daß Computersimulationen im Physikunterricht nichts zu suchen haben, plädiere ich hier für den Einsatz eines Computerprogramms zur Simulation von Einzelphotonen- und Einzel-Elektronen-Interferenzexperimenten.

In Abschnitt 4.5.4 erfolgt eine stark elementarisierte Betrachtung der Axiome der Quantenmechanik in der von Neumannschen Hilbertraum-Formulierung, ohne dabei natürlich irgendwelche funktionalanalytischen Begriffe an irgendeiner Stelle zu verwenden. Gelegentlich findet man Versuche, Operatoren in den gymnasialen Physikunterricht einzuführen; ich halte das für völlig überzogen. Dennoch sollte man die Funktionalanalysis unausgesprochen im Blick behalten. Das hat den Vorteil, daß man mit einem stabilen, physikalisch hieb und stichfesten begrifflichen Fundament arbeitet. Alle anderen Axiomatiken der Quantenmechanik, insbesondere auch tiefergehende, bei welchen die Hilbertraumstruktur selbst mit abgeleitet werden soll, sind für die Schule unbrauchbar.

Abschnitt 4.5.5 widmet sich der Unschärferelation. Hier wird leider in der Schulbuchliteratur oft gehudelt, was reihenweise zu unkorrekten Formulierungen führt. Dabei bleibt insbesondere häufig unklar, was unter dem Begriff „Unschärfe“ (oder „Unbestimmtheit“, wie gelegentlich dazu gesagt wird) überhaupt zu verstehen ist, was unweigerlich zu Fehlinterpretationen führt. Deshalb kommt man um eine präzise Definition der quantenmechanischen Unschärfe einer physikalischen Größe nicht herum; nur dann ist eine präzise Formulierung der Unschärferelation möglich.

Wie jede physikalische Theorie sollte sich auch die Quantenmechanik nicht zuletzt mit der Beschreibung physikalischer Systeme beschäftigen. Laut Bildungsplan könnte man sich hier zwar auf eine qualitative Beschäftigung beschränken; dann müßte man sich aber fragen, wozu der ganze bisherige (und im Bildungsplan auch geforderte) Aufwand eigentlich gut sein soll. Daher nehmen wir uns in Abschnitt 4.5.6 zwei einfache Systeme vor, aber ohne behelfsmäßige Plausibilitätsbetrachtungen, sondern ganz konkret mit Hilfe der entsprechenden Schrödingergleichungen und deren Lösungen. Selbstverständlich muß man sich hier auf stationäre Probleme beschränken. Der lineare Potentialtopf mit unendlich hohen Wänden bietet sich als ideales Beispiel an, um eine Eigenwertaufgabe komplett mit allen Schikanen zu lösen und dabei die Schülerinnen und Schüler die allgemeine Vorgehensweise exemplarisch und mit schulischen Mitteln kennenlernen zu lassen. Das im Bildungsplan geforderte zeitgemäße Atommodell vermittelt man am besten, indem man die Schrödingergleichung des Wasserstoffatoms zwar nicht löst, was in der Oberstufe natürlich auch nicht zu leisten wäre, aber vorgegebene Wellenfunktionen auf ihre Eigenschaft, Lösungen der stationären Schrödingergleichung zu sein, überprüfen läßt. Dabei bekommt man die Rydbergformel gewissermaßen als Nebenprodukt geschenkt. In diesem Abschnitt braucht man an zwei Stellen partielle Integration, was

bekanntlich im Fach Mathematik nicht mehr im Bildungsplan enthalten ist. Wer die (sehr einfache) Einführung im Physikunterricht vermeiden möchte, sollte die beiden Integrale direkt mitteilen.

Etwas aus dem Rahmen fällt Abschnitt 4.5.7, sowohl was den Umfang, als auch Inhalt und Schwierigkeitsgrad betrifft. Er enthält einen Vorschlag für einen in der gymnasialen Oberstufe gangbaren Weg zur Einführung verschränkter Systeme und berücksichtigt dabei auch das EPR-Paradoxon, das Bellsche Theorem und die daraus folgende experimentelle Bestätigung der Quantenmechanik gegenüber beliebigen lokalen Theorien mit verborgenen Variablen. Das Thema Verschränkung ist kein Pflichtthema, kommt folglich auch im schriftlichen Abitur nicht vor, und man kann zweifellos darüber streiten, ob man es im Gymnasium überhaupt bringen sollte. Ich habe es hier dennoch aufgenommen und noch dazu sehr ausführlich, und zwar aus folgenden Gründen. Erstens findet man dazu sowohl in Schul- wie auch in Hochschulbüchern viel Unsinn; darüberhinaus hat die Thematik Einzug in unterschiedlichste esoterische Diskussionen gehalten, was zu noch abwegigeren Behauptungen und Spekulationen führt. Es dürfte sinnvoll sein, zu versuchen, hier der Orientierungslosigkeit und den verbreiteten Fehlinformationen etwas entgegenzuwirken. Zweitens und folglich kann der Abschnitt zunächst interessierten Kolleginnen und Kollegen als Einstieg in diese Thematik dienen. Ob man es dann auch unterrichtet, kann jede/jeder selbst entscheiden. Der Schwierigkeitsgrad dürfte allerdings typische Oberstufenkurse nicht selten überfordern. Ich halte es dennoch für sinnvoll, solche Sachverhalte zumindest zu erwähnen, denn drittens wird erst im Zusammenhang mit verschränkten Systemen so richtig klar, wie groß der Unterschied zwischen klassischer Physik und Quantenmechanik tatsächlich ist und warum wir überhaupt so etwas wie eine klassische Welt um uns herum wahrnehmen. Die Wichtigkeit dieses Themas kann somit gar nicht überschätzt werden.

Wieder ist der Umfang der beschriebenen Unterrichtseinheit sehr groß; das relativiert sich jedoch schon durch das soeben zu ihrem letzten Abschnitt gesagte. Darüberhinaus ist die Einheit zunächst sowohl für Grund- als auch für Leistungskurse konzipiert. Erstere verfügen etwa in Baden-Württemberg über zwei, letztere über vier Wochenstunden, woraus sich von vorneherein eine unterschiedlich ausführliche Durchführung ergibt. Grundsätzlich sollten die Abschnitte 4.5.2, 4.5.3, 4.5.5 und 4.5.6 stets unterrichtet werden; Abschnitt 4.5.4 kann in Grundkursen überschlagen werden, da man sich dort auf die Erwähnung der Unschärferelation beschränken kann. Ansonsten sollte der Unterschied zwischen Grund- und Leistungskursen hauptsächlich im Umfang der betrachteten Anwendungen und Aufgaben liegen.

Bevor wir den Unterrichtsgang nun im Detail beschreiben, sind noch einige Anmerkungen zu den beiden Fundamenten angebracht, auf denen dabei aufgebaut wird, das heißt also zur Axiomatik der Quantenmechanik und zum Prinzip des Welle-Teilchen-Dualismus. Das erfolgt einerseits, weil dazu in der hier benötigten Form in einschlägigen Lehrbüchern wenig bis gar nichts nachzulesen ist, und andererseits, um das Einhalten der oben stets ausgesprochenen Forderung, den Unterricht stets im Einklang mit der etablierten Physik zu halten, dennoch zu ermöglichen, indem die hierfür erforderlichen Grundlagen bereitgestellt werden.

## 4.2 Axiomatische Basis der Quantenmechanik

Wir betrachten hier das Axiomensystem, auf dem die Hilbertraum-Formulierung der Quantenmechanik aufgebaut ist. Es stammt ursprünglich von J. von Neumann, der damit gleichzeitig die Grundlage zur modernen, mathematisch präzisen Form der Quantenmechanik lieferte<sup>2</sup>. Es herrscht keine Einigkeit darüber, ob dieses Axiomensystem wirklich das grundlegendste ist; es gibt Vorschläge für tieferliegende Systeme, aus welchen die hier beschriebenen Axiome wiederum ableitbar sind<sup>3</sup>. Darüberhinaus existieren weitere, alternative quantenmechanische Axiomensysteme<sup>4</sup>. Auf einer sehr viel allgemeineren Ebene steht die axiomatische Quantenfeldtheorie, die in Anmerkung 50 auf Seite 115 erwähnt wurde. Sie stellt nach wie vor eher ein Forschungsprojekt als einen ausformulierten Theorienrahmen dar.

### 4.2.1 Die Axiome der Quantenmechanik

Es hat sich gezeigt, daß die nichtrelativistische Quantenmechanik der Einteilchen-Systeme im wesentlichen auf sechs Axiomen aufgebaut werden kann. Auf diese sechs Axiome werden wir uns hier beschränken. Betrachtet man auch Mehrteilchen-Systeme, so kommen weitere Axiome dazu, etwa was Bosonen und Fermionen und die Symmetrieeigenschaften der sie beschreibenden Zustände angeht. Die relativistische Quantenmechanik und die Quantenfeldtheorie erfordern gleichfalls weitere und auch abgewandelte Axiome.

Wir verwenden die formale Sprache der mathematischen Physik, wie sie auch in Kapitel 2 zum Einsatz kam, abgesehen davon, daß das  $\hbar$  hier wiedereingeführt wird. Natürlich ist das weit davon entfernt, im Physikunterricht einsetzbar zu sein; wie man die hier beschriebenen Axiome in eine Form bringt, die in der Schule verwendbar ist, steht in Abschnitt 4.5.1. Wir listen die Axiome zunächst nur auf, um sie im nächsten Abschnitt etwas zu kommentieren.

*Axiom 1: Die möglichen Zustandsvektoren eines quantenmechanischen Systems sind genau die Elemente eines unendlichdimensionalen komplexen Hilbertraumes  $\mathcal{H}$ .*

*Axiom 2: Jeder meßbaren physikalischen Größe eines quantenmechanischen Systems entspricht ein selbstadjungierter Operator auf  $\mathcal{H}$ .*

*Axiom 3: Einzig mögliche Meßwerte einer physikalischen Größe sind die Elemente des Spektrums  $\sigma(\hat{A})$  des zugehörigen selbstadjungierten Operators  $\hat{A}$ .*

*Axiom 4: Befindet sich ein quantenmechanisches System im normierten Zustand  $|\psi\rangle \in \mathcal{H}$  und wird die zum selbstadjungierten Operator  $\hat{A}$  gehörende physikalische Größe gemessen, so ist die Wahrscheinlichkeit, einen Meßwert aus der Menge  $\mathcal{M} \subseteq \sigma(\hat{A})$  zu*

<sup>2</sup>Von Neumanns berühmtes Buch [392] beschreibt auf knapp 270 Seiten in auch heute noch beeindruckender Knappheit und Präzision den Stand der mathematischen Quantenmechanik der frühen dreißiger Jahre.

<sup>3</sup>Hier wäre insbesondere die Genfer Schule zu nennen mit J. M. Jauch und C. Piron als ihren wichtigsten Vertretern [275], [408]. Diese Axiomatik startet mit Booleschen Verbänden und leitet die Hilbertraumstruktur daraus ab. Vergleiche auch [355] und [396].

<sup>4</sup>Beispiele sind die Vorschläge von L. Hardy [212] - [215] und von G. Ludwig [351], [352], [353].

finden,

$$\mathfrak{P}(\mathcal{M}) = \int_M d\langle \hat{E}_\lambda \psi | \psi \rangle. \quad (4.1)$$

Dabei ist  $\{\hat{E}_\lambda\}_{-\infty \leq \lambda \leq \infty}$  die Spektralschar von  $\hat{A}$ .

**Axiom 5:** Falls eine Messung der zum Operator  $\hat{A}$  gehörenden physikalischen Größe am durch den Zustand  $|\psi\rangle$  beschriebenen quantenmechanischen System den Meßwert  $\lambda \in \sigma(\hat{A})$  ergeben hat, befindet sich das System unmittelbar nach der Messung im Zustand

$$|\psi_\lambda\rangle := \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon > 0}} \frac{(\hat{E}_{\lambda+\varepsilon} - \hat{E}_{\lambda-\varepsilon}) |\psi\rangle}{\sqrt{\langle \psi | \hat{E}_{\lambda+\varepsilon} - \hat{E}_{\lambda-\varepsilon} | \psi \rangle}}.$$

**Axiom 6:** Zu jedem quantenmechanischen System gehört ein eindeutig bestimmter selbstadjungierter Operator  $\hat{H}$ , der die zeitliche Entwicklung des Systems beschreibt. Befindet sich das System zum Zeitpunkt  $t_0$  im normierten Zustand  $|\psi(t_0)\rangle$  so ist es zum Zeitpunkt  $t \geq t_0$  im Zustand

$$|\psi(t)\rangle = \hat{U}(t, t_0) |\psi(t_0)\rangle.$$

Dabei ist  $\hat{U}(t, t_0)$  der Zeitentwicklungsoperator; er ist Lösung der Operator-Differentialgleichung

$$i\hbar \frac{\partial \hat{U}(t, t_0)}{\partial t} = \hat{H}(t) \hat{U}(t, t_0)$$

mit der Randbedingung  $\hat{U}(t_0, t_0) = \mathbf{1}$ .

## 4.2.2 Bemerkungen zu den Axiomen

Die nun folgenden Interpretationen der sechs Axiome erheben nicht den Anspruch, die einzig möglichen und uneingeschränkt akzeptierten zu sein. Sobald man in der Quantenmechanik den formal mathematischen Aspekt verläßt, herrscht im allgemeinen bis auf eine gewisse minimale Schnittmenge kein Konsens mehr unter den Fachleuten. Wir vertreten hier wie im gesamten Buch eine Sichtweise, die einen sehr weitgehenden Theorienrealismus beinhaltet und stellenweise deutlich über die statistische Minimalinterpretation hinausgeht.

### 4.2.2.1 Zustände und Zustandsvektoren

Wir können die physikalische Bedeutung des ersten Axioms verdeutlichen, indem wir den folgenden Zusatz anfügen: *Der Zustand eines quantenmechanischen Systems ist durch die Angabe des zugehörigen Zustandsvektors vollständig beschrieben*<sup>5</sup>. Das heißt, alles was man über ein quantenmechanisches System überhaupt wissen kann, steckt in dessen Zustandsvektor. Es ist physikalisch unmöglich, mehr Information über ein System

<sup>5</sup>Wir beschränken uns hier auf die Beschreibung reiner Zustände und lassen die Erweiterung auf reine und gemischte Zustände und die Verwendung statistischer Operatoren als Repräsentanten quantenmechanischer Zustände außer Acht, da dies für die Schulphysik völlig ohne Belang ist.

zu besitzen, als in dessen Zustandsvektor enthalten ist. Insbesondere hat diese Unmöglichkeitsaussage nichts mit irgendwelchem technischem Unvermögen oder dergleichen zu tun, sie bringt vielmehr eine grundlegende Eigenschaft der Natur zum Ausdruck.

Da sich quantenmechanische Systeme in der Zeit entwickeln können, sind die ihren Zustand beschreibenden Zustandsvektoren zeitabhängig<sup>6</sup>:

$$|\psi\rangle = |\psi(t)\rangle.$$

Wir bleiben hierbei zunächst bei der oben verwendeten darstellungsfreien Formulierung, wählen also keine spezielle Basis. Diese abstrakte Formulierung beinhaltet bereits sämtliche Strukturen und damit sämtliche Informationen über das betrachtete physikalische System. Für konkrete Betrachtungen wird man dennoch auf konkrete Darstellungen und damit auf die Wahl spezieller Basissysteme zurückgreifen. Das bedeutet, daß man die Zustandsfunktionen nach einem vollständigen Orthonormalsystem des Hilbertraums entwickelt. Hier kommen insbesondere eigentliche oder verallgemeinerte Eigensysteme selbstadjungierter Operatoren zum Einsatz.

Die zentrale Aussage des ersten Axioms der Quantenmechanik ist die Hilbertraumstruktur der Gesamtheit der möglichen Zustandsvektoren eines quantenmechanischen Systems. Dies beinhaltet insbesondere auch, daß wenn irgendwelche Wellenfunktionen  $|\psi_i\rangle$  in diesem Hilbertraum  $\mathcal{H}$  liegen, damit auch beliebige Linearkombinationen  $\sum_i |\psi_i\rangle$

Elemente von  $\mathcal{H}$  sind. Mit jedem möglichen Zustand eines Systems sind somit auch beliebige Superpositionen solcher Zustände wieder mögliche Zustände des Systems. Was aber sollen Superpositionen oder Überlagerungen von Zuständen eigentlich überhaupt sein?

Auf den ersten Blick könnte man meinen, daß es sich hier um etwas altbekanntes handelt. An das Superpositionsprinzip hat man sich aus der klassischen Wellentheorie gewöhnt, ganz egal, ob es sich dabei um mechanische Wellen, Schallwellen, elektromagnetische Wellen oder was auch immer handelt. Auch hier bilden die Gesamtheit der jeweiligen Lösungen der zugehörigen Wellengleichungen jeweils Vektorräume. In diesem Kontext bedeutet das Superpositionsprinzip, daß sich Wellen ungestört durchdringen und am Ort ihres Zusammentreffens überlagern, ihre Amplituden sich also addieren. Dies hat die wohlbekannten *Interferenzerscheinungen* zur Folge, mit deren Hilfe sich Wellen normalerweise zweifelsfrei als solche identifizieren lassen.

Mathematisch gesehen handelt es sich beim quantenmechanischen Superpositionsprinzip um genau dasselbe. Aus physikalischer Sicht ist hier *alles vollkommen anders*. Denn die Größen, die in der Quantenmechanik überlagert werden, sind von ganz anderer Art als diejenigen in den klassischen Wellentheorien. Dort handelt es sich üblicherweise um Feldstärken, also um *reelle Größen*, die (mehr oder weniger) direkt und vor allen Dingen zumindest prinzipiell stets beliebig genau meßbar sind und die sich direkt interpretieren lassen, auch was deren Überlagerung betrifft. In der Quantenmechanik dagegen werden *Zustandsvektoren* überlagert, also *Elemente eines komplexen Hilbertraums*, die insbesondere keine Feldstärken irgendwelcher Art repräsentieren, sondern Zustände des Systems. Wir haben es hier folglich mit *Superpositionen unterschiedlicher Zustände* des betrachteten physikalischen Systems zu tun, die noch dazu *interferenzfähig* sind. Hierin

---

<sup>6</sup>Wir befinden uns hier im Schrödingerbild. Auf die Diskussion der Möglichkeit, die Quantenmechanik in unterschiedlichen Bildern wie Schrödinger-, Heisenberg- oder Wechselwirkungsbild darzustellen, können wir hier ebenfalls verzichten.



liegen die Ursachen für all die Merkwürdigkeiten, für die die Quantenmechanik berühmt und berüchtigt geworden ist. Auch wenn wir uns hier wiederholen: *Der* wesentliche Unterschied zwischen der klassischen Physik und der Quantenmechanik liegt in der *Existenz interferenzfähiger Superpositionen* in der Quantenmechanik.

Einige zusätzliche Bemerkungen zum Begriff des Raums der Zustände seien angefügt. Die Klasse der unendlichdimensionalen komplexen Hilberträume bildet ein sehr weites Feld unterschiedlicher Räume. Häufig kann man sich jedoch auf eine Teilklasse beschränken, nämlich auf die unendlichdimensionalen komplexen *separablen* Hilberträume. Dies ist sehr häufig der Fall; Ausnahmefälle, das heißt *nichtseparable* Hilberträume, gibt es in der statistischen Mechanik und der Quantenfeldtheorie, nicht jedoch in der Einteilchen-Quantenmechanik. Da alle separablen unendlichdimensionalen Hilberträume isomorph und sogar isometrisch zum Raum  $\mathcal{L}^2$  der quadratisch Lebesgue-integrierbaren Funktionen sind, kann letzterer üblicherweise als schön anschaulicher Representant für quantenmechanische Zustandsräume verwendet werden. Das ist insbesondere im allgemeinen dann der Fall, wenn von Wellenfunktionen die Rede ist. Die Separabilität garantiert dabei stets auch die Existenz abzählbarer vollständiger Orthonormalsysteme.

#### 4.2.2.2 Observablen und Operatoren

Das zweite Axiom gibt darüber Auskunft, in welcher Form meßbare Größen in der Quantenmechanik repräsentiert werden. Warum dafür gerade selbstadjungierte lineare Operatoren verwendet werden, wird gleich bei der Diskussion des dritten Axioms klar werden. Was jedoch nicht klar wird, ist die Antwort auf die Frage, *was denn nun genau die beobachtbaren oder meßbaren physikalische Größen sind*. Hier befinden wir uns bereits mitten in unsicherem Territorium.

Die naheliegendste Möglichkeit wäre es, die Aussage des zweiten Axioms einfach umzudrehen. Das führt zu der Behauptung, daß jedem selbstadjungierten Operator auf dem betrachteten Zustandsraum eine physikalische Größe entspricht, die an dem zugehörigen quantenmechanischen System zumindest im Prinzip gemessen werden kann. Diese Behauptung muß aus zweierlei Perspektive näher geprüft werden. Einerseits stellt sich die Frage, ob die Eigenschaft, selbstadjungiert zu sein, für einen Operator bereits ausreicht, um eine meßbare Größe zu repräsentieren. Andererseits muß auch hinterfragt werden ob tatsächlich jeder geeignete selbstadjungierte Operator (was immer das heißt) einer Größe entspricht, die auch wirklich gemessen werden *kann*. Beide Fragestellungen gehen weit über die Axiomatik der Quantenmechanik hinaus. Wir werden sie uns dennoch im Zusammenhang mit dem vierten Axiom erneut vornehmen.

Eine Besonderheit, die durch die Verwendung unendlichdimensionaler Hilberträume verursacht ist, sollte eigens erwähnt werden, auch weil sie in den einschlägigen elementaren Einführungen der Quantenmechanik außerhalb der mathematischen Physik gewöhnlich unterschlagen wird. Quantenmechanische Operatoren sind meist unbeschränkt, und in der Funktionalanalysis gilt folgende Merkregel: Bei beschränkten Operatoren ist alles ganz einfach, bei unbeschränkten Operatoren ist alles kompliziert. Unbeschränkte Operatoren sind an jeder Stelle ihres Definitionsbereichs unstetig und immer nur auf echten Teilmengen des zugehörigen Hilbertraums definiert. Die erwähnten Schwierigkeiten, die bei allzu sorglosem Umgang mit dem Diracschen Bra-Ket-Formalismus auftreten können, sind hierauf zurückzuführen. Eine weitere Konsequenz aus der Unbeschränktheit wird im nächsten Abschnitt auftauchen.

### 4.2.2.3 Meßwerte und Spektren

Im Zusammenhang mit meßbaren Größen ist bisher nur klar, durch welche mathematischen Objekte solche Größen in der Quantenmechanik repräsentiert werden. Das dritte Axiom beschreibt nun, wie diese Repräsentation funktioniert und insbesondere wie die Verbindung zu den *Meßwerten* herzustellen ist, die bei Messungen herauskommen.

Die Meßwerte, die bei der Messung einer physikalischen Größe herauskommen können, sind die Elemente des Spektrums des zugehörigen selbstadjungierten Operators<sup>7</sup>. Das heißt, als Meßwerte können nicht nur Eigenwerte herauskommen, sondern auch alle sonstigen nichtregulären Punkte des Operators. Damit wird der Tatsache Rechnung getragen, daß bei weitem nicht alle Operatoren, die Observablen repräsentieren, reine Punktspektren besitzen; die Gleichsetzung von Eigenwerten und Meßwerten ist zu speziell. Durch die Ausweitung der möglichen Meßwerte auf alle möglichen Elemente von Spektren aller Art werden auch Operatoren mit ganz oder teilweise kontinuierlichen Spektren quantenmechanisch handhabbar. Wie schon mehrfach erwähnt ist das eher die Regel als die Ausnahme.

Damit wird aber erst recht sofort klar, warum es gerade selbstadjungierte Operatoren sein müssen. Meßwerte sind stets Zahlen, die auf irgendwelchen Anzeigegeräten direkt oder indirekt *abgelesen werden*. Mit anderen Worten, Meßwerte müssen *reell* sein. Wenn wir Meßwerte mit den Elementen der Spektren von linearen Operatoren identifizieren wollen, so müssen diese Spektren *ebenfalls reell* sein; vom Spektrum eines beliebigen linearen Operators ist zunächst jedoch nur klar, daß es eine Teilmenge der Menge der komplexen Zahlen ist. Die linearen Operatoren mit rein reellen Spektren sind *genau die selbstadjungierten Operatoren*<sup>8</sup>. Es bleibt uns also gar nichts anderes übrig, als gerade diese Klasse von Operatoren zur Repräsentation von meßbaren physikalischen Größen heranzuziehen. Dieser Sachverhalt ist im übrigen Auslöser einer weiteren weitverbreiteten, mit der im allgemeinen vorliegenden Unbeschränktheit der betrachteten Operatoren zusammenhängenden Ungenauigkeit. In vielen Lehrbüchern der Quantenmechanik ist anstelle von selbstadjungierten nur von hermiteschen Operatoren die Rede, oder es wird sogar behauptet, beides sei ein und dasselbe. Das ist jedoch nur für beschränkte Operatoren der Fall; bei unbeschränkten Operatoren folgt aus der Selbstadjungiertheit zwar die Hermitizität, nicht aber umgekehrt. Man kann zeigen, daß für die Quantenmechanik die Forderung nach Hermitezität nicht ausreicht: Erst die Selbstadjungiertheit von Operatoren garantiert, daß deren Spektren rein reell sind.

Der Nachweis der Hermitezität eines unbeschränkten Operators ist im allgemeinen einfach, der Nachweis, daß er selbstadjungiert ist, erweist sich dagegen meist als eher aufwendig, manchmal so aufwendig, daß diese Aufgabenstellung als eigene Sparte der Funktionalanalysis betrachtet werden darf. Oft erweisen sich quantenmechanische Operatoren in Tat als nicht selbstadjungiert, man kann in solchen Fällen jedoch normalerweise

---

<sup>7</sup>Der Begriff des *Spektrums eines Operators* wurde 1912 von Hilbert eingeführt [246], also zu einer Zeit, als noch niemand auch nur die Spur einer Ahnung hatte, welche Anwendungen die Spektraltheorie wenig später in der Quantenmechanik finden würde. Nach eigenen Aussagen wählte Hilbert diese Bezeichnung, ohne sich etwas bestimmtes dabei zu denken. Es ist in der Tat schon ein ziemlich verblüffender Zufall, daß beispielsweise die Bestimmung der *Spektren* von Atomen und Molekülen zu diesen Verwendungsmöglichkeiten der *Spektraltheorie* gehören.

<sup>8</sup>Ein Beweis dieses Sachverhalts steht in jedem Standard-Lehrbuch der Funktionalanalysis, beispielsweise in [204].



eine selbstadjungierte Erweiterung, also einen entsprechenden Operator mit größerem Definitionsbereich finden. Operatoren mit selbstadjungierten Erweiterungen heißen *wesentlich selbstadjungiert*, und diese Eigenschaft reicht für die Quantenmechanik bereits aus, da man dann stattdessen die Erweiterungen zur Beschreibung der betrachteten Observablen verwenden kann.

#### 4.2.2.4 Quantenmechanische Wahrscheinlichkeiten

Das vierte Axiom ist eine präzise Fassung der berühmten Wahrscheinlichkeitsinterpretation der quantenmechanischen Zustandsfunktionen. Im Gegensatz zur klassischen Physik macht die Quantenmechanik *Aussagen über Wahrscheinlichkeiten*. Dabei erklärt das Axiom nicht nur, worüber Wahrscheinlichkeitsaussagen gemacht werden, sondern auch, wie diese Wahrscheinlichkeiten zu berechnen sind. Es muß hier (wie mehrfach woanders in diesem Buch) aber ausdrücklich betont werden, daß diese Beschränkung auf Wahrscheinlichkeitsaussagen nichts mit technischem oder sonstigem Unvermögen, Unkenntnis oder dergleichen zu tun hat, wie es in der klassischen statistischen Mechanik der Fall ist; es handelt sich hierbei vielmehr um ein grundlegendes Naturgesetz. Die Natur läßt nicht mehr als Wahrscheinlichkeitsaussagen zu. Die strikte Unterscheidung von zweierlei Wahrscheinlichkeiten, einerseits klassische, andererseits quantenmechanische, ist dabei wesentlich und macht sich sowohl am Formalismus als auch in Gestalt der physikalischen Folgen deutlich bemerkbar.

Die hier gewählte, etwas ungewohnte Formulierung des vierten Axioms ist ihrer mathematischen Strenge geschuldet und wird sogleich verständlicher, wenn wir es in gewohnte Varianten übersetzen. Dazu betrachten wir (4.1) und schreiben dafür zunächst

$$\mathfrak{P}(M) = \int_M d \|\hat{E}_\lambda |\psi\rangle\|^2.$$

Wir unterscheiden nun zwei Fälle. Zunächst betrachten wir die Wahrscheinlichkeit, einen einzelnen festgelegten Meßwert  $\lambda \in \sigma(\hat{A})$  zu finden. Es geht gemäß dem vierten Axiom also darum, den Ausdruck

$$d \mathfrak{P}(\lambda) = d \|\hat{E}_\lambda |\psi\rangle\|^2.$$

genauer unter die Lupe zu nehmen. Dabei handelt es sich eigentlich um ein Spektralmaß, man kann die rechte Seite aber auch konkret auswerten und erhält

$$d \mathfrak{P}(\lambda) = \|(\hat{E}_{\lambda+d\lambda} - \hat{E}_\lambda) |\psi\rangle\|^2. \quad (4.2)$$

Dies ist eine infinitesimale Wahrscheinlichkeit, nämlich die Wahrscheinlichkeit, einen Meßwert aus dem infinitesimalen Intervall  $(\lambda, \lambda + d\lambda)$  zu finden; versucht man, daraus die Wahrscheinlichkeit dafür zu berechnen, genau  $\lambda$  zu messen, erhält man einen Grenzwert der Form  $\lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon > 0}} \|(\hat{E}_{\lambda+\varepsilon} - \hat{E}_\lambda) |\psi\rangle\|^2$ . Dies liefert nur für solche Elemente des Spektrums, die Sprungpunkte sind, also für Eigenwerte, ein nichtverschwindendes Resultat. Gehört  $\lambda$  dagegen als Stetigkeitspunkt zum rein kontinuierlichen Teil des Spektrums, so folgt  $\mathfrak{P}(\lambda) = 0$ ; dies entspricht der wohlbekannten Tatsache, daß Einzelereignisse bei kontinuierlichen Wahrscheinlichkeitsverteilungen die Wahrscheinlichkeit Null besitzen. Die Frage lautet daher zunächst, wie groß die Wahrscheinlichkeit ist, bei einer Messung gerade einen Eigenwert des die gemessene Größe repräsentierenden Operators zu erwischen,

vorausgesetzt dieser hat überhaupt welche. Aus dem Projektor  $\hat{E}_{\lambda+\varepsilon} - \hat{E}_\lambda$  wird dann im Grenzwert  $\varepsilon \rightarrow 0$  aufgrund der Sprungstelle bei  $\lambda$  der Projektor  $\hat{P}_\lambda$ , der auf den Eigenraum des Eigenwerts  $\lambda$  projiziert; die Wahrscheinlichkeit, genau diesen Eigenwert zu messen, ist dann

$$\mathfrak{P}(\lambda) = \|\hat{P}_\lambda |\psi\rangle\|^2, \quad (4.3)$$

und damit haben wir gerade jenen Spezialfall wiedergefunden, der der Formulierung in den üblichen elementaren Einführungen entspricht. Die allgemeine Formulierung mit Hilfe von Spektralscharen wird jedoch mit beliebigen Spektren aller Art fertig. Startpunkt ist wieder (4.2); aus solchen infinitesimalen Wahrscheinlichkeiten erhält man endliche Wahrscheinlichkeiten durch Überintegrieren. Daher nehmen wir uns jetzt die Wahrscheinlichkeit vor, einen Wert aus einem Intervall des Spektrums, etwa  $[\lambda, \mu]$ , als Meßergebnis zu erhalten. Diese ist gemäß dem vierten Axiom

$$\mathfrak{P}([\lambda, \mu]) = \int_\lambda^\mu d\langle \hat{E}_\lambda \psi | \psi \rangle.$$

Mit Hilfe von

$$d\hat{E}_\lambda = \hat{E}_{\lambda+d\lambda} - \hat{E}_\lambda = \hat{P}_\lambda d\lambda$$

( $\hat{P}_\lambda$  ist dabei der Projektor, der auf den zu  $\lambda$  gehörenden echten oder verallgemeinerten Eigenraum projiziert) erhalten wir daraus

$$\mathfrak{P}([\lambda, \mu]) = \int_\lambda^\mu \langle \hat{P}_{\lambda'} \psi | \psi \rangle d\lambda' = \int_\lambda^\mu \|\hat{P}_{\lambda'} |\psi\rangle\|^2 d\lambda',$$

und damit haben wir auch hier die übliche Formulierung wiedergefunden.

Wir untersuchen nun, was für Folgerungen aus den Berechnungsvorschriften für quantenmechanische Wahrscheinlichkeitsaussagen gezogen werden können. Wir beginnen wieder mit dem Fall diskreter Spektren und berechnen die Wahrscheinlichkeit, bei einer Messung der durch den Operator  $\hat{A}$  beschriebenen Größe am durch den Zustandsvektor  $|\psi\rangle$  beschriebenen System den Eigenwert  $a_n$  zu messen; dabei gelte  $\hat{A} |u_n\rangle = a_n |u_n\rangle$ . Gemäß (4.3) ergibt sich

$$\mathfrak{P}(a_n) = |\langle u_n | \psi \rangle|^2.$$

Falls der Eigenwert  $a_n$   $g_n$ -fach entartet ist,  $\hat{A} |u_n^i\rangle = a_n |u_n^i\rangle$ ,  $i = 1, 2, \dots, g_n$ , ergibt sich stattdessen

$$\mathfrak{P}(a_n) = \sum_{i=1}^{g_n} |\langle u_n^i | \psi \rangle|^2.$$

Bei kontinuierlichen Spektren findet man rein formal gesehen nur geringfügige Änderungen, abgesehen davon, daß jetzt anstelle von Eigenvektoren Eigendistributionen von  $\hat{A}$  zum Einsatz kommen. Mit  $\hat{A} |v(\lambda)\rangle = \lambda |v(\lambda)\rangle$  gilt für die Wahrscheinlichkeit, einen Wert im Intervall  $(\lambda, \lambda + d\lambda)$  zu finden,

$$d\mathfrak{P}(\lambda) = |\langle v(\lambda) | \psi \rangle|^2 d\lambda,$$

das ist wieder genau (4.2) und folglich eine infinitesimale Wahrscheinlichkeit; endliche Wahrscheinlichkeiten erhält man daraus in üblicher Weise durch Integration. Die Wahrscheinlichkeit, einen Meßwert im Intervall  $[\lambda, \mu]$  (einer Teilmenge des Spektrums von  $\hat{A}$ ) zu finden, ist somit

$$\mathfrak{P}([\lambda, \mu]) = \int_{\lambda}^{\mu} |\langle v(\lambda') | \psi \rangle|^2 d\lambda'.$$

Die Größe  $|\langle v(\lambda) | \psi \rangle|^2$  ist demnach eine *Wahrscheinlichkeitsdichte* und die Projektion  $\langle v(\lambda) | \psi \rangle \equiv \psi_v(\lambda)$  eine *Wahrscheinlichkeitsamplitude*.

Wahrscheinlichkeiten beziehungsweise Wahrscheinlichkeitsdichten für das Auftreten von Meßwerten erhält man also gerade aus den Entwicklungskoeffizienten des Zustands  $|\psi\rangle$  in der Eigenbasis des Operators  $\hat{A}$ . Dabei wird aber offenbar vorausgesetzt, daß die Eigenvektoren von  $\hat{A}$  tatsächlich ein vollständiges System im Zustandsraum bilden. Die Messung soll am betrachteten quantenmechanischen System ja in jedem möglichen Zustand durchgeführt werden können, folglich soll jeder Zustandsvektor des Systems in der Basis aus Eigenvektoren von  $\hat{A}$  entwickelt werden können. Faßt man den Begriff der Eigenvektoren als verallgemeinerte Eigenfunktionen im Distributionensinn auf, dann erfüllen gerade wieder die selbstadjungierten Operatoren diese Anforderung.

Von besonderem Interesse ist der spezielle Fall  $\hat{A} = \hat{\vec{R}} = (\hat{X}_1, \hat{X}_2, \hat{X}_3)$  des Ortsoperators; dieser ist unbeschränkt und hat keine Eigenwerte, dafür aber ein rein kontinuierliches Spektrum und ein vollständiges System  $\{ |\vec{r}\rangle \mid \vec{r} \in \mathbb{R}^3 \}$  verallgemeinerter Eigenfunktionen oder Eigendistributionen. Diese genügen einer verallgemeinerten Vollständigkeitsrelation und sind auf Delta-Funktionen orthonormiert, das heißt, es gilt

$$\int_{\mathbb{R}^3} |\vec{r}\rangle \langle \vec{r}'| d\vec{r} = \mathbf{1}$$

sowie

$$\langle \vec{r} | \vec{r}' \rangle = \delta^3(\vec{r} - \vec{r}')$$

für alle  $\vec{r}, \vec{r}' \in \mathbb{R}^3$ . Die Entwicklung von Zustandsvektoren nach den Eigendistributionen des Ortsoperators,

$$|\psi(t)\rangle = \int_{\mathbb{R}^3} |\vec{r}\rangle \langle \vec{r} | \psi(t)\rangle d\vec{r}$$

führt auf die Entwicklungskoeffizienten

$$\psi(\vec{r}, t) := \langle \vec{r} | \psi(t)\rangle$$

und damit auf genau das, was man landläufig als *Wellenfunktionen* bezeichnet. Die aus der Normierung von Zustandsvektoren automatisch folgende Normierung von Wellenfunktionen ist damit gleichbedeutend mit der Aussage, daß die Gesamtheit der Wellenfunktionen eines quantenmechanischen Systems mit dem Hilbertraum  $\mathcal{L}^2$  der quadratisch Lebesgue-integrierbaren Funktionen ist. Diesen Sachverhalt müssen wir allerdings gleich noch etwas präzisieren.

Die Wahrscheinlichkeit, ein durch den Zustandsvektor  $|\psi\rangle$  beschriebenes Quantenobjekt bei einer Ortsmessung im räumlichen Gebiet  $\mathcal{G} \subset \mathbb{R}^3$  anzutreffen, beträgt damit

$$\begin{aligned} \mathfrak{P}(\mathcal{G}) &= \int_{\mathcal{G}} |\langle \vec{r} | \psi \rangle|^2 d\vec{r} = \int_{\mathcal{G}} \langle \vec{r} | \psi \rangle \langle \psi | \vec{r} \rangle d\vec{r} \\ &= \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \int_{\mathcal{G}} |\vec{r}'\rangle \langle \vec{r}' | \psi \rangle \langle \psi | \vec{r} \rangle \langle \vec{r}'' | d\vec{r} d\vec{r}' d\vec{r}'' \\ &= \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \int_{\mathcal{G}} \delta^3(\vec{r}' - \vec{r}) \psi^*(\vec{r}) \delta^3(\vec{r} - \vec{r}'') \psi(\vec{r}) d\vec{r} d\vec{r}' d\vec{r}'' \\ &= \int_{\mathcal{G}} \psi^*(\vec{r}) \psi(\vec{r}) d\vec{r} = \int_{\mathcal{G}} |\psi(\vec{r})|^2 d\vec{r}, \end{aligned}$$

womit wir ein weiteres wohlbekanntes Ergebnis wiedergefunden haben.

Daraus ergibt sich eine wichtige allgemeine Konsequenz für Zustandsvektoren. Die Wahrscheinlichkeit, ein Quantenobjekt bei einer Ortsmessung überhaupt irgendwo im Raum anzutreffen, ist natürlich 1, das heißt, für beliebige Zustandsvektoren ist stets die Relation

$$\int_{\mathbb{R}^3} |\langle \vec{r} | \psi \rangle|^2 d\vec{r} = \int_{\mathbb{R}^3} |\psi(\vec{r})|^2 d\vec{r} = 1$$

zu fordern. Diese läßt sich wegen

$$\int_{\mathbb{R}^3} |\langle \vec{r} | \psi \rangle|^2 d\vec{r} = \int_{\mathbb{R}^3} \langle \psi | \vec{r} \rangle \langle \vec{r} | \psi \rangle d\vec{r} = \langle \psi | \mathbf{1} | \psi \rangle = \langle \psi | \psi \rangle$$

auch in der Form

$$\langle \psi | \psi \rangle = 1$$

schreiben. Mit anderen Worten: Zustandsvektoren müssen auf 1 normiert sein. Das ist stets erreichbar, denn wenn ein Zustandsvektor  $|\psi\rangle$  diese Forderung nicht erfüllt, läßt er sich normieren, indem man ihn durch den neuen Zustandsvektor

$$|\varphi\rangle := \frac{|\psi\rangle}{\langle \psi | \psi \rangle}$$

ersetzt.

Meßbare physikalische Größen nennt man auch *Observable*, wobei die Verwendung dieses Begriffs in der Literatur uneinheitlich ist. Teilweise werden auch die meßbare Größen repräsentierenden Operatoren selbst als Observable bezeichnet. Hierbei liegen allerdings nur auf den ersten Blick zwei verschiedene Begriffe für ein und dasselbe vor; auf den zweiten Blick ist die Sache sogleich nicht mehr so ganz klar. Wie beim zweiten Axiom schon angedeutet, ist durch die Festlegung, welche Klasse von Operatoren zur Beschreibung von meßbaren physikalischen Größen verwendet werden, natürlich überhaupt nichts darüber gesagt, welches denn nun genau die meßbaren physikalischen Größen sind. Das

ist eine umstrittene und keineswegs geklärte Angelegenheit. Beim zweiten Axiom wurde eine Möglichkeit bereits angedeutet, die wir nun präziser formulieren können. Es handelt sich um die Annahme, daß *jedem selbstadjungierten Operator eine physikalische Größe entspricht, die im Prinzip gemessen werden kann*. Diese Aussage wird auch als von Neumannsches Axiom bezeichnet; manche Autoren nehmen es mit in die Gruppe der Axiome der Quantenmechanik auf<sup>9</sup>. Ob diese Annahme zutrifft oder nicht, ist selbstverständlich im Rahmen der mathematischen Quantenmechanik nicht zu klären; diese liefert nur das Resultat, daß sämtliche selbstadjungierte Operatoren gleichermaßen als Observablen geeignet sind. Die Antwort auf diese Frage geht wie gesagt über den in diesem Abschnitt gesteckten axiomatischen Rahmen der Quantenmechanik hinaus und hängt davon ab, wie allgemein man den Begriff der meßbaren physikalischen Größen verstanden wissen will. Steht dabei der Aspekt der technischen Grenzen, die dem Vorgang des Messens gesetzt sein könnten, im Vordergrund, das heißt die Frage, was für Messungen tatsächlich durchführbar sind, dann ist das von Neumannsche Axiom natürlich in Frage zu stellen oder sogar zu verwerfen. Das ist aus unserer Sicht jedoch wieder eine viel zu anthropozentrische Sicht der Dinge. Wir betrachten die Quantenmechanik als Sammlung grundlegender Prinzipien der Natur und damit als einen deren mathematische Struktur manifestierenden Theorienrahmen, nicht aber als Ausdruck der menschlichen Fähigkeiten und deren Grenzen, die Natur zu beschreiben. Damit repräsentieren selbstadjungierte Operatoren mit ihren rein reellen Spektren grundsätzlich prinzipiell meßbare Größen und damit Observable, unabhängig davon, ob diese Messungen technisch tatsächlich durchführbar sind.

#### 4.2.2.5 Das Projektionspostulat

Das fünfte Axiom ist vielleicht das problematischste der sechs Axiome. So richtig deutlich wird dies allerdings erst im Zusammenhang mit dem sechsten. Wir können aber immerhin schon jetzt genauer untersuchen, was dieses fünfte Axiom eigentlich aussagt. Dazu betrachten wir ein quantenmechanisches System, dessen Zustandsvektor sich zunächst, wenn wir nichts an ihm tun, ungestört zeitlich entwickelt:

$$|\psi\rangle = |\psi(t)\rangle.$$

Wie er sich genau entwickelt, ist Inhalt des sechsten Axioms und wird gleich näher diskutiert. Zum Zeitpunkt  $t = t_0$  erfolge nun am System eine Messung der durch den selbstadjungierten Operator  $\hat{A}$  repräsentierten physikalischen Größe. Diese Messung liefere als Meßergebnis den Wert  $\lambda$  aus dem Spektrum des Operators  $\hat{A}$ . Gemäß dem fünften Axiom befindet sich das System *unmittelbar nach der Messung* nun im Zustand

$$|\psi_\lambda\rangle := \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon > 0}} \frac{(\hat{E}_{\lambda+\varepsilon} - \hat{E}_{\lambda-\varepsilon}) |\psi\rangle}{\sqrt{\langle \psi | (\hat{E}_{\lambda+\varepsilon} - \hat{E}_{\lambda-\varepsilon}) | \psi \rangle}}.$$

Was bedeutet das?  $\lambda$  ist aus dem Spektrum des Operators  $\hat{A}$  und folglich ein Wachstumspunkt, daher ändert sich die Spektralschar von  $\hat{A}$  an der Stelle  $\lambda$ . Der Zustandsvektor des Systems, an dem gemessen wird, ist also unmittelbar nach der Messung ein Element des

<sup>9</sup>Siehe zum Beispiel [153].

Projektionsraumes des Projektors  $\hat{P}_\lambda = \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon > 0}} \left[ \frac{1}{\varepsilon} (\hat{E}_{\lambda+\varepsilon} - \hat{E}_{\lambda-\varepsilon}) \right]$ . Je nachdem, ob  $\lambda$  ein

Sprungpunkt oder ein Stetigkeitspunkt ist, ist dieser Projektionsraum ein Teilraum des zugehörigen Hilbertraumes oder des diesen umfassenden Distributionenraumes eines passend gewählten Gelfandschen Raumtripels<sup>10</sup>, und entsprechend ist der Zustandsvektor unmittelbar nach der Messung ein Eigenvektor beziehungsweise eine Eigendistribution des Operators  $\hat{A}$ . Vor der Messung ist dies im allgemeinen nicht der Fall; die Messung der durch  $\hat{A}$  repräsentierten physikalischen Größe kann an beliebigen Systemen erfolgen.

Es ist lehrreich, diejenigen Fälle, bei denen die bei den Messungen herauskommenden Größen Eigenwerte von  $\hat{A}$  sind, gesondert zu betrachten. Dann ist gemäß dem fünften Axiom der Zustandsvektor des Systems unmittelbar nach der Messung

$$|\psi_n\rangle := \frac{\hat{P}_n |\psi\rangle}{\sqrt{\langle \psi | \hat{P}_n | \psi \rangle}}.$$

Wir schreiben den Zustand  $|\psi_n\rangle$  zunächst etwas um, indem wir uns an die genaue Form des darin vorkommenden Projektors erinnern,  $\hat{P}_n = \sum_{i=1}^{g_n} |u_n^i\rangle \langle u_n^i|$ . Dieser projiziert auf den zum Eigenwert  $a_n$  gehörenden Eigenraum  $\mathcal{H}_n$ ; für ihn gilt wie für jeden Projektionsoperator  $\hat{P}_n^2 = \hat{P}_n$ . Damit folgt für den Zustand unmittelbar nach der Messung

$$|\psi_n\rangle = \frac{\hat{P}_n |\psi\rangle}{\sqrt{\langle \psi | \hat{P}_n^2 | \psi \rangle}} = \frac{\sum_{i=1}^{g_n} |u_n^i\rangle \langle u_n^i | \psi \rangle}{\left\| \sum_{i=1}^{g_n} |u_n^i\rangle \langle u_n^i | \psi \rangle \right\|} \equiv \frac{\sum_{i=1}^{g_n} c_n^i |u_n^i\rangle}{\left\| \sum_{i=1}^{g_n} c_n^i |u_n^i\rangle \right\|};$$

dies ist eine *normierte Linearkombination aus Eigenvektoren von  $\hat{A}$* , also selbst ein *Eigenvektor von  $\hat{A}$* . Vor der Messung ist das im allgemeinen nicht so.

Ganz egal, was für ein Zustand vor der Messung vorlag, wir finden unmittelbar nach der Messung in jedem Fall entweder einen echten oder einen verallgemeinerten Eigenzustand der gemessenen Observablen vor, auch und insbesondere wenn das vor der Messung nicht der Fall war. Die Messung einer durch den Operator  $\hat{A}$  repräsentierten Observablen projiziert somit den Zustand des gemessenen Systems auf einen der Eigenräume von  $\mathcal{A}$ , weswegen das fünfte Axiom auch *Projektionspostulat* genannt wird. Wir haben damit in beiden beschriebenen Fällen die abrupte Änderung des Zustandsvektors des Systems bei der Messung in einen Eigenzustand der gemessenen Größe oder mit anderen Worten die *Reduktion des Zustandsvektors* wiedergefunden. Diese ist also zumindest formal im axiomatischen Aufbau der Quantenmechanik mit angelegt.

#### 4.2.2.6 Zeitentwicklungsoperator und Schrödingergleichung

Das sechste Axiom beschreibt, wie sich Zustandsvektoren im Laufe der Zeit verändern. Dazu wird ein spezieller selbstadjungierter Operator  $\hat{H}$  eingeführt, der die dynamische Entwicklung des Systems beschreibt. Der Operator  $\hat{H}$  wird *Hamilton-Operator* des betrachteten Systems genannt. Da mit ihm in der im sechsten Axiom beschriebenen Art

<sup>10</sup>Vergleiche Abschnitt 2.6.

die Zeitentwicklung von Zustandsvektoren gegeben ist, muß es sich dabei um denjenigen Operator handeln, *der die Gesamtenergie des zugehörigen quantenmechanischen Systems beschreibt*.

Damit lassen sich verschiedene Folgerungen hinsichtlich der zeitlichen Entwicklung von Zustandsvektoren ziehen. Zunächst betrachten wir den im sechsten Axiom eingeführten *Zeitentwicklungsoperator* etwas näher<sup>11</sup>, mit dem die zeitliche Entwicklung von Zuständen beschreibbar ist gemäß

$$|\psi(t)\rangle = \hat{U}(t, t') |\psi(t')\rangle.$$

Das gilt in dieser Form für alle Zeiten, also beispielsweise auch für

$$|\psi(t')\rangle = \hat{U}(t', t'') \hat{U}(t', t'') |\psi(t'')\rangle.$$

Nimmt man diese beiden Relationen zusammen, erhält man

$$\begin{aligned} |\psi(t)\rangle &= \hat{U}(t, t') |\psi(t')\rangle \\ &= \hat{U}(t, t'') |\psi(t'')\rangle \end{aligned}$$

und damit

$$\hat{U}(t, t') \hat{U}(t', t'') = \hat{U}(t, t'').$$

Dies ist die Gruppenmultiplikations-Vorschrift in der Gruppe der zeitlichen Translationen. Sei nun beispielsweise  $t'' = t$ ; daraus folgt sofort  $\hat{U}(t, t') \hat{U}(t', t) = \mathbf{1}$ . Entsprechend findet man, wenn man  $t$  und  $t'$  vertauscht,  $\hat{U}(t', t) \hat{U}(t, t') = \mathbf{1}$ . Damit ist der Zeitentwicklungsoperator *invertierbar*, und es gilt  $\hat{U}(t', t) = \hat{U}^{-1}(t, t')$ .

Dies hat eine sehr wichtige Konsequenz, die man sieht, wenn man berücksichtigt, daß für eine vernünftige statistische Interpretation der Quantenmechanik normierte Zustandsvektoren im Lauf der Zeit normiert bleiben sollten; Relationen der Form  $\langle\psi(t)|\psi(t)\rangle = 1$  gelten also *zeitunabhängig*. Mit der Zeitentwicklungsgleichung für Bra-Vektoren,

$$\langle\psi(t)| = \langle\psi(t_0)| \hat{U}^+(t, t_0),$$

lautet die Normierungsbedingung

$$\langle\psi(t)|\psi(t)\rangle = 1 = \langle\psi(t_0)| \hat{U}^+(t, t_0) \hat{U}(t, t_0) |\psi(t)\rangle.$$

Daraus folgt  $\hat{U}^+(t, t_0) \hat{U}(t, t_0) = \mathbf{1}$  und weiter  $\hat{U}(t, t_0) = \hat{U}^{-1}(t, t_0)$ . Der Zeitentwicklungsoperator  $\hat{U}(t, t_0)$  muß also *unitär* sein. Zustandsvektoren entwickeln sich, so lange man nichts mit ihnen anstellt, unitär, was zur Folge hat, daß sich Skalarprodukte währenddessen nicht ändern und damit auch Normierungen und Wahrscheinlichkeitsdichten erhalten bleiben.

Gemäß dem sechsten Axiom gilt für Zeitentwicklungsoperatoren die Differentialgleichung

$$i\hbar \frac{\partial \hat{U}(t, t_0)}{\partial t} = \hat{H}(t) \hat{U}(t, t_0) \quad (4.4)$$

<sup>11</sup>Der Zeitentwicklungsoperator wurde in Kapitel 2 aus anderer Perspektive bereits ausführlich diskutiert.



mit der Randbedingung  $\hat{U}(t_0, t_0) = \mathbf{1}$ . Die formale Lösung dieser Differentialgleichung lautet

$$\hat{U}(t, t_0) = \mathbf{1} - \frac{i}{\hbar} \int_{t_0}^t \hat{H}(t') \hat{U}(t', t_0) dt', \quad (4.5)$$

aber damit ist sie eben nur formal gelöst oder anders gesagt, aus einer Differentialgleichung ist eine Integralgleichung geworden. Für Zustandsvektoren folgt daraus entsprechend

$$|\psi(t)\rangle = \hat{U}(t, t_0) |\psi(t_0)\rangle = |\psi(t_0)\rangle - \frac{i}{\hbar} \int_{t_0}^t \hat{H}(t') |\psi(t')\rangle dt',$$

also ebenfalls eine Integralgleichung. Wenn  $\hat{H}$  *zeitabhängig* ist, kann man solche Gleichungen abgesehen von Spezialfällen nicht allgemein lösen<sup>12</sup>.

Dafür kann man mit Hilfe von (4.5) ganz allgemein die Unitarität der Zeitentwicklungsoperatoren beweisen und damit sicherstellen, daß diese physikalische Notwendigkeit auch mathematisch zwingend erfüllt ist. Man betrachte dazu einen Zeitpunkt  $t'$ , ein infinitesimales Zeitintervall  $dt$  und den zugehörigen Zeitentwicklungsoperator; für diesen folgt aus (4.5)

$$\hat{U}(t' + dt, t') = \mathbf{1} - \frac{i}{\hbar} \hat{H}(t') dt. \quad (4.6)$$

Da der Hamilton-Operator im allgemeinen selbstadjungiert ist oder zumindest wesentlich selbstadjungiert und damit durch seinen eindeutig definierten selbstadjungierten Abschluß ersetzt werden kann, wird durch (4.6) eine infinitesimale unitäre Transformation definiert. Zeitentwicklungsoperatoren für endliche Zeitintervalle erhält man durch Hintereinanderausführen infinitesimaler Transformationen, so daß sich der gesamte Zeitentwicklungsoperator  $\hat{U}(t, t_0)$  als Operatorprodukt aus unendlich vielen infinitesimalen Transformationen ergibt. Durch Nacheinanderausführen unitärer Transformationen erhält man jedoch stets wieder solche, womit gezeigt ist, daß Zeitentwicklungsoperatoren generell unitär sind.

Wie gesagt sind Integralgleichungen der eben beschriebenen Form im allgemeinen schwer zu handhaben bis unlösbar. Bei konservativen Systemen sieht das anders aus. Aufgrund des Energieerhaltungssatzes ist bei solchen Systemen  $\hat{H}$  *zeitunabhängig*. Aus der Differentialgleichung (4.4) für  $\hat{U}$  folgt damit sofort (Nachrechnen)

$$\hat{U}(t, t_0) = \hat{U}(t - t_0) = \exp\left(-i \frac{t - t_0}{\hbar} \hat{H}\right). \quad (4.7)$$

Die Unitarität läßt sich hier natürlich leicht direkt nachprüfen. Es gilt

$$\hat{U}(t - t_0) \hat{U}^\dagger(t - t_0) = \exp\left(-i \frac{t - t_0}{\hbar} \hat{H}\right) \exp\left(i \frac{t - t_0}{\hbar} \hat{H}\right)$$

<sup>12</sup>Wie häufig in der Physik ist man auch hier meist auf Näherungsverfahren angewiesen. Beispielsweise gibt es Fälle, in denen man die Integralgleichung (4.5) *iterativ* lösen kann, das heißt, man startet mit einer Näherung, löst das Integral, setzt das Ergebnis wieder ein, löst das Integral erneut, und so weiter. Man erhält so unendliche Reihen, die man Dyson Reihen nennt. Diese können je nach gewünschter Genauigkeit abgebrochen werden. Dabei bilden Fragen nach der Konvergenz solcher Reihen gleich wieder ein eigenes Spezialgebiet der mathematischen Physik. – Dieses Verfahren bildet die Grundlage der *zeitabhängigen Störungstheorie* und damit auch ein wichtiges Werkzeug im Rahmen der Quantenfeldtheorie. Wir haben in Abschnitt 2.5.2 davon Gebrauch gemacht.

$$\begin{aligned}
&= \exp\left(-i \frac{t-t_0}{\hbar} \hat{H}\right) \exp\left(i \frac{t-t_0}{\hbar} \hat{H}\right) \\
&= \exp\left(-i \frac{t-t_0}{\hbar} \hat{H} + i \frac{t-t_0}{\hbar} \hat{H}\right) = \mathbf{1},
\end{aligned}$$

und analog

$$\hat{U}^+(t-t_0) \hat{U}(t-t_0) = \mathbf{1}.$$

Für Systeme mit zeitunabhängigem Hamilton-Operator gilt daher für die Zeitentwicklung von Zuständen

$$|\psi(t)\rangle = \exp\left(-i \frac{t-t_0}{\hbar} \hat{H}\right) |\psi(t_0)\rangle.$$

Hieran wird erneut die grundlegende Bedeutung der Selbstadjungiertheit quantenmechanischer Operatoren deutlich. Denn nur für selbstadjungierte Operatoren  $\hat{A}$  sind Ausdrücke der Form  $\hat{U} = e^{i\hat{A}}$  unitäre Operatoren; Abgeschlossenheit und Hermitezität genügen auch hier nicht. Die Selbstadjungiertheit des Hamilton-Operators garantiert also die Unitarität des Zeitentwicklungsoperators.

Üblicherweise werden in der Physik dynamische Entwicklungen durch Differentialgleichungen beschrieben. Die Lösungen dieser Differentialgleichungen sind dann den Zustand des Systems beschreibende zeitabhängige Funktionen. Entsprechend läßt sich auch hier eine Differentialgleichung konstruieren, für welche die obigen Ausdrücke für  $|\psi(t)\rangle$  Lösungen darstellen. Beachten wir, daß (4.4) eine *operatorwertige Differentialgleichung* ist und wenden wir die rechte und die linke Seite dieser Differentialgleichung jeweils auf den Anfangszustand  $|\psi(t_0)\rangle$  an, so erhalten wir

$$i\hbar \frac{\partial \hat{U}(t, t_0)}{\partial t} |\psi(t_0)\rangle = \hat{H}(t) \hat{U}(t, t_0) |\psi(t_0)\rangle.$$

Wegen der Zeitentwicklungsgleichung  $|\psi(t)\rangle = \hat{U}(t, t_0) |\psi(t_0)\rangle$  erfüllt somit die Differentialgleichung

$$i\hbar \frac{d|\psi(t)\rangle}{dt} = \hat{H} |\psi(t)\rangle \quad (4.8)$$

genau die obige Forderungen. Wir haben es hier mit der *abstrakten Formulierung der Schrödingergleichung* zu tun; die im Hauptteil des vorliegenden Buches an verschiedenen Stellen auftauchenden Formen der Schrödingergleichung stellen einfach Spezialfälle dieser allgemeinen Form dar, in Ortsdarstellung und mit speziellen Hamilton-Operatoren. Obwohl die obige Formulierung mit dem Zeitentwicklungsoperator auf den ersten Blick mit der Voraussetzung *normierter Zustandsvektoren* eine zusätzliche Anforderung an die Kets zu beinhalten scheint, ist sie doch etwas allgemeiner als eine entsprechende Formulierung mit der Schrödingergleichung. Denn erstens läßt sich jeder Zustandsvektor normieren (gegebenenfalls unter Zuhilfenahme einer etwas allgemeineren Auffassung des Begriffs der Normierung), sodaß dies keine zusätzliche Forderung darstellt, und zweitens ist, um die Schrödingergleichung überhaupt sinnvoll aufstellen zu können, nun gerade hier eine zusätzliche Forderung an die Zustandsvektoren zu stellen, nämlich  $|\psi_0\rangle \in \mathcal{D}(\hat{H})$ . Für die Praxis spielt dies jedoch kaum eine Rolle.

Eine sehr wesentliche Rolle spielt jedoch die Tatsache, daß es sich sowohl bei der Operator-Differentialgleichung (4.4) als auch bei der Schrödingergleichung (4.8) um *lineare Differentialgleichungen* handelt. Dies ist auf den ersten Blick trivial, denn es kann gar nicht anders sein. Nichtlineare Differentialgleichungen wären mit dem Hilbertraum-Formalismus nicht verträglich, denn nur die Lösungsmannigfaltigkeiten von linearen Differentialgleichungen tragen eine lineare Struktur. Dennoch handelt es sich hier um eine physikalisch gesehen sehr folgenreiche Angelegenheit, die sich zusammen mit der Hilbertraumstruktur der Zustandsräume im *Superpositionsprinzip* manifestiert. Dazu wurde bereits ausführliches gesagt.

Betrachtet man nun wieder den Spezialfall von Systemen, bei denen  $\hat{H}$  zeitunabhängig ist, also Systeme mit Zeitentwicklungsoperatoren der Form (4.7), so sieht man, daß mit der Anfangsbedingung  $|\psi(t_0)\rangle = |\varphi\rangle$ , der aus (4.7) resultierende Zustandsvektor

$$|\psi(t)\rangle = \exp\left(-i\frac{t-t_0}{\hbar}\hat{H}\right)|\varphi\rangle \quad (4.9)$$

die eindeutig bestimmte Lösung der Schrödingergleichung (4.8) ist. – Man könnte annehmen, daß es sich hier um sehr spezielle Sonderfälle handelt. Die Annahme der Zeitunabhängigkeit ist jedoch eine weniger starke Einschränkung als es auf den ersten Blick zu sein scheint, denn zeitabhängige Hamilton-Operatoren sind automatisch mit der Verletzung des Energieerhaltungssatzes für die zugehörigen Systeme verbunden; das wiederum bedeutet, daß es sich dabei nicht um isolierte Systeme handeln kann, sondern um offene Systeme, also solche, die mit irgendeiner Umgebung, das heißt mit anderen Systemen in Wechselwirkung stehen. Konsequenterweise faßt man das betrachtete System mit den mit ihm wechselwirkenden Systemen zu einem neuen, größeren System zusammen, das dann abgeschlossen ist, den Energieerhaltungssatz erfüllt und folglich durch einen zeitunabhängigen Hamilton-Operator beschrieben wird. Natürlich gibt es genaugenommen nur ein ideal abgeschlossenes System, nämlich das ganze Universum, und auch hier ist im Zusammenhang mit Paralleluniversen und dergleichen noch nicht das letzte Wort gesprochen. Kleinere isolierte Systeme sind daher nicht realisierbare Idealisierungen. Es zeigt sich jedoch, daß diese Idealisierungen häufig großen praktischen Nutzen haben, da man es in der Quantenmechanik verbreitet mit Systemen zu tun hat, die man in sehr guter Näherung als abgeschlossen betrachten kann.

Ein noch speziellerer, aber nicht minder wichtiger Spezialfall liegt vor, wenn man sich (wie häufig in der Physik) für sogenannte *stationäre Zustände* interessiert, das heißt für Zustände, die von der Zeit unabhängig sind. Hierbei läßt sich das folgende grundlegende Resultat beweisen, von dem in der angewandten Quantenmechanik rege Gebrauch gemacht wird: *Ein Zustand eines quantenmechanischen Systems ist genau dann stationär, wenn der diesen Zustand beschreibende Zustandsvektor  $|\psi\rangle$  ein Eigenzustand des zum System gehörenden Hamilton-Operators  $\hat{H}$  ist*<sup>13</sup>. Das System befindet sich dann also in einem *Energieeigenzustand*, das heißt, die Gesamtenergie solcher Systeme ist scharf definiert und kann daher (zumindest im Prinzip) streuungsfrei gemessen werden. Diese taucht dann in der entsprechenden Eigenwertgleichung

$$\hat{H}|\psi\rangle = E|\psi\rangle$$

<sup>13</sup>Den etwas technischen Beweis dieses Satzes findet der Leser beispielsweise in [507].

in Form des Eigenwerts  $E$  auf. Aus (4.9) wird dann aber ( $t_0$  sei ohne Beschränkung der Allgemeinheit  $= 0$  gesetzt)

$$|\psi(t)\rangle = \exp\left(\frac{-iEt}{\hbar}\right) |\varphi\rangle.$$

Der Vorfaktor  $\exp(-iEt/\hbar)$  wird üblicherweise dynamischer Phasenfaktor genannt; eingesetzt in die Schrödingergleichung ergibt sich

$$\hat{H} |\psi(t)\rangle = \hat{H} \exp\left(\frac{-iEt}{\hbar}\right) |\varphi\rangle = E \exp\left(\frac{-iEt}{\hbar}\right) |\varphi\rangle$$

und daraus, wenn man den dynamischen Phasenfaktor hinauswirft,

$$\hat{H} |\varphi\rangle = E |\varphi\rangle. \quad (4.10)$$

Das ist eigentlich wieder nichts anderes als die Eigenwertgleichung des Hamilton-Operators, aber darüberhinaus genau das, was üblicherweise als *zeitunabhängige* oder *stationäre Schrödingergleichung* bezeichnet wird. Konsequenterweise nennt man die ausführliche Variante (4.8) auch *zeitabhängige Schrödingergleichung*. Die  $|\varphi\rangle$ 's, die von den  $|\psi(t)\rangle$ 's nach Abspalten des dynamischen Phasenfaktors übrigbleiben, werden ebenfalls stationäre Zustände des betrachteten Systems genannt.

Da es sich bei (4.10) um eine Eigenwertgleichung handelt, hat man es beim Aufsuchen von stationären Zustände aus mathematischer Sicht natürlich mit der Spektraltheorie des das System beschreibenden Hamilton-Operators zu tun. Das ist auch schon deswegen völlig klar, weil man sich für die Gesamtenergie des betrachteten physikalischen Systems interessiert, die im Falle von zeitunabhängigen Hamilton-Operatoren wiederum nichts anderes als die durch letztere beschriebene Meßgröße ist, und gemäß Axiom 3 sind die einzig möglichen Meßwerte, die bei der Messung der durch einen quantenmechanischen Operator repräsentierten Größe herauskommen können, die Elemente des Spektrums dieses Operators. Es ist als besonders wichtiger Erfolg der mathematischen Physik zu werten, daß man für die wichtigsten – oder einfachsten<sup>14</sup> – Hamilton-Operatoren der Quantenmechanik in den meisten Fällen nachweisen kann, daß sie selbstadjungiert sind<sup>15</sup>. Die Spektraltheorie dieser sogenannten Schrödinger-Operatoren hat sich zu einer eigenständigen mathematischen Disziplin entwickelt. Hier ist stattdessen in erster Linie von Interesse, welche *physikalischen* Aussagen man aus dem Spektrum des Hamilton-Operators eines Systems ableiten kann, da dies für die Interpretation des Axioms 6 von grundlegender Bedeutung ist.

Hamilton-Operatoren besitzen als selbstadjungierte Operatoren natürlich rein reelle Spektren, dies können jedoch wie bei jedem selbstadjungierten Operator rein diskret, rein kontinuierlich oder auch gemischt sein. Folglich repräsentieren die eben erwähnten stationären Zustände tatsächlich nicht die allgemeine Situation, mit der man bei der Beschreibung quantenmechanischer Systeme durch Hamilton-Operatoren konfrontiert ist. Wie gesagt sind die stationären Zustände genau die Eigenzustände zu den Eigenwerten des Hamilton-Operators, womit ausdrücklich nicht verallgemeinerte Eigenwerte gemeint

<sup>14</sup>Das ist reine Ansichtssache.

<sup>15</sup>Siehe Fußnote 19 in Abschnitt 2.3.1.2.

sind. Denn andererseits gilt ganz allgemein, daß zu den Elementen des kontinuierlichen Spektrums eines selbstadjungierten Operators keine Eigenzustände desselben im eigentlichen Sinne gehören, jedenfalls keine, die im zugehörigen Hilbertraum enthalten sind. Dafür gibt es zu diesen Elementen des Spektrums im allgemeinen verallgemeinerte Eigenzustände, also Eigendistributionen, die Elemente des den Hilbertraum umfassenden Distributionenraumes des zugehörigen Gelfandschen Raumtripels sind. Wir haben gesehen, daß solche Eigendistributionen zwar keine Hilbertraumelemente sind, man jedoch unter geeigneten Voraussetzungen, die im Fall von Hamilton-Operatoren erfüllt sind, alle Hilbertraumelemente nach ihnen entwickeln kann, oder anders gesagt Wellenpakete aus ihnen konstruieren kann, die dann wieder im Hilbertraum enthalten sind. Zusammengefaßt gehören zum Punktspektrum stationäre Zustände, also Energieeigenzustände und zum kontinuierlichen Spektrum andere wilde Zustände, die man zwar zu Wellenpaketen bändigen kann, die sich aber irgendwie nicht so schön brav verhalten wie die übrigen. Dieser Unterschied läßt sich auch noch anders charakterisieren. Für Lösungen  $|\psi\rangle = e^{-i\hat{H}t/\hbar}|\varphi\rangle$  der zeitabhängigen Schrödingergleichung läßt sich folgendes, mit der Struktur des Spektrums des zugehörigen, dafür als zeitunabhängig vorauszusetzenden Hamilton-Operators<sup>16</sup> zusammenhängendes Theorem beweisen<sup>17</sup>:

**Satz:** (Ruelle)

- a) Ist  $\varphi$  ein Element der linearen Hülle der Menge der Eigenvektoren von  $\hat{H}$ , dann gilt

$$\lim_{R \rightarrow \infty} \int_{\|x\| \geq R} |\psi|^2 dx = \lim_{R \rightarrow \infty} \int_{\|x\| \geq R} |\varphi|^2 dx = 0.$$

- b) Ist  $\varphi$  ein Element des orthogonalen Komplements der linearen Hülle der Menge der Eigenvektoren von  $\hat{H}$ , dann gilt

$$\lim_{R \rightarrow \infty} \int_{\|x\| \leq R} |\psi|^2 dx = \lim_{R \rightarrow \infty} \int_{\|x\| \leq R} |\varphi|^2 dx = 0.$$

Das muß man sich sehr genau durchlesen. Zunächst handelt es sich dabei im Bezug auf die Zustandsvektoren um genau dieselbe Unterteilung wie oben. Der Unterscheid der beiden Fälle in dem Theorem liegt darüberhinaus in der Richtung des Größer-Kleiner-Zeichens. Die in a) beschriebenen Zustände bleiben, was ihre Aufenthaltswahrscheinlichkeit betrifft, im wesentlichen innerhalb von geeignet gewählten räumlichen Bereichen; sie sind in gewissem Sinne *lokalisiert*. Solche Zustände nennt man *gebundene Zustände* und bringt damit zum Ausdruck, daß die durch sie beschriebenen Systeme räumlich stabil sind. Die in b) beschriebenen Zustände dagegen neigen dazu, jeden beliebigen vorgegebenen räumlichen Bereich zu verlassen. Deswegen nennt man sie *freie Zustände* oder *Streuzustände*. Für das Spektrum des Hamilton-Operators eines quantenmechanischen Systems bedeutet dies, daß dessen diskretem Teil die gebundene Zustände und dem kontinuierlichen Teil die Streuzustände zugeordnet sind.

<sup>16</sup>Man beachte dazu die Bemerkungen, die weiter oben zu zeitabhängigen und zeitunabhängigen Hamilton-Operatoren gemacht wurden.

<sup>17</sup>Dieses Resultat wurde 1969 von D. Ruelle entdeckt [438]. Weitergehende Informationen über Schrödingeroperatoren findet man in [207] sowie besonders ausführlich in [111], [250] und [426] - [429].

Da Streuzustände folglich nicht auf 1, sondern nur auf Deltafunktionen normierbar sind, Einteilchen-Zustandsräume der Standard-Quantenmechanik jedoch stets isometrisch isomorph zum Raum  $\mathcal{L}^2(\mathbb{R}^3)$  sind, sind Streuzustände *keine Hilbertraum-Elemente*. Durch eine geeignete Erweiterung des Hilbertraums kann man sie jedoch mit ins Boot nehmen. Man erreicht dies wiederum durch die Einführung geeigneter Gelfandscher Raumtripel. Grundlage dafür ist das folgende fundamentale Resultat<sup>18</sup>.

**Satz:** (Gelfand, Wilenkin)  $\hat{A}$  sei ein linearer beschränkter hermitescher Operator in einem nuklearen Raum  $\Phi$ , der im zugehörigen Gelfandschen Raumtripel  $\Phi \subset \mathcal{H} \subset \Phi'$  zu einem selbstadjungierten – in  $\mathcal{H}$  nicht notwendig beschränkten – Operator  $\hat{A} = \hat{A}^+$  erweitert werden kann. Dann hat  $\hat{A}$  in  $\Phi'$  ein vollständiges System von Eigendistributionen  $|\tilde{\psi}_\lambda\rangle$ , nach denen jedes Element von  $\mathcal{H}$  entwickelt werden kann.

Der Operator  $\hat{A}$  ist in  $\Phi$  wesentlich selbstadjungiert und kann folglich zu einem selbstadjungierten Operator erweitert werden. Wie in der Schreibweise der Physik üblich, bezeichnen wir diesen erweiterten Operator mit dem selben Symbol; der Operator  $\hat{A}$  ist in  $\mathcal{H}$  sowie in  $\Phi'$  selbstadjungiert. Verallgemeinerte Eigenvektoren sind *keine* Hilbertraumelemente, aber sie liefern beim Überintegrieren mit geeigneten Koeffizienten *dennoch* Hilbertraumelemente. Man nennt das Entwickeln von Hilbertraumvektoren nach verallgemeinerten Eigenvektoren auch *Bilden von Wellenpaketen*.

Natürlich ist der allgemeinste Fall derjenige von selbstadjungierten Operatoren, deren Spektren sowohl diskrete als auch kontinuierliche Anteile aufweisen. Die zugehörigen vollständigen Orthonormalsysteme bestehen dabei sowohl aus echten Eigenvektoren, die auf Kronneckersymbole normiert sind, als auch aus Eigendistributionen, die auf Deltafunktionen normiert sind. Die Maße  $\sigma(\lambda)$ , die in den Entwicklungen von Vektoren nach solchen vollständigen Orthonormalsystemen auftauchen, sind dementsprechend Summen aus Maßen, die auf isolierte Punkte konzentriert sind, und aus Maßen, die (ohne Beschränkung der Allgemeinheit, wie wir gesehen haben) bezüglich des Lebesgue-Maßes absolut stetig sind. Berücksichtigt man zusätzlich, daß verallgemeinerte Eigenvektoren die echten als Spezialfälle mit umfassen, dann sieht man, daß sich die Entwicklung des beliebigen Hilbertraumvektors  $|\varphi\rangle$  nach Eigenelementen,

$$|\varphi\rangle = \int c(\lambda) |\tilde{\psi}_\lambda\rangle d\sigma(\lambda) = \int \langle \tilde{\psi}_\lambda | \varphi \rangle |\tilde{\psi}_\lambda\rangle d\sigma(\lambda)$$

im allgemeinen Fall in ein Integral und eine Summe zerlegen läßt,

$$\begin{aligned} |\varphi\rangle &= \sum_i c_i |\psi_i\rangle + \int c(\lambda) |\tilde{\psi}_\lambda\rangle d\sigma(\lambda) \\ &= \sum_i \langle \psi_i | \varphi \rangle |\psi_i\rangle + \int \langle \tilde{\psi}_\lambda | \varphi \rangle |\tilde{\psi}_\lambda\rangle d\sigma(\lambda), \end{aligned}$$

und auch die Vollständigkeitsrelation

$$\int |\tilde{\psi}_\lambda\rangle \langle \tilde{\psi}_\lambda| d\sigma(\lambda) = \mathbf{1}$$

<sup>18</sup>Einen Beweis findet man in [186].



läßt sich entsprechend schreiben,

$$\sum_i |\psi_i\rangle \langle \psi_i| + \int |\tilde{\psi}_\lambda\rangle \langle \tilde{\psi}_\lambda| d\sigma(\lambda) = \mathbf{1}.$$

Zusammenfassend können wir das wie folgt formulieren:

*Zu jedem (beschränkten oder unbeschränkten) selbstadjungierten Operator  $\hat{A}$  gibt es ein vollständiges orthonormiertes System aus Eigenvektoren; dieses besteht entweder, falls  $\hat{A}$  ein reines Punktspektrum hat, aus echten Eigenvektoren  $|\psi_i\rangle \in \mathcal{H}$ , oder, falls  $\hat{A}$  ein rein kontinuierliches Spektrum besitzt, aus verallgemeinerten Eigenvektoren (Eigendistributionen)  $|\tilde{\psi}_\lambda\rangle \in \Phi' \setminus \mathcal{H}$ , oder bei gemischtem Spektrum von  $\hat{A}$  sowohl aus echten als auch aus verallgemeinerten Eigenvektoren.*

Aus mathematischer Sicht ist die Bedeutung von Eigendistributionen damit klar: Sie ermöglichen das Entwickeln jedes Hilbertraumelementes nach Eigenelementen jedes selbstadjungierten Operators. Was jetzt noch fehlt, ist eine vernünftige physikalische Interpretation der Elemente der Distributionenräume Gelfandscher Raumtripel, die wir in Anlehnung an die verallgemeinerten Eigenvektoren *verallgemeinerte Zustandsvektoren* nennen werden. Die Aussage, daß es keine Hilbertraumelemente sind, ist so offensichtlich richtig wie physikalisch nichtssagend, wenn man nicht dazusagt, was sie sonst sind. Den entscheidenden Hinweis dafür liefert die Tatsache, daß verallgemeinerte Eigenvektoren immer zu *kontinuierlichen* verallgemeinerten Eigenwerten gehören, während echte Eigenvektoren stets *diskreten* echten Eigenwerten zugeordnet sind. Den verallgemeinerten Eigenvektoren eines selbstadjungierten Operators entsprechen daher Zustände des zugehörigen physikalischen Systems, bei denen die durch den Operator repräsentierte Meßgröße in Form eines scharf definierten Wertes dieser Meßgröße aus einem kontinuierlichen Wertevorrat, nämlich dem kontinuierlichen Spektrum des Operators, realisiert ist. Bei echten Eigenvektoren sind es dagegen nur Werte aus einem diskreten Wertevorrat. Physikalische Systeme mit scharf definierten Werten kontinuierlicher Meßgrößen stellen aber stets Idealisierungen dar. Das hat nichts mit mangelnder Meßgenauigkeit zu tun; die Ursache liegt vielmehr in der allgemeinen Robertson'schen Unschärferelation. Sobald eine physikalische Größe an einem quantenmechanischen System scharf vorliegen, das heißt mit anderen Worten streuungsfrei meßbar sein soll, hat das zwangsläufig unendliche Streuung bei Messungen der dazu konjugierten Größe zur Folge. Entsprechend können die von verallgemeinerten Zustandsvektoren beschriebenen Zustände in der Natur nur näherungsweise realisiert werden. Das manifestiert sich im quantenmechanischen Formalismus in der Entwickelbarkeit jedes Hilbertraumelements nach Eigendistributionen selbstadjungierter Operatoren mit ganz oder teilweise kontinuierlichen Spektren, also darin, was man volkstümlich als das Bilden von Wellenpaketen bezeichnet.

Die geläufigsten Beispiele dafür sind lokalisierte Zustände, die durch Deltafunktionen beschrieben werden, oder Zustände mit definiertem Impuls, die durch ebene Wellen repräsentiert werden. Beides sind keine Hilbertraumelemente, und in beiden Fällen sind die Vorstellungen eines beliebig genau festliegenden Ortes beziehungsweise eines beliebig scharf definierten Impulses des betrachteten Systems Idealisierungen, die so in der Natur nicht vorkommen. Ein weiteres Beispiel stellen die oben beschriebenen Streuzustände dar. Sie sind verallgemeinerte Eigenzustände zum kontinuierlichen Teil des Spektrums



des Hamilton-Operators, der zum betrachteten System gehört. Das Faszinierende daran ist insbesondere auch die Tatsache, daß der mathematische Formalismus selbst sehr genau zu unterscheiden weiß, welche Zustände Idealisierungen und welche tatsächlich realisiert sind, eine Fähigkeit, die zu erreichen den Entdeckern der Hilbertraum-Quantenmechanik keineswegs im Sinn stand. Der Formalismus leistet das von selbst<sup>19</sup>.

### 4.3 Welle-Teilchen-Dualismus oder: Was sind Quantenobjekte?

Mit dem Begriff „Welle-Teilchen-Dualismus“ wurde ursprünglich und wird verbreitet nach wie vor ganz einfach zum Ausdruck gebracht, daß sich Quantenobjekte anscheinend manchmal wie Wellen und manchmal wie Teilchen verhalten, je nachdem was man für Experimente mit ihnen anstellt. Dabei scheint auch die zusätzliche Aussage mit drinzustecken, diese beiden Eigenschaften seien komplementär in dem Sinne, daß sie niemals gleichzeitig und damit bei allen Experimenten fein säuberlich getrennt im Sinne von entweder-oder auftreten. Als Beispiele werden dafür typischerweise Experimente mit Elektronen oder mit Photonen angeführt. So werden im Falle von Elektronen Teilchenspuren in Nebelkammern oder Elektronenkanonen in Brownschen Röhren oder Fernsehapparaten als Beispiele für teilchenartiges und Elektronenbeugung an Kristallen oder Gittern als Beispiel für wellenartiges Verhalten erwähnt; entsprechend werden optische Interferenzphänomene aller Art als Beispiele dafür herangezogen, daß Photonen als Wellen auftreten, und der Photoeffekt als Beispiel dafür, daß sie als Teilchen in Erscheinung treten. Wie wir sehen werden, greift diese Definition des Welle-Teilchen-Dualismus viel zu kurz; allerdings haben die Schwierigkeiten, die dieses Prinzip zu bereiten scheint, andere Ursachen als nur eine unpräzise Begriffsbeschreibung. Sie liegen einerseits in der aus klassischer Sicht naheliegenden Frage, was denn Quantenobjekte dann *wirklich* seien, Wellen *oder* Teilchen. Stellt man sie, so wird sehr schnell deutlich, warum die Sache als problematisch empfunden wird: Der (in der klassischen Sicht verhaftete) gesunde Menschenverstand tut sich ungemein schwer, solcherart sich beliebig widersprechende Vorstellungen wie Wellen und Teilchen unter einen Hut zu bekommen. Das führt dazu, den Welle-Teilchen-Dualismus zu einer rein heuristischen Vorstellung mit allenfalls didaktischer Bedeutung bei Einführungen in die Quantenmechanik zu degradieren, ihn als offenes Problem oder Unvollständigkeit der quantenmechanischen Naturbeschreibung aufzufassen oder gar dazu, sich zu der Überzeugung zu versteigen, der Welle-Teilchen-Dualismus müsse „überwunden“ werden, was immer das auch heißen soll<sup>20</sup>. Andererseits steht dem korrekten Verständnis eine Auffassung im Wege, die jahrzehntelange Tradition hat und nach wie vor verbreitet in Lehrbüchern wie Physikergehirnen anzutreffen ist, nämlich diejenige, der Welle-Teilchen-Dualismus sei eine Konsequenz der Heisenberg-

<sup>19</sup>Ein solches Eigenleben von Theorien der mathematischen Physik im Sinne des zusätzlichen Lieferns von nicht erwarteten Aussagen mit realem physikalischen Gehalt ist auch in anderen Gebieten zu finden. Ein oben bereits erwähntes, prominentes Beispiel ist die kosmologische Konstante, die aus rein mathematischen Gründen als Bestandteil der Einsteinschen Feldgleichungen der Gravitation auftauchen muß, und inzwischen in Form der dunklen Energie ihr physikalisches Gegenstück im Universum gefunden hat.

<sup>20</sup>Hier tritt wieder einmal die gymnasiale Schulphysik insbesondere in ihrer in Baden-Württemberg häufig anzutreffenden Ausprägung nicht selten als negatives Beispiel in Erscheinung. Siehe dazu beispielsweise [36] oder ausführlicher [37].

schen Unschärferelation. Das ist vollkommen falsch; allerdings wurde hierüber erst in den 90er Jahren wirkliche Klarheit erlangt.

### 4.3.1 Das Komplementaritätsprinzip

Verschafft man sich einen Überblick über die Literatur zum Thema des vorliegenden Abschnitts, so fällt schnell auf, daß häufig von *Komplementarität* anstelle von *Welle-Teilchen-Dualismus* die Rede ist. Dieser Begriff geht ursprünglich auf Nils Bohr zurück [67], [68], der aber weder jemals sehr präzise noch jemals besonders einheitlich zum Ausdruck brachte, was nun genau unter dem Komplementaritätsprinzip zu verstehen sei. Dessenungeachtet lassen sich von Beginn an zwei grundlegende Aussagen herauskristallisieren, die unter der Rubrik „Komplementarität“ auftauchen:

- Observablen, die durch nichtkommutierende selbstadjungierte Operatoren repräsentiert werden, sind komplementär in dem Sinne, daß sie nicht gleichzeitig mit beliebig kleiner Streuung meßbar sind.
- Es ist nur möglich, *entweder* etwas über den Weg von Quantenobjekten herauszufinden *oder* Interferenzphänomene an ihnen zu beobachten, nicht aber beides gleichzeitig. Man sagt dabei kurz und präzise: Die Möglichkeit von Welcher-Weg-Information<sup>21</sup> verhindert das Auftreten von Interferenzerscheinungen.

Die erste Aussage ist einfach eine Folgerung aus der allgemeinen Unschärferelation; die zweite wurde traditionell als ein Sonderfall und damit als eine Folge der ersteren betrachtet, indem diese auf Ort und im Impuls bezogen wird. Der Ortsoperator mit seinen lokalisierten Eigenfunktionen wird dabei als Repräsentant für Teilchen- und der Impulsoperator mit Eigenfunktionen in Form ebener Wellen als solcher für Wellenverhalten betrachtet. Typischerweise wird dabei das Verschwinden des Interferenzmusters mit einem Impulsübertrag auf die betrachteten Quantenobjekte erklärt, der beim Beschaffen der Welcher-Weg-Information auftritt. In exemplarischer Weise wird diese Argumentation von Bohr bei seinen Debatten mit Einstein zu den Grundlagen der Quantenmechanik durchexerziert<sup>22</sup>. Einstein schlug dabei in einem berühmten Gedankenexperiment vor, beim Doppelspaltexperiment den jeweiligen Spalt, durch den ein Teilchen hindurchgegangen sein soll, durch eine Messung des Impulses herauszufinden, der vom Doppelspalt auf das Teilchen übertragen wird und der unterschiedlich ausfallen müßte, je nachdem welchen Spalt das Teilchen passiert hat [445]. Bohr hingegen konnte zeigen, daß dieser Impulsübertrag stets entweder so klein ist, daß er keine Rückschlüsse auf einen etwaigen Weg des Teilchens zuläßt, oder so groß, daß das Interferenzmuster dadurch zerstört wird [445]. Entscheidend ist dabei, daß die Gesetze der Quantenmechanik nicht nur für das Teilchen, sondern auch für den Doppelspalt gelten. Bohr betrachtete dazu einen Doppelspalt mit Spaltabstand  $\Delta x$  und ging davon aus, daß die Ortsunbestimmtheit  $\delta x$  auf jeden Fall sehr viel kleiner sein muß als  $\Delta x$ , um überhaupt ein Interferenzmuster zuzulassen,

<sup>21</sup>Die Begriff der Welcher-Weg-Information (Which-Way-Information) stammt von Scully, Englert und Walther. Eine Verallgemeinerung wurde von Sen, Basu und Sengupta vorgeschlagen [471]; sie sprechen dabei vom Zusammenhang zwischen Interferenzfähigkeit und Welcher-Zustand-Information (Which-State-Information).

<sup>22</sup>Eine ausführliche Diskussion dieser Debatte findet man in [220].

und die Impulsunschärfe  $\delta p_x$  sehr viel kleiner als der zu messende Impulsübertrag  $\Delta p_x$  sein muß, da sonst hieraus keine Rückschlüsse darüber möglich sind, welchen Spalt ein Teilchen passiert hat.  $\Delta p_x$  wird nun abgeschätzt, indem man davon ausgeht, daß das Teilchen durch den Doppelspalt an den Ort des ersten Maximums erster Ordnung abgelenkt werden soll. Ist  $p$  der Gesamtimpuls und  $\alpha$  der zu diesem Maximum gehörende Ablenkwinkel, dann gilt einerseits  $\Delta p_x = p \sin \alpha$ , andererseits auch  $\sin \alpha = \lambda / \Delta x = h / p \Delta x$ , und zusammengenommen liefert das

$$\Delta p_x = \frac{h}{\Delta x}.$$

Wegen  $\delta p_x \ll \Delta p_x$  und  $\delta x \ll \Delta x$  folgt daraus jedoch

$$\delta p_x \delta x \ll h,$$

in eklatantem Widerspruch zur Heisenbergschen Unschärferelation

$$\delta p_x \delta x \geq \frac{\hbar}{2}.$$

Die Unschärferelation sorgt somit in dieser Situation dafür, daß der Versuch, Welcher-Weg-Information zu ergattern, zum Verschwinden des Interferenzmusters führt. Dies wurde lange Zeit für einen Sachverhalt gehalten, der in völliger Allgemeinheit gilt.

Erst zu Beginn der 90er Jahre begann klar zu werden, daß das nicht der Fall ist, sondern daß es sich bei der Heisenbergschen Unschärferelation und dem Welle-Teilchen-Dualismus um zwei grundsätzlich verschiedene Dinge handelt. Das bedeutet nicht, daß die soeben beschriebene Analyse etwa falsch wäre; in dieser Situation wird das Verschwinden des Interferenzmusters bei Beschaffung von Welcher-Weg-Information völlig korrekt mit Hilfe der Heisenbergschen Unschärferelation beschrieben. Es bedeutet jedoch, daß es Situationen gibt, in denen das Verschwinden von Interferenzmustern nicht durch Unschärfereaktionen erklärt werden kann. Die Ursache liegt vielmehr generell in Korrelationen zwischen den betrachteten Quantenobjekten und dem System, das zur Beschaffung von Welcher-Weg-Information dient. Ausgehend von zur damaligen Zeit neuartigen quanten- und atomoptischen Techniken beschrieben Scully, Englert und Walther 1991 ein Experiment mit Atomen und Mikromasern, bei dem die Unschärferelation keine Rolle spielt, aber die eigentlich auftretenden Interferenzerscheinungen verhindert werden, sobald Welche-Weg-Information gewonnen werden kann [470]<sup>23</sup>. Es deutete sich an, daß im Welle-Teilchen-Dualismus mehr steckt als einfach eine Folgeerscheinung der Unschärferelation; im Gegenteil handelt es sich hierbei um ein eigenes grundlegendes Prinzip der Quantenmechanik. Englert veröffentlichte 1996 in einem bahnbrechenden Aufsatz eine Ungleichung, mit der die Sache quantitativ beschreibbar ist [146]. Der experimentelle Nachweis dieses Sachverhalts gelang 1998 auf spektakuläre Weise Rempe, Dürr und Nonn mit Hilfe eines Atom-Interferometers, bei dem Welcher-Weg-Information ausgelesen werden kann, ohne die beteiligten Atome einem nennenswerten Impulsübertrag auszusetzen [127], [128], [129]. Damit ist klar, daß auch hier die Unschärferelation keine Rolle spielt, außerdem konnte auch Englerts Ungleichung bestätigt werden. Mehr davon im übernächsten und dem darauffolgenden Abschnitt.

<sup>23</sup>Siehe auch [149] und [151].

Offensichtlich ist das Komplementaritätsprinzip eine allgemeinere Angelegenheit als verbreitet angenommen. Das wird in folgender, von Scully, Englert und Walther eingeführten Definition des Begriffs „komplementär“ deutlich [149], [470]: *Zwei Observablen heißen komplementär, falls bei genauer Kenntnis der einen alle möglichen Meßwerte der anderen mit gleicher Wahrscheinlichkeit vorkommen.* Anders ausgedrückt: Sind zwei Meßgrößen komplementär, dann streuen die Meßwerte der einen beliebig, wenn die andere streuungsfrei gemessen wird und umgekehrt. Das Komplementaritätsprinzip lautet dann wie folgt [149], [470]<sup>24</sup>:

*Zu jedem Freiheitsgrad gibt es in der Quantenmechanik zwei dynamische Variablen, die komplementär sind.*

Daraus folgt insbesondere, daß es bei jedem quantenmechanischen System und jeder möglichen Präparierung desselben eine Meßgröße gibt, bei deren Messung das Ergebnis nicht vorhergesagt werden kann. Formal besagt das Komplementaritätsprinzip für zwei komplementäre Observablen  $A$  und  $B$  und beliebige Eigenzustände  $|\psi_A\rangle$  und  $|\psi_B\rangle$  der zugehörigen Operatoren  $\hat{A}$  und  $\hat{B}$ , daß die bedingten Wahrscheinlichkeiten, einen Eigenzustand von  $A$  anzutreffen, wenn vor der Messung von  $A$  einer von  $B$  vorliegt oder umgekehrt, also die Größen  $|\langle\psi_A|\psi_B\rangle|^2 = |\langle\psi_B|\psi_A\rangle|^2$ , unabhängig von den Eigenwerten von  $\hat{A}$  und  $\hat{B}$  sind. Man beachte daß diese Formulierungen des Komplementaritätsprinzips nichts mit der Heisenbergschen Unschärferelation zu tun haben<sup>25</sup>.

Wir betrachten einige Beispiele für komplementäre Größen [149]. Die naheliegendsten sind wenig überraschend Ort und Impuls, wie man leicht nachrechnen kann. Die zu betrachtenden Größen erhält man hier gerade aus den Ortsdarstellungen der Impulseigenzustände, also ebenen Wellen der Form

$$\langle x|p\rangle = \frac{1}{(2\pi\hbar)^{3/2}} \exp\left(\frac{ipx}{\hbar}\right),$$

und für diese gilt

$$|\langle x|p\rangle|^2 = \frac{1}{(2\pi\hbar)^3},$$

unabhängig vom Ort  $x$  und vom Impuls  $p$ . Das ist nichts anderes als die bekannte Tatsache, daß, falls bei einem Teilchen der Ort genau feststeht, das Ergebnis einer Impulsmessung nicht vorhergesagt werden kann und umgekehrt. Hintergrund ist hier natürlich die Heisenbergsche Unschärferelation, aber es muß erneut darauf hingewiesen werden, daß das kein allgemeiner Zusammenhang ist.

Weitere Beispiele liefern Spinzustände. Bekanntlich lassen sich Spin-Operatoren durch sogenannte Spinmatrizen darstellen, im Fall von Spin- $\frac{1}{2}$ -Teilchen durch die Paulischen Spinmatrizen

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

<sup>24</sup>Vergleiche auch [87]. Englert und Walther betonen in [149], daß das von ihnen formulierte Komplementaritätsprinzip von Schwinger in einer früheren Arbeit [469] in gewissem Sinne bereits vorweggenommen wurde.

<sup>25</sup>Eine Kommutatorrelation der Form  $[\hat{A}, \hat{B}] = i\hbar\mathbf{1}$ , wie sie die Grundlage der allgemeinen Unschärferelation ist, ist für die Eigenschaft von  $A$  und  $B$ , komplementär zu sein, weder notwendig noch hinreichend [149]. Beispiele für nichtkomplementäre Observablen, die kanonische Vertauschungsrelationen erfüllen, liefern [53] und [178].

Der Spin-Operator lautet damit

$$\hat{S} = \begin{pmatrix} \hat{S}_x \\ \hat{S}_y \\ \hat{S}_z \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} \hat{\sigma}_x \\ \hat{\sigma}_y \\ \hat{\sigma}_z \end{pmatrix}$$

Dabei können beispielsweise  $\hat{S}_x$  und  $\hat{S}_z$  als dynamische Variable aufgefaßt werden, da man jede komplexe  $2 \times 2$ -Matrix als Funktion von  $\hat{\sigma}_x$  und  $\hat{\sigma}_z$  darstellen kann. Eigenzustände des Operators  $\hat{S}_x$  sind die 2er-Spinoren

$$|x \uparrow\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad |x \downarrow\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

diejenigen von  $\hat{S}_z$  lauten

$$|z \uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |z \downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

Bildet man Skalarprodukte aus diesen Eigenvektoren, so erhält man

$$\langle x \uparrow | z \uparrow \rangle = \langle x \uparrow | z \downarrow \rangle = \frac{1}{\sqrt{2}}$$

und

$$\langle x \downarrow | z \uparrow \rangle = \langle x \downarrow | z \downarrow \rangle = -\frac{1}{\sqrt{2}}$$

und damit

$$|\langle x \uparrow | z \uparrow \rangle|^2 = |\langle z \uparrow | x \uparrow \rangle|^2 = \frac{1}{2},$$

$$|\langle x \uparrow | z \downarrow \rangle|^2 = |\langle z \downarrow | x \uparrow \rangle|^2 = \frac{1}{2},$$

$$|\langle x \downarrow | z \uparrow \rangle|^2 = |\langle z \uparrow | x \downarrow \rangle|^2 = \frac{1}{2},$$

$$|\langle x \downarrow | z \downarrow \rangle|^2 = |\langle z \downarrow | x \downarrow \rangle|^2 = \frac{1}{2},$$

und das ist jeweils unabhängig von den Eigenwerten der Spinoperatoren. Folglich sind die Größen  $S_x$  und  $S_z$  komplementär, womit wir auf die altbekannte Tatsache gestoßen sind, das die  $x$ - und die  $z$ -Komponente des Spins eines Spin- $\frac{1}{2}$ -Teilchens nicht gleichzeitig bestimmt sein können.

Betrachtet man stattdessen Spin-1-Teilchen, so kann man beispielsweise

$$\hat{J}_x = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}, \quad \hat{J}_y = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \hat{J}_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

als Spinmatrizen wählen und erhält dann als Spin-Operator

$$\hat{S} = \begin{pmatrix} \hat{S}_x \\ \hat{S}_y \\ \hat{S}_z \end{pmatrix} = \hbar \begin{pmatrix} \frac{1}{\sqrt{2}} \hat{J}_x \\ \frac{1}{\sqrt{2}} \hat{J}_y \\ \hat{J}_z \end{pmatrix}.$$

Wieder können  $\hat{S}_x$  und  $\hat{S}_z$  als dynamische Variable gewählt werden, denn jede komplexe  $3 \times 3$ -Matrix kann als Linearkombination der 9 Matrizen  $\hat{J}_x^k \hat{J}_z^l$  mit  $k, l = 1, 2, 3$  dargestellt werden. Die Eigenvektoren von  $\hat{S}_x$  sind

$$|x \uparrow\rangle = \frac{1}{\sqrt{3}} \begin{pmatrix} e^{2\pi i/3} \\ 1 \\ e^{-2\pi i/3} \end{pmatrix}, \quad |x \rightarrow\rangle = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad |x \downarrow\rangle = \frac{1}{\sqrt{3}} \begin{pmatrix} e^{-2\pi i/3} \\ 1 \\ e^{2\pi i/3} \end{pmatrix},$$

diejenigen von  $\hat{S}_z$

$$|z \uparrow\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad |z \rightarrow\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad |z \downarrow\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Man rechnet leicht nach, daß für sämtliche Eigenzustände

$$|\langle x \dots | z \dots \rangle|^2 = |\langle z \dots | x \dots \rangle|^2 = \frac{1}{3}$$

gilt, wieder unabhängig von den Eigenwerten der Spinoperatoren. Damit sind auch hier  $S_x$  und  $S_z$  komplementär; auch bei Spin-1-Teilchen können, wie bei allen Systemen mit beliebigem quantenmechanischem Drehimpuls, dessen  $x$ - und  $z$ -Komponente nicht gleichzeitig bestimmt sein. Entsprechendes gilt natürlich jeweils auch für die  $x$ - und  $y$ - beziehungsweise für die  $y$ - und  $z$ -Komponente. Insbesondere ist an diesen Beispielen gut erkennbar, daß es sich beim Komplementaritätsprinzip nicht wirklich um ein eigenes Naturprinzip, sondern eher um eine spezielle Formulierung einer wohlbekannten Folgerung aus dem statistischen Charakter der Quantenmechanik handelt<sup>26</sup>.

Trotz dieses sehr allgemeinen Geltungsbereichs des Komplementaritätsprinzips werden wir uns von nun an wieder den Observablen Ort und Impuls zuwenden, da diese gerade wieder zum Welle-Teilchen-Dualismus und damit zum eigentlichen Thema dieses Abschnitts zurückführen. Dazu werden wir nun gewisse merkwürdige Beobachtungen analysieren, die man bei quantenmechanischen Interferenzexperimenten machen kann. Sie werden uns im vierten Kapitel erneut begegnen.

## 4.3.2 Welcher-Weg-Information und Sichtbarkeit von Interferenzmustern

### 4.3.2.1 Interferenzexperimente mit einzelnen Quantenobjekten

Die Tatsache, daß nicht nur Photonen, sondern auch Teilchen mit Ruhemasse wie Elektronen und dergleichen unter geeigneten Bedingungen Interferenzphänomene zeigen, darf

<sup>26</sup>Das soll die Leistung von Scully, Englert und Walther keinesfalls schmälern, die in erster Linie darin besteht, einem bis dahin diffusen Begriff eine präzise Definition verschafft zu haben.



angesichts verbreiteter Präsenz solcher Sachverhalte selbst in populärwissenschaftlichen Zeitschriften und gymnasialen Lehrplänen inzwischen beinahe zur Allgemeinbildung gezählt werden. Wie wir noch sehen werden, gelingt dies inzwischen mit Atomen und sogar mit weit schwereren Teilchen wie Fluorofullerenen, und es ist nicht abzusehen, wohin das noch führt. Die Erklärung gelingt formal scheinbar mühelos durch Einführen der De Broglie-Wellenlänge  $\lambda = h/p$ : Man stellt sich vor, die Teilchen wären Wellen, die den Doppelspalt, das Gitter oder was auch immer durchlaufen, dadurch gebeugt werden und je nach Gangunterschied durch destruktive oder konstruktive Interferenz die Beobachtung von Minima und Maxima nach sich ziehen. Das alles scheint gerade jenen naiven Dualismus von Teilchen und Wellen zu bestätigen, von dem oben die Rede war.

Daß die Sache doch nicht ganz so einfach ist, sieht man, wenn man die Intensität der Teilchenquelle so weit vermindert, bis nur noch einzelne Teilchen emittiert werden, und das in so großen zeitlichen Abständen, daß sich stets nicht mehr als höchstens ein Teilchen, niemals aber mehrere Teilchen gleichzeitig im Interferometer befinden<sup>27</sup>. Damit scheint eine Überlagerung von aus verschiedenen, aber ununterscheidbaren Wegen kommenden Beiträgen und folglich die Voraussetzung für Interferenzerscheinungen nicht gegeben zu sein. In der Tat beobachtet man auf einem geeignet installierten Schirm, der die Teilchen nach Durchlaufen des Interferometers auffängt, zunächst scheinbar stochastisch auftretende Punkte, aber nach und nach baut sich genau dasselbe Interferenzmuster auf wie bei höheren Intensitäten. Die Sache läßt sich sogar noch verschärfen, wenn man gewisse mit technischem Aufwand verbundene Probleme außer Acht läßt. Führt man das Experiment nämlich mit sehr vielen identischen Interferometern und jeweils nur einem Teilchen durch und überlagert anschließend die Ergebnisse auf den Schirmen, so erhält man wieder ein Interferenzmuster. Man beachte dabei: In jedem Fall befindet sich stets nur höchstens ein Teilchen zur gleichen Zeit in der Apparatur.

Damit steht man vor dem Problem, *weder mit Wellen- noch mit Teilchenvorstellungen etwas ausrichten zu können*. Klassische Teilchen können sich eigentlich nicht überlagern, schon gar nicht, wenn jeweils nur eines davon vorhanden ist. Aber auch klassische Wellen helfen nicht weiter, da sie etwas kontinuierlich ausgesendetes sind und gerade nicht in einzelnen Portionen abgegeben werden. Ein solches Verhalten ist *weder mit Teilchen- noch mit Wellenvorstellungen* erklärbar, und entsprechend ist die These, Quantenobjekte treten je nach experimenteller Situation mal mal als Wellen, mal als Teilchen in Erscheinung, nicht haltbar. *Quantenobjekte sind weder Wellen noch Teilchen*, sie sind etwas drittes, was wir mit klassischen Bildern nicht beschreiben können. Insbesondere wird daran auch deutlich, daß die Vorstellung, der Wellen- und der Teilchencharakter würden stets sauber getrennt auftreten, nicht haltbar ist.

Die Frage, was dieses dritte denn nun genau ist, liegt natürlich nahe, wir werden jedoch sehen, daß wir beim Versuch einer Antwort auf extrem unanschauliche Bilder zurückgreifen müssen. Wollen wir diesen Bildern eine experimentell zugängliche Seite abgewinnen, landen wir gerade beim Welle-Teilchen-Dualismus in seiner richtigen, exakten Form.

---

<sup>27</sup>Das ist unabhängig von der betrachteten Teilchensorte technisch heutzutage kein Problem.



### 4.3.2.2 Photonen im Mach-Zehnder-Interferometer

Um diese exakte Form zu verstehen, werden wir nun das Verhalten von Quantenobjekten im Zusammenhang mit Interferenzphänomenen sehr genau unter die Lupe nehmen. Als Anschauungsmaterial sind dafür eigentlich Zwei-Wege-Interferometer aller Art geeignet, wie Doppelspalte, Michelson-Interferometer, Mach-Zehnder-Interferometer und dergleichen. Wir werden dafür jedoch zunächst ausschließlich das Mach-Zehnder-Interferometer betrachten, weil es was die möglichen Wege angeht besonders übersichtlich aufgebaut ist und die Welche-Weg-Information dabei auf besonders einfache Weise erzielt werden kann. Anschließend betrachten wir dann weitere denkbare und insbesondere auch weitere bereits realisierte Varianten, in denen der Welle-Teilchen-Dualismus eindrucksvoll zur Geltung kommt.

Ein Mach-Zehnder-Interferometer ist ein Zwei-Wege-Interferometer für Licht, die hier auftauchenden Quantenobjekte sind daher Photonen. Das wirft zwar eigentlich ein Problem auf, da die Beschreibung von Photonen mit Hilfe von Wellenfunktionen – und solche werden in diesem Kapitel natürlich im Mittelpunkt stehen – vorsichtig formuliert zumindest problematisch ist. Der Grund liegt in der relativistischen Natur der Photonen, die es genaugenommen unmöglich macht, für sie Wellenfunktionen im Rahmen der elementaren nichtrelativistischen Quantenmechanik aufzustellen [340]. Es gibt jedoch im Rahmen der Quantenfeldtheorie mehr als vollwertigen Ersatz hierfür, folglich soll uns das hier nicht weiter stören<sup>28</sup>.

Ein Mach-Zehnder-Interferometer besteht aus zwei halbdurchlässigen und zwei vollverspiegelten Spiegeln sowie zwei Detektoren oder Schirme<sup>29</sup>. Der erste halbdurchlässige Spiegel teilt den einlaufenden Lichtstrahl in zwei rechtwinklig auseinanderlaufende Strahlen, diese werden durch die beiden vollverspiegelten Spiegel wieder an einen Punkt zusammengeführt, und dort erzeugt der zweite halbdurchlässige Spiegel erneut zwei senkrecht auseinanderlaufende Lichtstrahlen. Diese werden durch die beiden Detektoren beziehungsweise Schirme aufgefangen. Das bedeutet, es gibt für die Photonen zwei zunächst ununterscheidbare Wege, und hinter dem zweiten Strahlteiler kommt es jeweils zur Überlagerung der auf den beiden Wegen ankommenden Lichtstrahlen. Folglich erwarten wir dort Interferenzeffekte; was auf den beiden Schirmen genau zu beobachten ist, hängt klassisch gesehen von den jeweiligen Phasendifferenzen der beiden Wege ab und in quantenmechanischer Sprache ausgedrückt von den Zuständen, in denen sich die Photonen nach Durchlaufen der einzelnen Wege befinden.

Um das genauer zu untersuchen, muß man zunächst beachten, daß gemäß den Gesetzen der Optik bei geradlinigem Durchgang durch einen Strahlteiler überhaupt kein Phasensprung, bei Reflexion an einem solchen einer um  $\pi/2$  und bei Reflexion an einem voll verspiegelten Spiegel ein Phasensprung um  $\pi$  stattfindet. Wir nehmen an daß ein einlaufendes Photon durch die normierte Wellenfunktion  $|\psi_0\rangle$  beschrieben wird; ein

<sup>28</sup>Für Leser, die sich damit nicht zufrieden geben wollen, beschreiben Kuhn und Strnad in [331] wie man eine Näherung finden kann, die für die Diskussion von Interferenzerscheinungen und damit für die Anforderungen dieses Kapitels völlig ausreicht. Geeignet gewählte Matrixelemente aus der Quantenelektrodynamik verhalten sich nämlich weitestgehend wie nichtrelativistische Wellenfunktionen; man kann damit sogar eine formale Schrödingergleichung für Photonen angeben [535]. Näheres dazu findet man auch in [420], [481] und [506] sowie speziell in Bezug auf die Behandlung von Doppelspaltexperimenten mit Photonen in Lehrbüchern in [364] und [501].

<sup>29</sup>Siehe auch Abschnitt 4.5.3, wo sich Abbildungen dazu befinden, da sie dort besser aufgehoben sind.

Phasensprung um  $\varphi$  macht sich dann durch einen Faktor  $e^{i\varphi}$  und die Reflexion an einem oder der Durchgang durch einen Strahlteiler, die mit einer Wahrscheinlichkeit von jeweils  $1/2$  stattfinden, durch einen Faktor  $1/\sqrt{2}$  bemerkbar, da  $(1/\sqrt{2})^2 = 1/2$  ist und Wahrscheinlichkeitsdichten Betragsquadrate von Wellenfunktionen sind. Entsprechend lautet die Wellenfunktion eines durch das Interferometer laufenden Photons, das an einem Strahlteiler reflektiert wurde,

$$|\psi\rangle = \frac{1}{\sqrt{2}} e^{i\pi/2} |\psi_0\rangle = \frac{i}{\sqrt{2}} |\psi_0\rangle \quad (4.11)$$

und die Wellenfunktion eines an einem Spiegel reflektierten Photons

$$|\psi\rangle = e^{i\pi} |\psi_0\rangle = -|\psi_0\rangle. \quad (4.12)$$

Damit können wir nun berechnen, was an den Schirmen zu beobachten ist. Nennen wir die Quelle  $Q$ , den Strahlteiler links oben  $S_1$ , den rechts unten  $S_4$ , den Spiegel rechts oben  $S_2$ , den links unten  $S_3$  sowie den unteren Schirm  $D_1$  und den rechten  $D_2$ , dann führen zu  $D_1$  die beiden ununterscheidbaren Wege  $\Gamma_1 = Q S_1 S_2 S_4 D_1$  und  $\Gamma_2 = Q S_1 S_3 S_4 D_1$ , zu  $D_2$  dagegen  $\Gamma_3 = Q S_1 S_2 S_4 D_2$  sowie  $\Gamma_4 = Q S_1 S_3 S_4 D_2$ . So erhalten wir aus (4.11) und (4.12) Schritt für Schritt am Schirm  $D_1$  für Photonen, die auf dem Weg  $\Gamma_1$  dorthin gelangt sind,

$$|\psi_1\rangle = \frac{1}{\sqrt{2}} e^{i\pi} \frac{1}{\sqrt{2}} |\psi_0\rangle = \frac{1}{2} e^{i\pi} |\psi_0\rangle = -\frac{1}{2} |\psi_0\rangle \quad (4.13)$$

und für solche, die über  $\Gamma_2$  gereist sind,

$$|\psi_2\rangle = \frac{1}{\sqrt{2}} e^{i\pi/2} e^{i\pi} \frac{1}{\sqrt{2}} e^{i\pi/2} |\psi_0\rangle = \frac{1}{2} e^{2i\pi} |\psi_0\rangle = \frac{1}{2} |\psi_0\rangle. \quad (4.14)$$

Diese beiden werden nun überlagert, so daß man insgesamt als Wellenfunktion

$$|\psi_{D_1}\rangle = |\psi_1\rangle + |\psi_2\rangle = 0$$

findet. Das heißt die Wahrscheinlichkeit, das Photon am Schirm  $D_1$  zu finden, ist

$$\| |\psi_{D_1}\rangle \|^2 = 0,$$

es kommt also *überhaupt nichts* bei  $D_1$  an. Man kann das auch an der Phasenverschiebung erkennen: Aus (4.13) und (4.14) erhält man dafür  $\Delta\varphi_{D_1} = \pi$ , und das ist gerade eine Bedingung für *destruktive Intefferenz*. Entsprechend erhalten wir wieder aus (4.11) und (4.12) am Schirm  $D_2$  für  $\Gamma_3$

$$|\psi_3\rangle = \frac{1}{\sqrt{2}} e^{i\pi} \frac{1}{\sqrt{2}} e^{i\pi/2} |\psi_0\rangle = -\frac{1}{2} e^{3i\pi/2} |\psi_0\rangle = -\frac{i}{2} |\psi_0\rangle \quad (4.15)$$

und für  $\Gamma_4$

$$|\psi_4\rangle = \frac{1}{\sqrt{2}} e^{i\pi/2} e^{i\pi} \frac{1}{\sqrt{2}} |\psi_0\rangle = -\frac{1}{2} e^{3i\pi/2} |\psi_0\rangle = -\frac{i}{2} |\psi_0\rangle, \quad (4.16)$$

und überlagert man diese beiden, so erhält man

$$|\psi_{D_2}\rangle = |\psi_3\rangle + |\psi_4\rangle = -i|\psi_0\rangle.$$

Man findet folglich ein Photon am Schirm  $D_2$  mit der Wahrscheinlichkeit

$$\| |\psi_{D_2}\rangle \|^2 = |-i|^2 \| |\psi_0\rangle \|^2 = 1,$$

es kommt also *immer etwas* bei  $D_2$  an. Auch das kann man durch Berechnen der Phasenverschiebung bestätigen; gemäß (4.15) und (4.16) beträgt diese hier  $\Delta\varphi_{D_2} = 0$ , eine Bedingung für *konstruktive Interferenz*. – Genaugenommen gilt das natürlich nur für die exakte optische Achse des Interferometers; da jedoch selbst ein Laserstrahl natürlich keine mathematische Gerade darstellt und eine gewisse Ausdehnung besitzt, durchläuft die Phasenverschiebung ausgehend von der optischen Achse radial nach außen kontinuierlich weitere Werte, und man erhält als Interferenzmuster Maxima und Minima in Form konzentrischer Kreise, wobei am Schirm  $D_1$  ein Minimum und am Schirm  $D_2$  ein Maximum im Zentrum auftritt.

Zunächst ist das jedoch in keiner Weise aufregend, sondern lediglich ein altbekanntes Phänomen der Wellenoptik. So richtig spannend wird die Sache auch hier wieder, wenn man die Intensität der Photonenquelle immer weiter herunterdreht, bis sich stets nicht mehr als höchstens ein Photon zur gleichen Zeit im Interferometer befindet. An obiger Rechnung ändert sich dabei überhaupt nichts, und experimentell bestätigt sich das genau: Zunächst treten wieder scheinbar zufällig einzelne Punkte auf den Schirmen auf, aber nach und nach bauen sich genau dieselben Interferenzmuster auf, wie sie im Fall höherer Intensität anzutreffen sind. Auch die Variante mit der Überlagerung der Positionen der Photonen auf den Schirmen bei sehr vielen Mach-Zehnder-Interferometern mit jeweils nur einem einzigen Photon liefert ebenfalls dieselben beiden Interferenzmuster.

Wieder gilt: Mit dem üblichen Wellenmodell ist das nicht zu erklären und mit irgendeiner Teilchenvorstellung schon gar nicht. Wenn sich irgendetwas aus beiden möglichen Wegen überlagern soll, dann sollte wohl auch etwas auf beiden Wegen daherkommen; die Möglichkeit, daß Photonen aus beiden Wegen zusammenkommen und miteinander interferieren, fällt jedoch von vorneherein flach, da niemals mehrere Photonen gleichzeitig in der Apparatur vorhanden sind. Es bleibt nur der Schluß, daß jedes Photon nur mit sich selbst interferiert und daher irgendwie gleichzeitig durch beide Wege gelangen muß, was jeder klassischen Vorstellung beliebig zuwiderläuft. Wir folgern daraus, daß Quantenobjekte weder Wellen noch Teilchen sondern etwas drittes sind. Allerdings sind Quantenobjekte eben in der Lage, Interferenzerscheinungen hervorzurufen, und sie können auch durch Detektoren lokal nachgewiesen werden. Sie zeigen also gelegentlich *wellenähnliches oder teilchenähnliches Verhalten*. Es gilt dabei jedoch kein striktes Entweder-Oder, sondern vielmehr so etwas wie ein Sowohl-Als-Auch: Diese beiden Aspekte müssen keineswegs stets sauber voneinander getrennt sein; sie können auch auf subtile Weise miteinander vermischt sein. Damit geht der Welle-Teilchen-Dualismus über das Komplementaritätsprinzip hinaus. Interessant wird es nun, wenn man dieses wellen- beziehungsweise teilchenähnliche Verhalten gezielt experimentell abfragt. Versucht man insbesondere, die Quantenobjekte auszutricksen und dazu zu verleiten, beides gleichzeitig zu tun, stellt man fest, daß sie einem immer irgendwie einen Schritt voraus sind. Das alles ist der wesentliche Inhalt des Welle-Teilchen-Dualismus, wie wir in den folgenden Abschnitten sehen werden.

### 4.3.2.3 Der Effekt der Welcher-Weg-Information auf die Ausbildung von Interferenzmustern

Wellenähnliches Verhalten läßt sich bei Quantenobjekten grundsätzlich durch die Beobachtung von Interferenzerscheinungen nachweisen; teilchenähnliches Verhalten dagegen ist mit lokalisierbarem Verhalten und damit mit dem Beschreiten eines bestimmten Weges verbunden. Möchte man daher eventuelle Zusammenhänge zwischen wellen- und teilchenähnlichem Verhalten von Quantenobjekten untersuchen, so bietet es sich an, zu versuchen, etwas über den Weg herauszufinden, den Quantenobjekte in Zwei-Wege-Interferometern einschlagen und gleichzeitig das Interferenzmuster im Auge zu behalten. Bevor wir das diskutieren, ist allerdings noch eine kurze Bemerkung zur Terminologie angebracht. Wir werden ständig von irgendwelchen Wegen reden, die Quantenobjekte angeblich einschlagen, behalten dabei aber stets im Blick, daß sie im allgemeinen das gerade nicht tun. Klassische Bahnen kommen in der Quantenmechanik nicht vor, diese Erkenntnis bestätigt sich bei allem, was mit Welle-Teilchen-Dualismus zu tun hat, erneut auf dramatische Weise. Wenn wir also Informationen über Wege von Quantenobjekten suchen, dann entweder um herauszufinden, daß das gar nicht möglich ist, oder aber, daß dadurch das System grundlegend verändert wird. Dazu gehen wir nun der Reihe nach unterschiedliche experimentelle Varianten mit dem Mach-Zehnder-Interferometer durch.

- a) Zunächst soll das Interferometer mit einzelnen Photonen ohne weitere Manipulationen zum Einsatz kommen. Wir betrachten dazu, was der quantenmechanische Formalismus ganz allgemein bei der Beschreibung von Zwei-Wege-Interferometern sagt. Den Photonen<sup>30</sup> bieten sich im Interferometer zwei unterschiedliche, aber für sie zunächst ununterscheidbare Wege; entsprechend bezeichnen wir den Zustand „Photon verfolgt Weg 1“ mit  $\psi_1$  und den Zustand „Photon verfolgt Weg 2“ mit  $\psi_2$ . Der allgemeinste Zustand, den ein Photon nach Durchlaufen der Apparatur aufweist, ist aufgrund der Ununterscheidbarkeit der Wege folglich

$$\psi = \frac{1}{\sqrt{2}} (\psi_1 + \psi_2); \quad (4.17)$$

der Faktor  $1/\sqrt{2}$  dient zur Normierung, und vom Unterschied der Interferenzmuster auf den beiden Schirmen des Mach-Zehnder-Interferometers sehen wir ab jetzt ab. Was man auf den Schirmen sieht, wird durch die Aufenthaltswahrscheinlichkeitsdichte der Teilchenzustände bestimmt, also vom Betragsquadrat der Wellenfunktion. Für diese erhält man

$$\begin{aligned} |\psi|^2 &= \left| \frac{1}{\sqrt{2}} (\psi_1 + \psi_2) \right|^2 \\ &= \frac{1}{2} (|\psi_1|^2 + |\psi_2|^2 + \psi_1^* \psi_2 + \psi_2^* \psi_1) \\ &= \frac{1}{2} [|\psi_1|^2 + |\psi_2|^2 + 2 \Re (\psi_1^* \psi_2)], \end{aligned} \quad (4.18)$$

<sup>30</sup>Unsere Betrachtung ist so allgemein gehalten, daß man sich anstelle von Photonen beliebige andere Quantenobjekte und anstelle von Mach-Zehnder-Interferometern beliebige andere Zwei-Wege-Interferometer denken kann.

einen Ausdruck, der die Summe der Aufenthaltswahrscheinlichkeitsdichten der Einzelwellenfunktionen und damit die Beiträge der beiden einzelnen Wege enthält und darüberhinaus einen zusätzlichen Term, der das Interferenzverhalten beschreibt und konsequenterweise *Interferenzterm* heißt. Genau das Auftreten eines solchen Terms signalisiert das Auftreten eines Interferenzmusters.

- b) Jetzt sehen wir nach, was passiert, wenn man versucht, etwas über die Wege herauszufinden. Das geht natürlich am einfachsten auf die grobe Art, indem man einen der beiden Wege versperrt. Denn dann weiß man mit Gewißheit, daß die Photonen den anderen Weg genommen haben müssen, sofern auf dem Schirm etwas ankommt. Damit besteht in beiden möglichen Fällen die Wellenfunktion (4.17) nur noch aus einem Summanden, und in der Wahrscheinlichkeitsdichte (4.18) verschwindet insbesondere der Interferenzterm. Das Versperren eines der beiden Wege führt erwartungsgemäß zum Verschwinden des Interferenzmusters. Damit hat man anscheinend wieder die ausschließende Entweder-Oder-Situation der traditionellen Auffassung des Welle-Teilchen-Dualismus. Wie wir gesehen haben, wird das dem gesamten Sachverhalt nicht gerecht; wir werden also subtiler vorgehen müssen.
- c) Man müßte den Photonen weiterhin beide Wege offenhalten und ihnen dennoch nach Durchlaufen des Interferometers ansehen können, welchen der beiden Wege sie genommen haben, oder anders gesagt, man muß dafür sorgen, daß die Photonen unterschiedlich markiert sind, je nachdem auf welchem Weg sie durchmarschieren. Das ist die Grundidee jeglicher Welcher-Weg-Messung in der Interferometrie. Im Fall von Photonen gelingt dies am einfachsten durch Polarisierung. Man bringt beispielsweise jeweils einen Polarisationsfilter mit drehbarer Polarisierungsebene in die beiden Arme eines Mach-Zehnder-Interferometers, und wenn man diese unterschiedlich orientiert, so kann man an der Polarisierung der Photonen, die das Interferometer durchlaufen haben, ablesen, auf welchem Weg sie das getan haben. Denn nach Durchlaufen der Polarisationsfilter weisen die Photonen, die vorher keine festgelegte Polarisierung haben, gerade die an ersteren eingestellte Polarisierung auf<sup>31</sup>. Mit anderen Worten: Die beiden bisher ununterscheidbaren Wege sind unterscheidbar geworden.

Das Photon hat nun externe und interne Freiheitsgrade; die externen werden hier weiterhin durch die beiden möglichen Zustände  $\psi_1$  und  $\psi_2$  repräsentiert, die internen dagegen durch den Winkel  $\alpha$  der Polarisierungsebenen, gemessen zu irgendeiner Vorzugsrichtung, etwa parallel zum Labortisch. Die entsprechenden normierten Zustände seien  $|\alpha\rangle$ . Man kann zeigen, daß jede Polarisierung, lineare, zirkulare und sonstige, als Linearkombination aus zwei orthogonalen linearen Polarisierungsrichtungen dargestellt werden kann, beispielsweise waagrechte und senkrechte lineare Polarisierung mit Zustandsvektoren  $|\rightarrow\rangle$  und  $|\uparrow\rangle$ ; entsprechend gilt für diese  $\langle\rightarrow|\rightarrow\rangle = \langle\uparrow|\uparrow\rangle = 1$  sowie  $\langle\rightarrow|\uparrow\rangle = \langle\uparrow|\rightarrow\rangle = 0$ . Sind die beiden Polarisatoren auf die Winkel  $\alpha_1$  beziehungsweise  $\alpha_2$  eingestellt, dann befinden sich die

---

<sup>31</sup>Die aus der Wellenoptik bekannte anschauliche Vorstellung von Polarisierung als Festlegung der Schwingungsebene der sich ausbreitenden Lichtwellen funktioniert bei einzelnen Photonen natürlich nicht. Dennoch läßt sich die Polarisierung auch bei einzelnen Photonen definieren; das hängt mit dem Spin der Photonen zusammen, gehört aber in den Bereich der relativistischen Quantenmechanik. Mehr darüber findet man in [331].

Photonen nach Durchlaufen des Interferometers im Zustand

$$\psi = \frac{1}{\sqrt{2}} (\psi_1 |\alpha_1\rangle + \psi_2 |\alpha_2\rangle); \quad (4.19)$$

das bedeutet, daß sie mit den Polarisatoren korreliert sind.

Stellt man zunächst versuchsweise beide Polarisatoren *parallel*, also auf denselben Winkel ein, sollte eigentlich nichts passieren<sup>32</sup>, denn die beiden Wege bleiben ununterscheidbar. Und in der Tat liefert die Wellenfunktion

$$\psi = \frac{1}{\sqrt{2}} (\psi_1 |\alpha\rangle + \psi_2 |\alpha\rangle)$$

als Aufenthaltswahrscheinlichkeitsdichte

$$\begin{aligned} |\psi|^2 &= \frac{1}{2} [ |\psi_1|^2 \langle \alpha | \alpha \rangle + |\psi_2|^2 \langle \alpha | \alpha \rangle + 2 \Re (\psi_1^* \psi_2 \langle \alpha | \alpha \rangle) ] \\ &= \frac{1}{2} [ |\psi_1|^2 + |\psi_2|^2 + 2 \Re (\psi_1^* \psi_2) ], \end{aligned}$$

mit unverändert vorhandenem Interferenzterm. Bei parallel ausgerichteten Polarisatoren erhält man ein Interferenzmuster.

- d) Jetzt verwenden wir die Polarisatoren wie geplant zum Markieren der Wege und stellen sie *senkrecht zueinander* ein; ohne Beschränkung der Allgemeinheit wählen wir dazu waagrechte und senkrechte Polarisation<sup>33</sup>. Die beiden Wege sind jetzt unterscheidbar: Nach Passieren des Interferometers könnte man den Photonen nun ansehen, welchen der beiden Wege sie genommen haben, indem man ihre Polarisation mißt, denn Messung von waagrechter Polarisation bedeutet mit Gewißheit, daß es Weg 1 war, und Messung von senkrechter Polarisation bedeutet mit derselben Gewißheit Weg 2. Dabei werden beide Wege mit gleicher Wahrscheinlichkeit vertreten sein. Diese Situation wird durch folgende Wellenfunktion zum Ausdruck gebracht:

$$\psi = \frac{1}{\sqrt{2}} (\psi_1 | \rightarrow \rangle + \psi_2 | \uparrow \rangle). \quad (4.20)$$

Sie bringt wieder zum Ausdruck, daß in dieser experimentellen Anordnung die Photonen mit den Polarisatoren korreliert sind. Berechnen der Aufenthaltswahrscheinlichkeitsdichte liefert

$$\begin{aligned} |\psi|^2 &= \frac{1}{2} ( |\psi_1|^2 \langle \rightarrow | \rightarrow \rangle + |\psi_2|^2 \langle \uparrow | \uparrow \rangle + \psi_1^* \psi_2 \langle \rightarrow | \uparrow \rangle + \psi_2^* \psi_1 \langle \uparrow | \rightarrow \rangle ) \\ &= \frac{1}{2} ( |\psi_1|^2 + |\psi_2|^2 ); \end{aligned} \quad (4.21)$$

wegen  $\langle \rightarrow | \uparrow \rangle = 0$  verschwindet jetzt der Interferenzterm. Bei gekreuzten Polarisatoren findet man kein Interferenzmuster auf den Schirmen. Wie (4.21) zeigt

<sup>32</sup>Es bleibt in dieser sehr elementaren Betrachtung natürlich unberücksichtigt, daß die Gesamtintensität nachläßt, aber das spielt für Aspekte der Interferenzfähigkeit auch keine Rolle.

<sup>33</sup>Man kann die Basis des Polarisationsraumes beliebig wählen, sodaß die Basisvektoren  $| \rightarrow \rangle$  und  $| \uparrow \rangle$  jedes beliebige Paar zueinander senkrechter Polarisationsrichtungen repräsentieren können.



– der Ausdruck enthält nur die Summe der Aufenthaltswahrscheinlichkeitsdichten für das Passieren jeweils der einzelnen Wege – entsteht nur eine stochastische Verteilung von Punkten. Ursache für das Verschwinden der Interferenzfähigkeit sind, wie die Wellenfunktion (4.20) zeigt, die Korrelationen zwischen dem Photon und den Polarisatoren, die zur Ermittlung der Welcher-Weg-Information zum Einsatz kommen.

- e) Die Welcher-Weg-Information wurde in Beispiel d) durch unterschiedliche Polarisation der Photonen zugänglich, je nachdem auf welchem Weg sie das Interferometer passieren. Was passiert jedoch, wenn man diese Information nach Durchlaufen der Apparatur wieder unkenntlich macht? Dazu setzen wir je einen Polarisationsfilter in Weg 1 und Weg 2, die Richtungen gekreuzt. Wie gesehen ergibt sich so kein Interferenzmuster. Jetzt fügen wir einen dritten Polarisationsfilter *nach* dem zweiten Strahlteiler hinzu, und zwar mit einer Ausrichtung von jeweils  $45^\circ$  zu den beiden anderen. Dieser Polarisator wird von den Photonen erst passiert, nachdem sie das Interferometer und damit beide Wege vollständig durchlaufen haben. Eine Polarisationsrichtung von  $45^\circ$  läßt sich als Linearkombination aus waagrechter und senkrechter Polarisation ausdrücken; weil sowohl waagrecht als auch senkrecht polarisierte Photonen einen solchen Polarisator mit 50 %-iger Wahrscheinlichkeit passieren und mit 50 %-iger Wahrscheinlichkeit von ihm absorbiert werden, lautet diese  $|\nearrow\rangle = \frac{1}{\sqrt{2}} (|\rightarrow\rangle + |\uparrow\rangle)$  und die Wellenfunktion der Photonen nach Durchgang durch diesen dritten Polarisator folglich

$$\begin{aligned}\psi &= \frac{1}{\sqrt{2}} (\psi_1 |\rightarrow\rangle + \psi_2 |\uparrow\rangle) |\nearrow\rangle \\ &= \frac{1}{2} (\psi_1 |\rightarrow\rangle + \psi_2 |\uparrow\rangle) (|\rightarrow\rangle + |\uparrow\rangle) \\ &= \frac{1}{2} [\psi_1 (|\rightarrow\rangle|\rightarrow\rangle + |\rightarrow\rangle|\uparrow\rangle) + \psi_2 (|\uparrow\rangle|\rightarrow\rangle + |\uparrow\rangle|\uparrow\rangle)].\end{aligned}$$

Wieder berechnen wir die zugehörige Aufenthaltswahrscheinlichkeitsdichte; das ist lästig, aber aufschlußreich und führt auf

$$\begin{aligned}|\psi|^2 &= \frac{1}{4} [|\psi_1|^2 (\langle\rightarrow|\langle\rightarrow|\rightarrow\rangle|\rightarrow\rangle + \langle\rightarrow|\langle\rightarrow|\rightarrow\rangle|\uparrow\rangle \\ &\quad + \langle\rightarrow|\langle\uparrow|\rightarrow\rangle|\rightarrow\rangle + \langle\uparrow|\langle\rightarrow|\rightarrow\rangle|\rightarrow\rangle) \\ &\quad + |\psi_2|^2 (\langle\uparrow|\langle\rightarrow|\uparrow\rangle|\rightarrow\rangle + \langle\uparrow|\langle\rightarrow|\uparrow\rangle|\uparrow\rangle \\ &\quad + \langle\uparrow|\langle\uparrow|\uparrow\rangle|\rightarrow\rangle + \langle\uparrow|\langle\uparrow|\uparrow\rangle|\uparrow\rangle) \\ &\quad + \psi_1^* \psi_2 (\langle\rightarrow|\langle\rightarrow|\uparrow\rangle|\uparrow\rangle + \langle\rightarrow|\langle\rightarrow|\uparrow\rangle|\uparrow\rangle \\ &\quad + \langle\rightarrow|\langle\uparrow|\uparrow\rangle|\rightarrow\rangle + \langle\rightarrow|\langle\uparrow|\uparrow\rangle|\uparrow\rangle) \\ &\quad + \psi_2^* \psi_1 (\langle\uparrow|\langle\rightarrow|\rightarrow\rangle|\rightarrow\rangle + \langle\uparrow|\langle\rightarrow|\rightarrow\rangle|\uparrow\rangle \\ &\quad + \langle\uparrow|\langle\uparrow|\rightarrow\rangle|\rightarrow\rangle + \langle\uparrow|\langle\uparrow|\rightarrow\rangle|\uparrow\rangle)] \\ &= \frac{1}{4} [|\psi_1|^2 + |\psi_2|^2 + 2 \Re(\psi_1^* \psi_2)].\end{aligned}$$



Der Interferenzterm taucht wieder auf, und entsprechend wird das Interferenzmuster wieder sichtbar. Und das, obwohl der gesamte Vorgang, der zuerst zur Ausbildung des Interferenzmusters und danach zur Auslöschung desselben geführt hat, zum Zeitpunkt, an dem die Photonen durch den dritten Polarisator hindurchgehen, *bereits abgeschlossen ist*. Das Interferenzmuster kann durch *nachträgliches Auslöschen der Welcher-Weg-Information* wieder hergestellt werden. Eine solche Anordnung nennt man *Quantenradierer*<sup>34</sup>.

An dieser Stelle ist eine Zwischenbilanz angebracht. Bei den Versuchen a) und c) sind die beiden möglichen Wege durch das Interferometer jeweils ununterscheidbar, das heißt, in beiden Fällen ist keine Welcher-Weg-Information verfügbar. Bei Versuch d) dagegen sind die Photonen nach Durchlaufen des Interferometers unterschiedlich markiert, je nachdem auf welchem der beiden Wege sie gekommen sind. Die beiden Wege sind damit unterscheidbar, und es besteht die Möglichkeit, Welcher-Weg-Information zu ermitteln. Man erhält nur dann ein Interferenzmuster, wenn die beiden Wege ununterscheidbar sind. Sobald die Wege unterscheidbar sind, verschwindet das Interferenzmuster. Damit bestätigt sich die Entweder-Oder-Fassung des Prinzips vom Welle-Teilchen-Dualismus: *Bei einem Zwei-Wege-Interferometer ist man entweder in der Lage, Welcher-Weg-Information zu erhalten, oder man findet ein Interferenzmuster*.

Dabei spielt ein wie auch immer gearteter Meßvorgang zum Einholen der Welcher-Weg-Information und damit irgendwelche vermeintliche Störungen oder Eingriffe in das System keinerlei Rolle. Das wird im Versuch e) deutlich; dort macht der dritte Polarisationsfilter die Welcher-Weg-Information nachträglich wieder unkenntlich, und das vormals verschwundene Interferenzmuster taucht wieder auf. *Entscheidend ist die Möglichkeit, Welcher-Weg-Information zu bekommen, nicht aber, ob diese überhaupt eingeholt wurde*.

Darüberhinaus läßt sich aus Versuchen mit Zwei-Wege-Interferometern und einzelnen Quantenobjekten, wie sie oben exemplarisch beschrieben wurden, eine fundamentale Folgerung hinsichtlich des Zustandekommens von Interferenzmustern ziehen: *Jedes Quantenobjekt interferiert jeweils mit sich selbst*. Es befindet sich in irgendeiner nicht-lokalen Weise in beiden Armen des Interferometers gleichzeitig und ist so in der Lage, Interferenzerscheinungen auszulösen, wenn der Vorgang sehr häufig stattfindet. Diese Erkenntnis zieht allerdings unmittelbar einige beunruhigende Fragen nach sich. Wie kann das Photon gleichzeitig beide Wege durchlaufen? Wie kann ein Photon, das sich anscheinend in einem Weg des Interferometers befindet, wissen, ob der andere Weg offen ist oder nicht? Wie kann ein Photon, das anscheinend einen Polarisationsfilter durchläuft, wissen, wie der andere Polarisationsfilter eingestellt ist? Und nicht zuletzt: Wie kann ein Photon wissen ob es beobachtet werden kann oder nicht? Teilen sich die Photonen etwa auf? Dann könnte je eine Hälfte durch beide Wege eilen, und diese Hälften könnten dann interferieren. Doch auch dieser sowieso nicht ganz ernst gemeinte Vorschlag läßt sich auf experimentellem Weg leicht ausschließen. Führt man nämlich Koinzidenzzählungen durch, das heißt, bringt man auf beiden Wegen des Interferometers Detektoren an, die gekoppelt sind, so stellt man fest, daß zwar stets einer der beiden Detektoren anspricht, niemals aber beide Detektoren gleichzeitig. *Wenn man nachsieht*, befindet sich das Photon durchweg *entweder* auf Weg 1 *oder* auf Weg 2.

<sup>34</sup>Der englische Ausdruck dafür ist *Quantum Eraser*.

Wie man es auch dreht und wendet, man kommt nicht um die Tatsache herum, daß interferenzfähige Photonen in Zwei-Wege-Interferometern sich in Superpositionen aus Zuständen befinden, in denen sie die Apparatur auf dem einen und auf dem anderen möglichen Weg durchlaufen, genau wie es durch Wellenfunktionen der Form (4.17) beschrieben wird. Etwas bodenständiger ausgedrückt: Die Photonen nehmen in nicht-lokaler Weise beide Wege gleichzeitig, um durch das Interferometer zu gelangen. Die oben erwähnte Form des Welle-Teilchen-Dualismus ist nichts anderes als eine kombinierte Folgerung aus dieser Tatsache und dem quantenmechanischen Meßpostulat, wonach sich Systeme nach Messung einer Observablen in Eigenzuständen dieser Observablen befinden. Daß dabei die Möglichkeit einer solchen Messung bereits ausreicht, die Interferenzfähigkeit zu zerstören und die Messung selbst gar nicht durchgeführt werden muß, darf als Hinweis darauf gelten, daß bei allem, was mit dem Problem des quantenmechanischen Meßprozesses zu tun hat, das Wissen über das betrachtete System oder das menschliche Bewußtsein oder sonstige anthropozentrische Vorstellungen ganz gewiß keine Rolle spielen.

Was damit jedoch nach wie vor fehlt, ist die Überwindung der traditionellen Entweder-Oder-Form des Welle-Teilchen-Dualismus. Diese wird erst bei einer weiteren, zuerst von Englert untersuchten Variante von Experimenten mit Zwei-Wege-Interferometern manifest, die wir uns im folgenden Abschnitt ansehen werden.

#### 4.3.2.4 Exakte Formulierung des Prinzips des Welle-Teilchen-Dualismus

Versuche wie die oben beschriebenen lassen sich insgesamt weder mit einem Teilchenbild noch mit einem Wellenbild erklären. Quantenobjekte sind weder Teilchen noch Wellen, sondern irgendetwas anderes, das wir uns nicht vorstellen können. Sie verhalten sich manchmal so ähnlich wie Teilchen und manchmal so ähnlich wie Wellen. Dabei wird wellenähnliches Verhalten typischerweise durch die Fähigkeit repräsentiert, Interferenzerscheinungen auszulösen und teilchenähnliches Verhalten durch die Möglichkeit, Welcher-Weg-Information gewinnen zu können. Das ist spektakulär, aber nichts wirklich neues.

Wirklich neu war jedoch Mitte der 90er Jahre die Erkenntnis, daß der Übergang zwischen wellen- und teilchenähnlichem Verhalten fließend ist. Man kann das wieder exemplarisch am Mach-Zehnder-Interferometer sehen, indem man untersucht was passiert, wenn man die beiden Polarisatoren zur Ermöglichung von Welcher-Weg-Information auf beliebige Winkel einstellt. Wir nehmen dazu an, der Polarisator im Weg 1 sei auf den Winkel  $\alpha_1$  und derjenige im Weg 2 auf  $\alpha_2$  eingestellt, schreiben beide Polarisationszustände als Linearkombinationen aus waagrechter und senkrechter Polarisation und erhalten so für die Wellenfunktion von Photonen, die durch dieses Interferometer gelaufen sind, den Ausdruck

$$\begin{aligned} \psi &= \frac{1}{\sqrt{2}} (\psi_1 |\alpha_1\rangle + \psi_2 |\alpha_2\rangle) \\ &= \frac{1}{\sqrt{2}} [\psi_1 (a_1 | \rightarrow \rangle + b_1 | \uparrow \rangle) + \psi_2 (a_2 | \rightarrow \rangle + b_2 | \uparrow \rangle)]. \end{aligned} \quad (4.22)$$

Wieder liegt eine Korrelation zwischen den Photonen und den Polarisatoren vor. Für die entsprechende Aufenthaltswahrscheinlichkeitsdichte findet man

$$\begin{aligned}
|\psi|^2 &= \frac{1}{2} [ |\psi_1|^2 (|a_1|^2 \langle \rightarrow | \rightarrow \rangle + a_1^* b_1 \langle \rightarrow | \uparrow \rangle + b_1^* a_1 \langle \uparrow | \rightarrow \rangle + |b_1|^2 \langle \uparrow | \uparrow \rangle) \\
&\quad + \psi_1^* \psi_2 (a_1^* a_2 \langle \rightarrow | \rightarrow \rangle + a_1^* b_2 \langle \rightarrow | \uparrow \rangle + b_1^* a_2 \langle \uparrow | \rightarrow \rangle + a_1^* a_2 \langle \uparrow | \uparrow \rangle) \\
&\quad + \psi_2^* \psi_1 (a_2^* a_1 \langle \rightarrow | \rightarrow \rangle + a_2^* b_1 \langle \rightarrow | \uparrow \rangle + b_2^* a_1 \langle \uparrow | \rightarrow \rangle + b_2^* b_1 \langle \uparrow | \uparrow \rangle) \\
&\quad + |\psi_2|^2 (|a_2|^2 \langle \rightarrow | \rightarrow \rangle + a_2^* b_2 \langle \rightarrow | \uparrow \rangle + b_2^* a_2 \langle \uparrow | \rightarrow \rangle + |b_2|^2 \langle \uparrow | \uparrow \rangle) ] \\
&= \frac{1}{2} \{ |\psi_1|^2 (|a_1|^2 + |b_1|^2) + |\psi_2|^2 (|a_2|^2 + |b_2|^2) + 2 \Re [\psi_1^* \psi_2 (a_1^* a_2 + b_2^* b_1)] \}
\end{aligned}$$

und damit einen Ausdruck mit Interferenzterm, der aber natürlich von den Koeffizienten  $a_1, a_2, b_1, b_2$  abhängt. Die bisher diskutierten Szenarien mit zwei Polarisatoren sind hierin als Spezialfälle enthalten:  $a_1 = a_2$  und  $b_1 = b_2$  liefert Versuch c) mit maximaler Interferenzfähigkeit, aber keinerlei Möglichkeit, Welcher-Weg-Information herauszufinden,  $a_1 = b_2 = 1$  und  $a_2 = b_1 = 0$  führen auf Versuch d) mit der Möglichkeit, vollständige Welcher-Weg-Information zu erlangen, aber für den Preis eines nicht vorhandenen Interferenzmusters. Zusätzlich sind nun alle möglichen Zwischenstufen denkbar. Das sehen wir uns nun etwas allgemeiner an. Die Wellenfunktion (4.22) ist von der Form

$$\psi = \frac{1}{\sqrt{2}} (\psi_1 |\varphi_1\rangle + \psi_2 |\varphi_2\rangle),$$

wobei die Zustandsvektoren  $|\varphi_1\rangle$  und  $|\varphi_2\rangle$  die Systeme beschreiben, die zur Beschaffung von Welcher-Weg-Information dienen, also beispielsweise die Polarisatoren beim Mach-Zehnder-Interferometer<sup>35</sup>. Das liefert nun als Aufenthaltswahrscheinlichkeitsdichte den Ausdruck

$$|\psi|^2 = \frac{1}{2} [ |\psi_1|^2 + |\psi_2|^2 + 2 \Re (\psi_1^* \psi_2 \langle \varphi_1 | \varphi_2 \rangle) ]$$

mit einem Interferenzterm, der vom Skalarprodukt  $\langle \varphi_1 | \varphi_2 \rangle$  der beiden Weg-Markierungszustände abhängt. Der Interferenzterm kann jeden Wert von 0 (keine Interferenzfähigkeit) bis  $\Re (\psi_1^* \psi_2)$  (maximale Interferenzfähigkeit) annehmen. Dabei ist  $\langle \varphi_1 | \varphi_2 \rangle$  umso größer, je ähnlicher sich die beiden Zustände  $|\varphi_1\rangle$  und  $|\varphi_2\rangle$  sind und um so kleiner, je weiter sie auseinanderliegen (je orthogonaler sie sind, um es etwas volkstümlich auszudrücken). Je ähnlicher sich diese Zustände sind, desto weniger ist ihre Messung jedoch zum Unterscheiden der beiden möglichen Wege des Interferometers geeignet. Die Möglichkeit der Gewinnung besserer Welcher-Weg-Information bedeutet folglich geringere Sichtbarkeit des Interferenzmusters und umgekehrt. Damit kann das allgemeine Prinzip des Welle-Teilchen-Dualismus wie folgt formuliert werden:

*Je besser sich verschiedene mögliche Wege von identischen Quantenobjekten unterscheiden lassen, desto geringer ist ihre Interferenzfähigkeit; je größer ihre Interferenzfähigkeit ist, desto weniger gut lassen sich ihre möglichen Wege unterscheiden.*

<sup>35</sup>Die Beschränkung auf reine Zustände ist hier zwar wenig realistisch, das spielt jedoch für die betrachteten Zusammenhänge keine Rolle.

Auch wenn das hier betrachtete Modell stark elementarisiert ist, gilt die daraus gezogene Folgerung in völliger Allgemeinheit. Hier liegt nun eine echte Erweiterung des Komplementaritätsprinzips und damit auch der traditionellen Auffassung vor. Dort stehen sich stets nur die beiden Extremfälle gegenüber, und von irgendwelchen Zwischenstufen ist keine Rede. Die allgemeine Form des Welle-Teilchen-Dualismus bringt jedoch klar zum Ausdruck, daß zwischen diesen Extremfällen der vollständigen Welcher-Weg-Information und der idealen Interferenzfähigkeit in kontinuierlicher Weise alle Mischformen realisiert werden können. Sie wird damit der Sowohl-Als-Auch-Natur gerecht, die Quantenobjekte hinsichtlich ihrer wellen- und teilchenähnlichen Eigenschaften aufweisen. Die Betonung liegt allerdings mit Nachdruck auf *ähnlich*: Quantenobjekte haben nichts klassisches an sich, schon gar nicht klassisches Wellen- oder Teilchenverhalten. Die Ursache liegt hier wie stets in der Existenz interferenzfähiger Superpositionen; diese äußern sich in den durch den Welle-Teilchen-Dualismus beschriebenen Situationen in Korrelationen zwischen den beteiligten Quantenobjekten und den Systemen, die zur Ermöglichung der Beschaffung von Welcher-Weg-Information über diese Quantenobjekte herangezogen werden. Diese Korrelationen sorgen für das Verschwinden der Interferenzfähigkeit der Quantenobjekte. Wir haben hier genau besehen spezielle Fälle von Dekohärenzeffekten vorliegen. Bei der Dekohärenz besorgt die Wechselwirkung des betrachteten Systems mit der Umgebung natürlich auch nichts anderes als die Bereitstellung von Welcher-Weg-Information, da hierdurch fortwährend Ortsmessungen stattfinden, und dadurch werden Interferenzeffekte un beobachtbar. Beim Welle-Teilchen-Dualismus interessiert man sich für die Wechselwirkung mit einem speziellen Teil der Umgebung, nämlich mit dem System, das in der betrachteten experimentellen Anordnung für die Welcher-Weg-Information zuständig ist; vom Rest der Umgebung ist die Interferometer üblicherweise weitgehend abgeschirmt, um weitere Dekohärenzeffekte zu verhindern.

Wir weisen auch an dieser Stelle nachdrücklich darauf hin, daß es nicht darauf ankommt, ob die Welcher-Weg-Information tatsächlich eingeholt wird. Die reine Möglichkeit, die Wege unterscheiden zu können, genügt bereits dafür, die Interferenzfähigkeit geringer werden zu lassen, selbst wenn von dieser Möglichkeit keinerlei Gebrauch gemacht wird. Im übrigen läßt sich das Prinzip weiter verallgemeinern. Oben wurde von idealisierten Bedingungen ausgegangen, weil wir für die Markierungsgeräte am Anfang der Experimente durchgehend *reine Zustände* angenommen haben. Das ist gleichbedeutend mit vollständiger Kenntnis des quantenmechanischen Zustandes dieser Markierer. Da das in realen Situationen üblicherweise nicht der Fall ist, handelt es sich bei den Zuständen, die die Markierer beschreiben, im allgemeinen um *gemischte Zustände*. Damit ist Welcher-Weg-Information von vorneherein nur eingeschränkt möglich, und diese zusätzliche Einschränkung ist unabhängig von irgendwelchen Interferenzfähigkeiten. Es ist also durchaus möglich, daß bei gewissen Versuchsanordnungen sowohl die Interferenzfähigkeit der beteiligten Quantenobjekte als auch die Unterscheidbarkeit ihrer möglichen Wege gering ausgeprägt bis gar nicht vorhanden sind. Es kann lediglich nicht beides gleichzeitig in beliebig hohem Maß auftreten. Die Konsequenz davon ist, *daß die Umkehrung von obigem Dualismusprinzip falsch ist*. Weder folgt aus geringerer Unterscheidbarkeit der Wege automatisch bessere Interferenzfähigkeit, noch bedeutet geringere Interferenzfähigkeit zwangsläufig, daß die Wege besser unterscheidbar sind. Eine derartige Umkehrung des Welle-Teilchen-Dualismus ist nur richtig, wenn die Anfangszustände der Markierungssysteme für die Wege reine Zustände sind.

Einige weitere Bemerkungen sind wohl angebracht. Zunächst muß man unbedingt bedenken, daß, wenn hier von möglichen Wegen die Rede ist, damit keine klassischen Bahnen gemeint sind, die die Teilchen etwa unbemerkt beschreiten. Erst wenn man nachschaut, ob sie auf einem solchen Weg sind, findet man sie dort auch; das ist nichts anderes als ein Beispiel für einen quantenmechanischen Meßprozeß.

Gelegentlich findet man auch folgende Formulierung: Je besser sich verschiedene mögliche Wege von identischen Quantenobjekten unterscheiden lassen, desto geringer ist die Sichtbarkeit von Interferenzerscheinungen. Hierbei ist jedoch Vorsicht angebracht, da in dieser Sprechweise nicht in der anderen Richtung von geringerer oder fehlender Sichtbarkeit von Interferenzmustern auf bessere Unterscheidbarkeit der Wege geschlossen werden darf. Denn selbst absolute Ununterscheidbarkeit der Wege führt bei speziellen Versuchsanordnungen im Falle destruktiver Interferenz nicht zum Auftreten von Interferenzerscheinungen.

Erneut betont werden sollte auch, daß der Welle-Teilchen-Dualismus nichts mit der Heisenbergschen Unschärferelation zu tun hat. Er ist ein eigenes, fundamentaleres Prinzip, das allen Quantenobjekten grundsätzlich zu eigen ist und letztlich aus dem quantenmechanischen Superpositionsprinzip folgt. Die zusammengesetzten Zustände aus Quantenobjekt und Wegmarkierungssystem, die uns oben begegnet sind, sind Musterbeispiele für sogenannte *verschränkte Zustände*, denen wir später einen eigenen Abschnitt widmen und die im Zentrum aller Merkwürdigkeiten stehen, welche die Quantenmechanik zu bieten hat. Sie sind ihrerseits aber genau deshalb möglich, weil in der Quantenmechanik das Superpositionsprinzip gilt. Das ist wieder einmal ein Beleg für die These, daß der hauptsächlichste Unterschied zwischen klassischer Physik und Quantenmechanik in der Existenz interferenzfähiger Superpositionen in letzterer besteht.

### 4.3.3 Englerts Dualitätsrelation

Besonders faszinierend am Welle-Teilchen-Dualismus ist, daß sich dieses oben rein qualitativ formulierte Prinzip quantifizieren läßt, wie B.-G. Englert 1996 zeigen konnte. Im Zentrum dieser quantitativen Beschreibung steht eine Ungleichung, die genau das zum Ausdruck bringt, was über die traditionelle Entweder-Oder-Aussage hinausgeht; ihre Bedeutung wird jedoch erst dadurch so richtig klar, daß man sie in formaler Strenge herleiten kann. Englert führte das in seiner grundlegenden Arbeit [146] von 1996 explizit vor<sup>36</sup>. Auch wenn das eine etwas technische Angelegenheit ist, werden wir der Bedeutung der Sache dennoch wohl nur gerecht, wenn wir die Herleitung der Dualitäts-Relation in allen Einzelheiten nachvollziehen. Da das etwas längere Rechnungen erforderlich macht, verteilen wir diese auf mehrere Abschnitte.

#### 4.3.3.1 Sichtbarkeit von Interferenzmustern

Zunächst definieren wir eine Größe zur quantitativen Erfassung der Fähigkeit eines Quantenobjekts, Interferenzmuster auszubilden. Wir starten dazu wieder mit der Betrachtung eines beliebigen symmetrischen Zwei-Wege-Interferometers, das von Quantenobjekten

<sup>36</sup>Vergleiche auch [147]. Dieses Resultat darf als eine der bedeutendsten physikalischen Entdeckungen der Neunziger Jahre betrachtet werden. Erste vage Ansätze wurden 1979 von Wootters und Zurek [530] veröffentlicht.

durchlaufen wird; diese sind beschreibbar durch einen Zustandsvektor der Form

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|\psi_1\rangle + |\psi_2\rangle).$$

Der Dichteoperator des Quantenobjekts lautet damit

$$\rho_{Q,0} = |\psi\rangle\langle\psi| = \frac{1}{2} (|\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2| + |\psi_2\rangle\langle\psi_1| + |\psi_1\rangle\langle\psi_2|);$$

wie üblich signalisieren die Außerdiagonalelemente das Vorhandensein von Interferenzphänomenen. Zusätzlich soll der Versuchsaufbau ein Gerät zur Markierung der Wege enthalten, also einen Welcher-Weg-Detektor. Der Dichteoperator des Zustands, in den dieser Markierer anfänglich präpariert ist, sei  $\rho_{M,0}$ . Damit erreichen wir weitgehende Allgemeinheit, da die Beschränkung auf reine Zustände des Markierengerätes, die oben gemacht wurde, dadurch fallengelassen wird und sich dieses auch in gemischten Zuständen befinden darf. Die Ermöglichung von Welcher-Weg-Information besteht in einer Wechselwirkung des Markierers mit dem Quantenobjekt; diese ist beschreibbar durch unitäre Operatoren  $\hat{U}_1$  und  $\hat{U}_2$ , einen für jeden der beiden Wege. Für Weg 1 bedeutet das den Übergang

$$\rho_{M,0} \longrightarrow \hat{U}_1^+ \rho_{M,0} \hat{U}_1 \equiv \rho_{M,1}$$

und für Weg 2

$$\rho_{M,0} \longrightarrow \hat{U}_2^+ \rho_{M,0} \hat{U}_2 \equiv \rho_{M,2}.$$

Die Wechselwirkung des Quantenobjekts mit dem Markierer verursacht eine Verschränkung dieser beiden Systeme; der Dichteoperator des verschränkten Systems aus Quantenobjekt und Markierer ist

$$\hat{P} = \frac{1}{2} (|\psi_1\rangle\langle\psi_1| \rho_{M,1} + |\psi_2\rangle\langle\psi_2| \rho_{M,2} + |\psi_2\rangle\langle\psi_1| \tilde{\rho}_M + |\psi_1\rangle\langle\psi_2| \tilde{\rho}_M^+) \quad (4.23)$$

mit  $\tilde{\rho}_M \equiv \hat{U}_2^+ \rho_{M,0} \hat{U}_1$ . Durch die Bildung der entsprechenden Teilspuren bei (4.23) findet man den Dichteoperator des Markierers

$$\rho_M = \text{tr}_Q \hat{P} = \frac{1}{2} (\rho_{M,1} + \rho_{M,2})$$

und den Dichteoperator des Quantenobjekts

$$\rho_Q = \text{tr}_M \hat{P} = \frac{1}{2} (|\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2| + C |\psi_2\rangle\langle\psi_1| + C^* |\psi_1\rangle\langle\psi_2|),$$

mit der Abkürzung  $C = \text{tr}_M \tilde{\rho}_M = \text{tr}_M (\hat{U}_2^+ \rho_{M,0} \hat{U}_1)$ ; diese komplexe Zahl steht zusammen mit ihrer komplex konjugierten Zahl als Vorfaktor vor den Außerdiagonalelementen des Dichteoperators und kann folglich als *Kontrast-Faktor* für die Sichtbarkeit des Interferenzmusters interpretiert werden. Entsprechend seiner Definition ist dieser Kontrast-Faktor eine charakteristische Größe des Markierers.



Ein Interferenzmuster wird üblicherweise durch die Phasendifferenzen bestimmt, die bei Superpositionen zwischen den beteiligten Kanälen auftreten; ist die Phasendifferenz etwa  $\phi$ , dann lautet die entsprechende Superposition

$$|\psi(\phi)\rangle = \frac{1}{\sqrt{2}} (|\psi_1\rangle + e^{i\phi} |\psi_2\rangle).$$

Der zugehörige Dichteoperator ist

$$|\psi(\phi)\rangle \langle\psi(\phi)| = \frac{1}{2} (|\psi_1\rangle \langle\psi_1| + |\psi_2\rangle \langle\psi_2| + e^{i\phi} |\psi_2\rangle \langle\psi_1| + e^{-i\phi} |\psi_1\rangle \langle\psi_2|).$$

Zur Beschreibung des Interferenzmusters berechnen wir die Wahrscheinlichkeit  $p(\phi)$ , den Zustand  $|\psi(\phi)\rangle$  vorzufinden und erhalten

$$\begin{aligned} p(\phi) &= \text{tr}_Q (|\psi(\phi)\rangle \langle\psi(\phi)| \rho_Q) \\ &= \frac{1}{4} \text{tr}_Q (|\psi_1\rangle \langle\psi_1| + C^* |\psi_1\rangle \langle\psi_2| + |\psi_2\rangle \langle\psi_2| + C |\psi_2\rangle \langle\psi_1| \\ &\quad + e^{i\phi} |\psi_2\rangle \langle\psi_1| + e^{i\phi} C^* |\psi_2\rangle \langle\psi_2| + e^{-i\phi} |\psi_1\rangle \langle\psi_2| + e^{-i\phi} C |\psi_1\rangle \langle\psi_1|) \\ &= \frac{1}{4} (2 + e^{-i\phi} C + e^{i\phi} C^*) \\ &= \frac{1}{2} [1 + \Re(e^{i\phi} C^*)], \end{aligned} \tag{4.24}$$

also einen Ausdruck, der Werte im Bereich von

$$p_{\min} = \frac{1}{2} (1 - |C|)$$

bis

$$p_{\max} = \frac{1}{2} (1 + |C|)$$

annehmen kann. Definiert man nun die *Sichtbarkeit der Interferenzmuster*<sup>37</sup>  $\mathcal{V}$  gemäß

$$\mathcal{V} = \frac{p_{\max} - p_{\min}}{p_{\max} + p_{\min}},$$

so folgt durch Einsetzen unmittelbar

$$\mathcal{V} = |C|,$$

ein Ergebnis, das angesichts obiger Interpretation der Größe  $C$  einen sehr anschaulichen Charakter besitzt.  $\mathcal{V}$  ist ein Maß für die Fähigkeit der betrachteten Quantenobjekte, bei der verwendeten experimentellen Anordnung Interferenzerscheinungen zu verursachen.

---

<sup>37</sup>Englisch *Fringe Visibility*.



### 4.3.3.2 Unterscheidbarkeit der Wege

Als nächstes definieren wir eine Größe, mit der die Unterscheidbarkeit der möglichen Wege eines Quantenobjekts quantitativ beschrieben werden kann. Zur Unterscheidbarkeit der beiden möglichen Wege, die das Quantenobjekt in einem Zwei-Wege-Interferometer nehmen kann, muß irgendeine Messung vorgenommen werden, deren Meßresultat Schlußfolgerungen über den eingeschlagenen Weg zuläßt. Dazu betrachten wir einen Operator  $\hat{W}$ , der eine Observable zum Auslesen der Welcher-Weg-Information repräsentiert, mit Eigenwerten  $W$  und Eigenzuständen  $|W\rangle$ . Die möglichen Meßergebnisse  $W$  einer solchen Messung sollen Aufschluß über den Weg des Quantenobjekts liefern. Die Wahrscheinlichkeit  $P(W)$  des Auftretens von  $W$  bei Messung dieser Observablen ist

$$P(W) = \langle W|\rho_M|W\rangle = \frac{1}{2} \langle W|\rho_{M,1}|W\rangle + \frac{1}{2} \langle W|\rho_{M,2}|W\rangle \quad (4.25)$$

Einholen von Welcher-Weg-Information besteht nun darin, aus der Verteilung der Meßresultate bei Messungen der durch  $\hat{W}$  repräsentierten Observablen den Weg herauszufinden, den das Quantenobjekt wohl genommen hat. Das kann typisch quantenmechanisch im allgemeinen nicht mit Sicherheit herausgefunden werden, außer in dem trivialen Fall, daß auf einem der beiden Wege überhaupt nichts ankommt. Man wird aber wohl dazu neigen, jenen Weg als den richtigen anzunehmen, zu dem derjenige der beiden Summanden in (4.25) gehört, der am meisten zu  $P(W)$  beiträgt [530]. Damit ist die Wahrscheinlichkeit  $\mathcal{L}_W$ , den Weg richtig zu erraten, gegeben durch

$$\mathcal{L}_W = \sum_W \text{Max} \left\{ \frac{1}{2} \langle W|\rho_{M,1}|W\rangle, \frac{1}{2} \langle W|\rho_{M,2}|W\rangle \right\}$$

Dabei gilt ersichtlicherweise  $1/2 \leq \mathcal{L}_W \leq 1$ , was auch der Intuition entspricht, denn im ungünstigsten Fall rät man ins Blaue hinein mit einer Erfolgsaussicht von 50 %, im günstigsten Fall weiß man Bescheid und rät mit Erfolgchancen von 100 %. Daher definiert man sinnvollerweise die *Kenntnis*  $\mathcal{K}_W$  der Wege gemäß

$$\mathcal{K}_W = 2\mathcal{L}_W - 1;$$

für diese gilt dann

$$0 \leq \mathcal{K}_W \leq 1.$$

Unter Verwendung der Relation

$$\max\{x, y\} = \frac{1}{2}(x + y) + \frac{1}{2}|x - y|$$

erhält man dafür

$$\begin{aligned} \mathcal{K}_W &= \frac{1}{2} \sum_W \langle W|\rho_{M,1} + \rho_{M,2}|W\rangle + \frac{1}{2} \sum_W |\langle W|\rho_{M,1} - \rho_{M,2}|W\rangle| - 1 \\ &= \sum_W \langle W|\rho_M|W\rangle + \frac{1}{2} \sum_W |\langle W|\rho_{M,1} - \rho_{M,2}|W\rangle| - 1 \end{aligned}$$

$$= \frac{1}{2} \sum_W |\langle W | \rho_{M,1} - \rho_{M,2} | W \rangle|. \quad (4.26)$$

Je größer die Kenntnis der Wege ist, desto besser wird es gelingen, sie erfolgreich zu unterscheiden. Das führt unmittelbar zur Definition der *Unterscheidbarkeit der Wege*<sup>38</sup>  $\mathcal{D}$  in der Form

$$\mathcal{D} \equiv \max_W \mathcal{K}_W.$$

Diese gibt nicht die aktuell erreichte, sondern die prinzipiell *maximal mögliche* Unterscheidung der Wege an, womit der Tatsache Rechnung getragen wird, daß die Möglichkeit des Einholens von Welcher-Weg-Information bereits ausreicht, um die Interferenzfähigkeit des betrachteten Systems zu untergraben. Genauer gesagt ist  $\mathcal{K}_W$  die maximal mögliche Information, die man durch Messung der Observablen  $W$  über die beiden Wege erreichen kann; dagegen ist  $\mathcal{D}$  die von der Natur vorgegebene maximale Information über die Wege, die im System steckt. Daher gilt für diese beiden Größen

$$\mathcal{K}_W \leq \mathcal{D}$$

und folglich auch

$$\mathcal{L}_W \leq \frac{1}{2} (1 + \mathcal{D}).$$

Wie nahe diese Ungleichungen Gleichungen kommen, das heißt wie weitreichend die Kenntnis der Wege wird, hängt natürlich von einer möglichst geschickten Wahl der Observablen  $W$  ab. Diese ist optimal, wenn die Eigenzustände  $|W\rangle$  von  $\hat{W}$  auch Eigenzustände von  $|\rho_{M,1} - \rho_{M,2}|$  sind; dann hat man sogar

$$\mathcal{K}_W = \mathcal{D}$$

und

$$\mathcal{L}_W = \frac{1}{2} (1 + \mathcal{D})$$

und kann (4.26) direkt auswerten, da die Summe dort zu einer Summe über die Eigenwerte wird und damit zur Spur des Operators  $|\rho_{M,1} - \rho_{M,2}|$ . Für die Unterscheidbarkeit der Wege gilt daher

$$\mathcal{D} = \frac{1}{2} \text{tr}_M (|\rho_{M,1} - \rho_{M,2}|). \quad (4.27)$$

Ist der Markierer in einem reinen Zustand, gilt also für dessen Dichteoperator anfangs etwa

$$\rho_{M,0} = |M_0\rangle \langle M_0|,$$

dann folgt daraus

$$\rho_{M,1} = |M_1\rangle \langle M_1| \quad \text{mit} \quad |M_1\rangle = \hat{U}_1^+ |M_0\rangle, \quad (4.28)$$

$$\rho_{M,2} = |M_2\rangle \langle M_2| \quad \text{mit} \quad |M_2\rangle = \hat{U}_2^+ |M_0\rangle, \quad (4.29)$$

$$\tilde{\rho}_M = |M_2\rangle \langle M_1|, \quad (4.30)$$

<sup>38</sup>Englisch *Distinguishability*

und damit weiter

$$\rho_{M,1} - \rho_{M,2} = |M_1\rangle \langle M_1| - |M_2\rangle \langle M_2|.$$

Dieser Operator hat die nichtverschwindenden Eigenwerte  $\pm \sqrt{1 - |\langle M_1|M_2\rangle|^2}$ , und da die Spur eines Operators gleich der Summe von seinen Eigenwerten ist, liefert (4.27) damit

$$\mathrm{tr}_M (|\rho_{M,1} - \rho_{M,2}|) = \sqrt{1 - |\langle M_1|M_2\rangle|^2}. \quad (4.31)$$

Davon werden wir gleich Gebrauch machen.

### 4.3.3.3 Beweis der Dualitätsrelation

Um nun für  $\mathcal{V}$  und  $\mathcal{D}$  eine Relation herzuleiten, betrachten wir wieder allgemeine Zustände des Markierergerätes und die Spektraldarstellung von dessen Dichteoperator  $\rho_{M,0}$ ,

$$\rho_{M,0} = \sum_k m_k |M_{k,0}\rangle \langle M_{k,0}|$$

mit Koeffizienten  $m_k \geq 0$ , die zusätzlich der Relation  $\sum_k m_k = 1$  genügen, sowie Eigenvektoren  $|M_{k,0}\rangle$ , die orthonormiert sind, es gelte also  $\langle M_{j,0}|M_{k,0}\rangle = \delta_{jk}$ . Daraus folgt für  $\rho_{M,1}$ ,  $\rho_{M,2}$  und  $\tilde{\rho}_M$

$$\rho_{M,1} = \sum_k m_k |M_{k,1}\rangle \langle M_{k,1}| \quad \text{mit} \quad |M_{k,1}\rangle = \hat{U}_1^+ |M_{k,0}\rangle, \quad (4.32)$$

$$\rho_{M,2} = \sum_k m_k |M_{k,2}\rangle \langle M_{k,2}| \quad \text{mit} \quad |M_{k,2}\rangle = \hat{U}_2^+ |M_{k,0}\rangle, \quad (4.33)$$

$$\tilde{\rho}_M = \sum_k m_k |M_{k,2}\rangle \langle M_{k,1}|. \quad (4.34)$$

(4.32) und (4.33) liefern nun

$$\rho_{M,1} - \rho_{M,2} = \sum_k m_k (|M_{k,1}\rangle \langle M_{k,1}| - |M_{k,2}\rangle \langle M_{k,2}|),$$

und damit wird (4.27) zu

$$\mathcal{D} = \frac{1}{2} \mathrm{tr}_M \sum_k m_k (|M_{k,1}\rangle \langle M_{k,1}| - |M_{k,2}\rangle \langle M_{k,2}|).$$

Mit Hilfe der Dreiecksungleichung

$$\mathrm{tr} (|A - B|) \leq \mathrm{tr} (|A|) + \mathrm{tr} (|B|)$$

erhält man daraus

$$\mathcal{D} \leq \frac{1}{2} \sum_k m_k \mathrm{tr}_M (|M_{k,1}\rangle \langle M_{k,1}| + |M_{k,2}\rangle \langle M_{k,2}|)$$

und wegen (4.31) weiter

$$\mathcal{D} \leq \sum_k m_k \sqrt{1 - |\langle M_{k,1} | M_{k,2} \rangle|^2}. \quad (4.35)$$

Außerdem liefert (4.34) für die Sichtbarkeit der Interferenzmuster

$$\mathcal{V} = |\operatorname{tr}_M \tilde{\rho}_M| = \left| \sum_k m_k \langle M_{k,1} | M_{k,2} \rangle \right|. \quad (4.36)$$

Die  $|M_{k,1}\rangle$  und  $|M_{k,2}\rangle$  sind nicht notwendigerweise zueinander orthogonal, da sie aus  $|M_0\rangle$  zwar durch unitäre Transformationen entstehen, aber jeweils nicht beide aus derselben; dennoch liegen die Beträge von Skalarprodukten der Form  $\langle M_{k,1} | M_{k,2} \rangle$  natürlich im Intervall  $[-1, 1]$ . Man kann solche Skalarprodukte deshalb in der Form

$$\langle M_{k,1} | M_{k,2} \rangle = \sin \theta_k e^{i\phi_k}$$

als Funktion zweier Winkel  $\theta_k$  und  $\phi_k$  schreiben, mit  $0 \leq \theta_k \leq \pi/2$  und  $0 \leq \phi_k \leq 2\pi$ . Einsetzen in (4.35) und (4.36) führt auf

$$\mathcal{D} \leq \sum_k m_k \cos \theta_k$$

sowie

$$\mathcal{V} = \left| \sum_k m_k \sin \theta_k e^{i\phi_k} \right|,$$

und mit etwas Trigonometrie erhält man damit

$$\mathcal{D}^2 + \mathcal{V}^2 \leq \sum_j \sum_k m_j m_k [\cos \theta_j \cos \theta_k + \sin \theta_j \sin \theta_k \cos(\phi_j - \phi_k)] \quad (4.37)$$

Für den Ausdruck in den eckigen Klammern gilt

$$\begin{aligned} & \cos \theta_j \cos \theta_k + \sin \theta_j \sin \theta_k \cos(\phi_j - \phi_k) \\ &= \sin \theta_j \cos \phi_j \sin \theta_k \cos \phi_k + \sin \theta_j \sin \phi_j \sin \theta_k \sin \phi_k + \cos \theta_j \cos \theta_k \\ &= \begin{pmatrix} \sin \theta_j \cos \phi_j \\ \sin \theta_j \sin \phi_j \\ \cos \phi_j \end{pmatrix} \cdot \begin{pmatrix} \sin \theta_k \cos \phi_k \\ \sin \theta_k \sin \phi_k \\ \cos \phi_k \end{pmatrix}; \end{aligned}$$

dieser Ausdruck hat genau die Gestalt eines Skalarprodukts aus zwei Einheitsvektoren in räumlichen Polarkoordinaten, und das bedeutet

$$\cos \theta_j \cos \theta_k + \sin \theta_j \sin \theta_k \cos(\phi_j - \phi_k) \leq 1.$$

Eingesetzt in (4.37) liefert das

$$\mathcal{D}^2 + \mathcal{V}^2 \leq \left( \sum_k m_k \right)^2 = (\operatorname{tr}_M \rho_{M,0})^2 = 1$$

und damit das Ergebnis

$$\mathcal{D}^2 + \mathcal{V}^2 \leq 1. \quad (4.38)$$

Das ist eine für beliebige zweikanalige Interferenzerscheinungen gültige fundamentale *Dualitäts-Relation* und damit die *quantitative Fassung des Welle-Teilchen-Dualismus*. Sie gibt die beschriebenen Situationen genau wieder:

- Je besser die verschiedenen möglichen Wege eines Quantenobjekts unterscheidbar sind (je größer  $\mathcal{D}$ ), desto geringer ist dessen Interferenzfähigkeit (desto kleiner  $\mathcal{V}$ ).
- Je besser die Interferenzfähigkeit eines Quantenobjekts ist (je größer  $\mathcal{V}$ ), desto geringer ist die Unterscheidbarkeit der Wege, die es genommen haben könnte.
- Maximale Unterscheidbarkeit der Wege ( $\mathcal{D} = 1$ ) bewirkt verschwindende Interferenzfähigkeit ( $\mathcal{V} = 0$ ).
- Maximale Interferenzfähigkeit ( $\mathcal{V} = 1$ ) bewirkt fehlende Unterscheidbarkeit der Wege ( $\mathcal{D} = 0$ ).

Wesentliche Bedeutung kommt dabei dem Kleiner-Gleich-Zeichen anstelle eines Gleichheitszeichens in (4.38) zu, denn darin zeigt sich die im allgemeinen nicht gegebene Umkehrbarkeit des Welle-Teilchen-Dualismus. Falls nämlich  $\mathcal{D}$  klein ist, folgt daraus keineswegs, daß  $\mathcal{V}$  automatisch groß ist, und ein kleiner Wert für  $\mathcal{V}$  zieht ebenso wenig zwangsläufig einen großen Wert für  $\mathcal{D}$  nach sich. Das gilt nur, wenn sich die Markierungssysteme zu Anfang in reinen Zuständen befinden. Aus (4.27) und (4.31) folgt für diesen Fall

$$\mathcal{D} = \sqrt{1 - |\langle M_1 | M_2 \rangle|^2}. \quad (4.39)$$

und aus (4.30)

$$\mathcal{V} = |\text{tr}_M \tilde{\rho}_M| = |\langle M_1 | M_2 \rangle|.$$

Dies liefert unmittelbar

$$\mathcal{D}^2 + \mathcal{V}^2 = 1, \quad (4.40)$$

das heißt, aus der Ungleichung wird im Falle eines Markierers mit reinem Anfangszustand eine Gleichung. Da eine realistische Interpretation der Quantenmechanik davon ausgehen muß, daß sich jedes physikalische System eigentlich in einem reinen Zustand befindet und wir diesen im Normalfall einfach nicht kennen, ist diese, durch Gleichung (4.40) repräsentierte Fassung des quantitativen Welle-Teilchen-Dualismus die fundamentalste Version desselben. Sie ist allerdings von metaphysischem Charakter, da reale experimentelle Situationen stets durch die entsprechende Ungleichung beschrieben werden. Es sollte jedoch nicht vergessen werden, daß der Versuch, fundamentale Eigenschaften der Natur zu beschreiben, stets genauso im physikalischen Auge behalten werden sollte wie das bloße Voraussagen experimenteller Sachverhalte, weswegen uns die beschriebene idealisierte Fassung des Welle-Teilchen-Dualismus mindestens genauso interessieren sollte.

Die Dualitätsrelation weist auf den ersten Blick eine gewisse Ähnlichkeit zur Heisenbergschen Unschärferrelation auf. Die beiden haben jedoch nichts miteinander zu tun, was erneut nichts anderes besagt, als daß der Welle-Teilchen-Dualismus ein von der Unschärferrelation unabhängiges fundamentaleres Prinzip der Quantenmechanik darstellt.

Fundamentaler deshalb, weil der Welle-Teilchen-Dualismus das Verschwinden der Interferenzfähigkeit von Quantenobjekten in völliger Allgemeinheit erklären kann, insbesondere auch in Situationen, in denen die Heisenbergsche Unschärferelation es nicht erklären kann.

Die Unabhängigkeit dieser beiden Gesetzmäßigkeiten voneinander wird insbesondere an deren Herleitungen deutlich. Diese erfolgt im Falle von Unschärferelationen mit Hilfe des Kommutators der beiden betrachteten Observablen. Die Herleitung der Dualitätsrelation funktioniert völlig anders; es tauchen dabei keinerlei Kommutatoren auf und insbesondere auch nur eine Observable, nämlich diejenige des Auslesens der Welcher-Weg-Information aus dem Markierer-System. Das zeigt daß die Dualitätsrelation und damit das Prinzip des Welle-Teilchen-Dualismus logisch unabhängig von jeglichen Unschärferelationen ist. Das Verschwinden der Interferenzfähigkeit von Quantenobjekten aufgrund der Möglichkeit der Gewinnung von Welcher-Weg-Information hat nichts mit Unschärferelationen zu tun, sondern mit Korrelationen aufgrund der Verschränkung dieser Quantenobjekte mit den die Welcher-Weg-Information ermöglichenden Markierer-Systemen.

#### 4.3.3.4 Verallgemeinerungen des Formalismus

Die beschriebene Herleitung der Dualitätsrelation ist für den Fall symmetrischer Zwei-Wege-Interferometer ausgelegt. Das sind solche bei denen die Wahrscheinlichkeit, daß die Quantenobjekte einen der beiden Wege nehmen, für beide Wege gleich und damit gleich  $1/2$  ist<sup>39</sup>. Die Verallgemeinerung auf asymmetrische Interferometer ist jedoch ohne weiteres möglich<sup>40</sup>. Wir nehmen an, daß die Zustände, welche die Quantenobjekte im einen oder im anderen Weg des Interferometers beschreiben, mit unterschiedlicher Gewichtung in deren Gesamtzustand auftauchen. Entsprechend ist die Wellenfunktion der Quantenobjekte nun von der Form

$$|\psi\rangle = c_1 |\psi_1\rangle + c_2 |\psi_2\rangle.$$

wobei die Wahrscheinlichkeiten, die Teilchen auf Weg 1 oder Weg 2 vorzufinden, durch  $w_1 = |c_1|^2$  beziehungsweise  $w_2 = |c_2|^2$  gegeben sind, und um zusätzlich Phasenverschiebungen auf den beiden Wegen zu berücksichtigen schreibt man wieder

$$|\psi(\phi)\rangle = c_1 |\psi_1\rangle + c_2 e^{i\phi} |\psi_2\rangle.$$

Analog zu (4.24) findet man damit für die Wahrscheinlichkeitsverteilung der Phasenverschiebung

$$\begin{aligned} p(\phi) = \operatorname{tr}_Q & \left( |c_1|^4 |\psi_1\rangle \langle \psi_1| + |c_1|^2 c_1^* c_2 C^* |\psi_1\rangle \langle \psi_2| \right. \\ & \left. + |c_2|^4 |\psi_2\rangle \langle \psi_2| + |c_2|^2 c_2^* c_1 C |\psi_2\rangle \langle \psi_1| \right) \end{aligned}$$

<sup>39</sup>Das ist eine zwar übliche, aber genaugenommen schlampige Ausdrucksweise, da sich die Quantenobjekte in einem Zustand der Überlagerung des Durchlaufens beider Wege befinden. Präziser gesprochen müßte es heißen: Wenn man versucht, nachzuschauen, welchen Weg die Quantenobjekte nehmen, findet man beide Wege jeweils mit der Wahrscheinlichkeit  $1/2$ ; die Eigenschaft, einen der beiden Wege genommen zu haben, nehmen die Quantenobjekte aber erst im Moment einer solchen Messung an.

<sup>40</sup>Entsprechende experimentelle Anordnungen werden beispielsweise in [359] und [466] beschrieben.

$$\begin{aligned}
& + |c_1|^2 c_2^* c_1 e^{i\phi} |\psi_2\rangle \langle \psi_1| + |c_1|^2 |c_2|^2 e^{i\phi} C^* |\psi_2\rangle \langle \psi_2| \\
& + |c_2|^2 c_1^* c_2 e^{-i\phi} |\psi_1\rangle \langle \psi_2| + |c_1|^2 |c_2|^2 e^{-i\phi} C |\psi_1\rangle \langle \psi_1| \\
= & |c_1|^4 + |c_2|^4 + |c_1|^2 |c_2|^2 (e^{i\phi} C^* + e^{-i\phi} C) \\
= & |c_1|^4 + |c_2|^4 + 2 |c_1|^2 |c_2|^2 \Re (e^{i\phi} C^*) \\
= & w_1^2 + w_2^2 + 2 w_1 w_2 \Re (e^{i\phi} C^*),
\end{aligned}$$

und weil dieser Ausdruck Werte im Bereich von

$$p_{\min} = w_1^2 + w_2^2 - 2 w_1 w_2 |C|$$

bis

$$p_{\max} = w_1^2 + w_2^2 + 2 w_1 w_2 |C| \quad (4.41)$$

annehmen kann, erhält man für die Sichtbarkeit der Interferenzmuster

$$\mathcal{V} = \frac{2 w_1 w_2 |C|}{w_1^2 + w_2^2}$$

Entsprechend findet man nun bei Messung der Welcher-Weg-Observablen für deren Wahrscheinlichkeitsverteilung in Analogie zu (4.25)

$$P(W) = w_1 \langle W | \rho_{M,1} | W \rangle + w_2 \langle W | \rho_{M,2} | W \rangle$$

und damit für die Unterscheidbarkeit der Wege

$$\mathcal{D} = \text{tr}_M (|w_1 \rho_{M,1} - w_2 \rho_{M,2}|).$$

Die Dualitätsrelation

$$\mathcal{D}^2 + \mathcal{V}^2 \leq 1$$

bleibt unverändert gültig.

Um den Vergleich zum symmetrischen Fall ziehen zu können, betrachten wir zunächst die Situation, die entsteht, wenn man den Welcher-Weg-Detektor ausschaltet oder gleich ganz abmontiert. Im ersten Fall bedeutet das für die unitären Operatoren, die dessen Wirkung beschreiben,  $\hat{U}_1 = \hat{U}_2$ , im zweiten sogar  $\hat{U}_1 = \hat{U}_2 = \mathbf{1}$ , und so oder so wird  $C = 1$ . Für symmetrische Interferometer hat das natürlich  $\mathcal{V} = 1$  und  $\mathcal{D} = 0$  zur Folge. Für asymmetrische Interferometer ist das jedoch nicht der Fall; es bleiben eine Sichtbarkeit der Interferenzmuster

$$\mathcal{V}_0 = \frac{2 w_1 w_2}{w_1^2 + w_2^2}$$

und eine Unterscheidbarkeit der Wege

$$\mathcal{P} = |w_1 - w_2|$$

übrig, die sich sehr anschaulich interpretieren lassen, und zwar auch dann, wenn man den Welcher-Weg-Detektor wieder einschaltet.  $\mathcal{V}_0$  ist die *a priori-Sichtbarkeit der Interferenzmuster*, die vorliegt, wenn das Interferometer ganz normal ohne irgendwelche Welcher-Weg-Detektoren in Betrieb ist; es gilt

$$\mathcal{V} = |C| \mathcal{V}_0, \quad (4.42)$$



das heißt die a priori-Sichtbarkeit wird um den Faktor  $|C|$  korrigiert, sobald ein solcher Detektor zum Einsatz kommt und folglich das Gewinnen von Welcher-Weg-Information ermöglicht wird. Es gilt  $\mathcal{V}_0 \leq \mathcal{V}$ , da  $\mathcal{V}$  wegen  $0 \leq |C| \leq 1$  das Optimum der Sichtbarkeit darstellt.  $\mathcal{P}$  ist die *Vorhersagbarkeit der Wege*, die allein aufgrund der Asymmetrie des Interferometers vorliegt, unabhängig davon, ob ein Welcher-Weg-Detektor eingeschaltet ist oder nicht, da man schon aufgrund des Aufbaus des Interferometers über Informationen bezüglich der Wege verfügt. Natürlich ist hier auch der Spezialfall symmetrischer Interferometer mit erfaßt; dann ist  $w_1 = w_2 = 1/2$  und folglich  $\mathcal{V}_0 = \mathcal{V}$  sowie  $\mathcal{P} = 0$ , denn hier gibt es nichts vorauszusagen, was die Wege angeht.

Für die a priori-Sichtbarkeit der Interferenzmuster und die Vorhersagbarkeit der Wege kann die Relation

$$\mathcal{P}^2 + \mathcal{V}_0^2 \leq 1 \quad (4.43)$$

bewiesen werden, eine Ungleichung, die 1988 von Greenberger und YaSin entdeckt wurde [201]<sup>41</sup>. Das sieht zwar formal genau gleich aus wie die Dualitätsrelation (4.38), die darin steckende physikalische Aussage ist jedoch eine gänzlich andere, und zwar gleich aus zwei Gründen. Während  $\mathcal{V}$  und  $\mathcal{V}_0$  in der beschriebenen Weise zusammenhängen, sind  $\mathcal{D}$  und  $\mathcal{P}$  zwei sehr unterschiedliche Größen. Zwar machen beide Aussagen über Welcher-Weg-Information, aber während die Unterscheidbarkeit  $\mathcal{D}$  die insgesamt maximal mögliche Kenntnis der Wege beschreibt, ist die durch die Vorhersagbarkeit  $\mathcal{P}$  beschriebene Kenntnis nur auf die Asymmetrie und damit den Aufbau des Interferometers zurückzuführen. Außerdem bezieht sich (4.43) auf quantenmechanische Eigenschaften des Anfangszustandes der Quantenobjekte, während (4.38) Aussagen über quantenmechanische Eigenschaften des Welcher-Weg-Detektors oder genauer gesagt über Korrelationen desselben mit den Quantenobjekten macht.

Man sieht daran, daß sich die Unterscheidbarkeit  $\mathcal{D}$  der Wege aus zwei sehr unterschiedlichen Anteilen zusammensetzt. Der eine beruht auf der genauen experimentellen Anordnung, das heißt dem Aufbau des Interferometers und der Quelle der Quantenobjekte oder anders gesagt deren Anfangszustand, und resultiert in einer Asymmetrie des Interferometers; dieser Anteil wird gerade durch die Vorhersagbarkeit  $\mathcal{P}$  beschrieben. Der andere entsteht durch Wechselwirkung der Quantenobjekte mit einem Welcher-Weg-Detektor und damit durch eine Verschränkung der ersteren mit dem letzteren, wodurch die Zustände, die das Durchlaufen des einen oder anderen Weges durch die Quantenobjekte repräsentieren, mit bestimmten Zuständen des Detektors korreliert werden und damit die Beschaffung von Welcher-Weg-Information ermöglicht wird. Für den Fall symmetrischer Interferometer macht dieser zweite Anteil allein die gesamte Unterscheidbarkeit aus. Martinez-Linares und Harmin nehmen das zum Anlaß, die *Qualität eines Welcher-Weg-Detektors* zu definieren gemäß [358]

$$\mathcal{Q} = \frac{1}{2} \operatorname{tr}_M (|\rho_{M,1} - \rho_{M,2}|).$$

Für symmetrische Interferometer gilt natürlich  $\mathcal{Q} = \mathcal{D}$ , im allgemeinen Fall jedoch

$$\mathcal{D} = \operatorname{Max} \{ \mathcal{P}, \mathcal{Q} \}.$$

<sup>41</sup> Auch dieses Resultat findet seinen Ausgangspunkt in der bereits erwähnten Arbeit [530] von Wootters und Zurek, direkter als bei Englert, wie wir gleich sehen werden. Vergleiche dazu auch [29], [43] und [356].

Hierfür läßt sich nun eine weitere Dualitätsrelation herleiten [358]. Dazu notiert man zunächst, daß aus der ursprünglichen Dualitätsrelation (4.38) unmittelbar

$$Q^2 + \mathcal{V}^2 \leq 1$$

folgt; setzt man hierin (4.42) und (4.43) ein, so erhält man

$$(1 - \mathcal{P}^2) Q^2 + \mathcal{P}^2 + \mathcal{V}^2 \leq 1,$$

was man auch in der Form

$$\mathcal{V}^2 \leq (1 - \mathcal{P}^2) (1 - Q^2)$$

schreiben kann<sup>42</sup>. Dies ist eine alternative Möglichkeit, den Welle-Teilchen-Dualismus mit einer Ungleichung zu quantifizieren, wobei hier die Details deutlich werden. Die Extremfälle lauten hier:

- Aus  $\mathcal{V} = 1$  folgt  $\mathcal{D} = \mathcal{P} = Q = 0$ ,
- aus  $\mathcal{P} = 1$  folgt  $\mathcal{V} = 0$ ,
- aus  $\mathcal{D} = 1$  folgt  $\mathcal{V} = 0$ ,
- aus  $Q = 1$  folgt  $\mathcal{V} = 0$ .

Das beschreibt präzise die in 4.3.2.3 beschriebenen Szenarien.

Das Prinzip des Welle-Teilchen-Dualismus in seiner qualitativen Form gilt in völliger Allgemeinheit. Daher ist es naheliegend anzunehmen, daß das für die Dualitätsrelation ebenfalls gilt. Das ist, zumindest was deren formale Struktur angeht, zweifellos richtig. Dennoch sind allgemeinere Betrachtungen denkbar, da ihr Beweis in der von Englert entdeckten Form gewisse spezielle Voraussetzungen macht; insbesondere ist er explizit auf den Fall von Zwei-Wege-Interferometern zugeschnitten. Der Versuch der Herleitung einer Dualitätsrelation für Vielstrahlinterferometer ist gegenwärtig Gegenstand engagierter Diskussionen<sup>43</sup>.

Damit beschließen wir unsere Vorbetrachtungen und kommen nun zur angekündigten Unterrichtseinheit zur Quantenmechanik. Wir werden dabei insbesondere auf schulisches Niveau heruntertransformierte Versionen der quantenmechanischen Axiome und des Welle-Teilchen-Dualismus kennenlernen und auch sonst generell wieder stets die mathematische Physik als Hintergrund-Instanz bei der Begriffsbildung berücksichtigen.

## 4.4 Notwendige Vorkenntnisse

Entsprechend der fundamentalen Umwälzung, welche die Quantenmechanik innerhalb der Physik bedeutet, und der Zielsetzung, dies im Gymnasium mit Hilfe eines axiomatischen Zugangs und damit nicht auf traditionelle, eher an der historischen Entwicklung orientierten Weise zu unterrichten, stellt ein solcher Unterrichtsgang notwendigerweise

<sup>42</sup>Martinez-Linares und Vargas Medina beschreiben in [359] die Anwendung dieses Formalismus auf eine reale experimentelle Anordnung, nämlich auf Versuche mit asymmetrischen Ramsay-Interferometern. Siehe dazu auch [150].

<sup>43</sup>Vergleiche etwa [126], [55], [56] und [57].

einen Neuanfang im Physikunterricht dar. Folglich beschränken sich die Vorkenntnisse, die mitgebracht werden sollen, auf einen sehr kleinen Umfang. Andererseits steht die Quantenmechanik weiterhin in der Abschlußklasse des Gymnasiums auf dem Programm, so daß alles, was man sinnvollerweise überhaupt voraussetzen darf, schon vorhanden sein müßte. Die Schülerinnen und Schüler sollten<sup>44</sup>

- sich ein wenig an die Beschreibung des Verhaltens von Massenpunkten mit Hilfe von Bahnen erinnern, wie es exemplarisch beim waagrechten Wurf geschieht,
- die drei mechanischen Energieformen kennen und berechnen können,
- den Begriff der potentiellen Energie auf den Fall des radialen elektrischen Feldes anwenden können,
- harmonische Schwingungen mit Hilfe von Differentialgleichungen beschreiben können,
- die Begriffe der Beugung und der Interferenz kennengelernt haben und
- Intensitätsverteilungen bei Beugung am Doppelspalt und am Gitter elementar berechnen können.

Der dritte Punkt stellt insofern ein gewisses Problem dar, als radiale elektrische Felder in Baden-Württemberg in den zweistündigen Kursen überhaupt nicht mehr und in den vierstündigen Kursen nur noch als Wahlgebiet vorgesehen sind; verpflichtend sind inzwischen ausschließlich homogene elektrische Felder. Wir empfehlen nachdrücklich, dieses Thema im Rahmen der Elektrostatik in zwei- wie in vierstündigen Kursen dennoch durchzunehmen, wobei die Quantenmechanik hier nur eine nachgeordnete Begründung liefert. Radiale Felder sind die einzigen nichthomogenen Felder, die man am Gymnasium behandeln kann und bieten damit die einzige Chance, das Faradaysche Feldkonzept anhand eines nichttrivialen Beispiels zu verdeutlichen. Darüberhinaus kann man hier zur Abwechslung wirklich ganz direkt mit Analogien arbeiten – sofern man zuvor in der zehnten Klasse das Newtonsche Gravitationsgesetz und damit radiale Gravitationsfelder vernünftig unterrichtet hat. Wurden radiale elektrische Felder zuvor nicht behandelt, muß man sie im Vorfeld der Betrachtung des Wasserstoffatoms kurz einführen und den Ausdruck für die potentielle Energie eines geladenen Teilchens im radialen elektrischen Feld mitteilen und etwas erläutern. Beim Rest handelt es sich ausschließlich um Dinge, die gemäß dem Bildungsplan ohnehin auf dem Pflichtprogramm stehen.

## 4.5 Klassische Physik und Quantenmechanik

### 4.5.1 Rückblick: Klassische Physik

Möchte man eine vernünftige Einschätzung des Bruchs vermitteln, den die Quantenmechanik gegenüber der klassischen Physik darstellt, tut man gut daran, ganz zu Beginn über letztere ein paar zusammenfassende Worte zu verlieren. Dabei ist es hilfreich, aber

---

<sup>44</sup>Auch hier verwenden wir den Konjunktiv wieder durchaus absichtsvoll.

nicht ausreichend, die Schülerinnen und Schüler darauf hinzuweisen, daß *alles*, was sie bisher im Fach Physik gehört haben, *ausschließlich* zur klassischen Physik gehört.

Die klassische Physik mit ihren hauptsächlichen Teilgebieten Mechanik, Elektrodynamik, Thermodynamik und Statistik schien Ende des 19. Jahrhunderts zu einem Abschluß der Physik geführt zu haben. Diese Überzeugung geriet ins Wanken, als plötzlich unerwartete Probleme auftraten. Bis dahin stellte sie jahrhundertlang die Grundlage der physikalischen Weltbeschreibung dar, hatte triumphale Höhenflüge beispielsweise in der Beschreibung der Bewegung von Himmelskörpern, in der Mechanik deformierbarer Medien, der Maxwellschen Elektrodynamik, der Thermodynamik und statistischen Mechanik und insbesondere und abschließend in der Speziellen und der Allgemeinen Relativitätstheorie. Dennoch sollte sie mit Ausnahme der Relativitätstheorien zu Beginn des 20. Jahrhunderts komplett zu Fall kommen, was bekanntlich mit der Entstehung und Entwicklung der Quantenmechanik das Aufkommen eines vollkommen neuen und im Vergleich zu früher extrem anderen, fremdartigen Weltbildes einleitete. Auslöser war das Auftreten der erwähnten unlösbaren Schwierigkeiten; die klassische Physik konnte grundlegende, damals neue Beobachtungen nicht erklären, z.B.

- den Zusammenhang von Energiedichte und Wellenlänge bei der elektromagnetischen Strahlung von Körpern im thermodynamischen Gleichgewicht (Schwarze Strahler),
- die Stabilität und die Spektren der Atome,
- die spezifische Wärme der Festkörper.

Es gibt viele weitere Beispiele. Damit zeichnete sich mehr und mehr die Notwendigkeit einer neuen Physik ab.

Um diesen Übergang und das, worauf dann alles hinauslief, kennenlernen zu können, sollte man ein wenig über die Grundprinzipien der klassischen Physik Bescheid wissen; diese lassen sich vereinfacht etwa folgendermaßen zusammenfassen:

- Physikalische Systeme sind im Prinzip durch geeignete Systeme von Differentialgleichungen sowie Rand- und Anfangsbedingungen zu jeder Zeit beliebig genau beschreibbar. Man sagt auch, die klassische Physik sei vollständig *deterministisch*.
- Dabei sind sämtliche Eigenschaften des Systems zu jeder Zeit beliebig genau angebar und gleichzeitig meßbar.
- Messungen sind zumindest im Prinzip möglich, ohne das System dabei zu beeinflussen.
- Statistische Aussagen sind höchstens erforderlich, wenn man nicht genug über die betrachteten Systeme weiß (z. B. in der Thermodynamik und statistischen Mechanik).

Die Quantenmechanik stellt im Licht der Schwierigkeiten der klassischen Physik eine notwendige Neuformulierung dar. Wir werden nun sehen, wie die soeben aufgelisteten Grundprinzipien im Rahmen dieser Neuformulierung durch völlig andere Gesetzmäßigkeiten ersetzt werden.

## 4.5.2 Quantenmechanik als unvermeidliche Neuformulierung der Physik

Zwar nicht den historischen, wohl aber den begrifflichen Startpunkt zur Einarbeitung in die Quantenmechanik stellen gewisse Merkwürdigkeiten dar, die man beim Verhalten physikalischer Systeme feststellt, welche in der klassischen Physik ganz unzweifelhaft *entweder* als Teilchen *oder* als Wellen aufgefaßt wurden.

### 4.5.2.1 Teilchen, die sich wie Wellen zu verhalten scheinen

Wir beginnen mit einem der wenigen echten quantenmechanischen Versuche, die in der Schule tatsächlich durchführbar sind. Die Elektronenbeugungsröhre ist ein Paradebeispiel für eine experimentelle Anordnung, bei der sich vermeintliche Teilchen sehr seltsam verhalten. Elektronen werden in einer Elektronenkanone beschleunigt und treffen auf eine Graphitfolie, die aus vielen kleinen Kristallen besteht. Auf der fluoreszierenden Schicht der Elektronenbeugungsröhre beobachtet man ein Interferenzmuster in Form eines Hauptmaximums in der Mitte und konzentrischen Ringen als Nebenmaxima. Interferenzerscheinungen sind eigentlich ein typisches Wellenphänomen, die Elektronen scheinen sich also wie Wellen zu verhalten. Die beobachtete Ringstruktur des Interferenzmusters erklärt sich dabei folgendermaßen: Da die Graphitkristalle viele unterschiedliche Orientierungen aufweisen, ist die Streuung axialsymmetrisch zur ursprünglichen Richtung des Elektronenstrahls. Alle Streuwinkel treten daher zylindrisch in allen Richtungen auf, und es entstehen ringförmige Maxima und Minima mit einem gemeinsamen Mittelpunkt. In Abbildung 4.1 wird deutlich, daß die Winkel  $\delta$ , unter dem die ringförmigen Maxima auf dem Schirm erscheinen, doppelt so groß wie die tatsächlichen Ablenkungswinkel  $\alpha$  sind, unter denen die Maxima auftreten.

Wenn Elektronen sich hier offensichtlich so ähnlich wie Wellen verhalten, dann müßte man ihnen eine Wellenlänge zuschreiben können. Diese Idee geht auf L. de Broglie zurück [78] - [81], und seine Überlegungen vollziehen wir nun auf elementare Weise nach. Zunächst machen wir dazu eine Analogiebetrachtung mit Photonen. Dazu setzen wir zunächst einmal *rein formel* die beiden Beziehungen  $E = h\nu$  und  $E = mc^2$  gleich und erhalten die ebenfalls rein formale Relation  $h\nu = mc^2$ . Das Produkt aus  $m$  und  $c$  erinnert dabei an so etwas wie einen Impuls für Photonen, aber da nicht klar ist, was die „Photo-

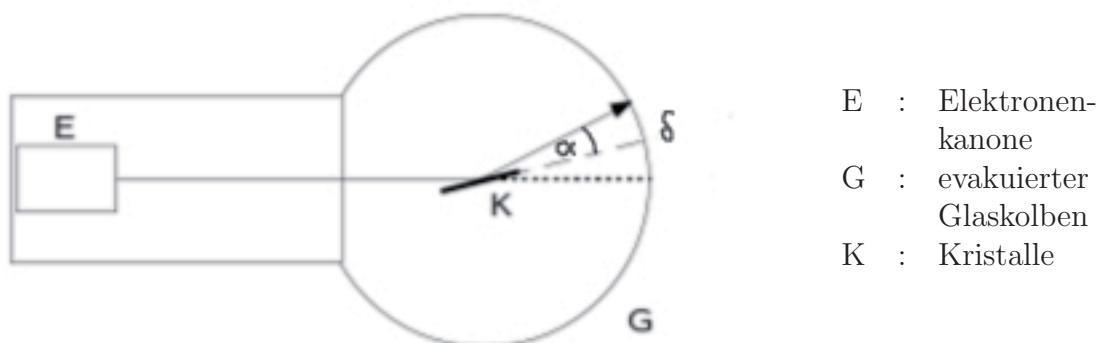


Abbildung 4.1: Aufbau der Elektronenbeugungsröhre

nenmasse“ sein soll, ersetzen wir sie gemäß  $m = E/c^2 = h\nu/c^2$  und erhalten den korrekten Ausdruck für den Photonenimpuls<sup>45</sup>

$$p_{\text{ph}} = \frac{h\nu}{c}.$$

Außerdem gilt bekanntlich  $c = \nu\lambda$ , und damit folgt

$$p_{\text{ph}} = \frac{h}{\lambda}.$$

Die Idee von de Broglie ist nun, für *beliebige Quantenobjekte*, egal ob es Photonen, Elektronen oder andere sind, allgemein die Relation

$$p = \frac{h}{\lambda}$$

zu postulieren. Damit kann man jedem Objekt mit Impuls  $p$  eine Wellenlänge zuordnen gemäß

$$\lambda_{\text{ph}} = \frac{h}{p}.$$

Diese Wellenlänge nennt man die *de Broglie-Wellenlänge* des Teilchens. Für materielle Teilchen mit Masse  $m$  und Geschwindigkeit  $v$  ergibt sich auf diese Weise die Möglichkeit, ihnen vermöge der Relation

$$\lambda_{\text{dB}} = \frac{h}{mv}$$

eine Wellenlänge zuzuordnen. Die Frage ist nun, ob dieser Wellenlänge irgendeine reale Bedeutung zukommt.

Um das zu überprüfen, stellen wir folgende Hypothese auf: Bei der Elektronenbeugung verhalten sich Elektronen so, als ob sie Wellen wären, deren Wellenlänge gleich der de Broglie-Wellenlänge  $\lambda_{\text{dB}} = h/mv$  ist. Da es sich hier um ein Kristallgitter und damit um ein räumliches Gitter handelt, liegt *Bragg-Reflexion* vor<sup>46</sup>. Diese Art der Beugung müssen wir daher etwas genauer betrachten.

Die Ionen eines Kristallgitters lassen sich als räumliches Beugungsgitter auffassen. Zur Herleitung einer Bedingung für Maxima betrachtet man exemplarisch zwei mögliche Wege der an Gitterionen gestreuten Elektronen (siehe Abbildung 4.2).

<sup>45</sup>Die Vorstellung eines Photonenimpulses taucht erstmals 1909 bei J. Stark auf [492].

<sup>46</sup>Bragg-Reflexion ist ebenfalls aus dem Pflichtprogramm des Physikunterrichts gestrichen worden, an dieser Stelle ist es aber sehr hilfreich, sie dennoch zu thematisieren.

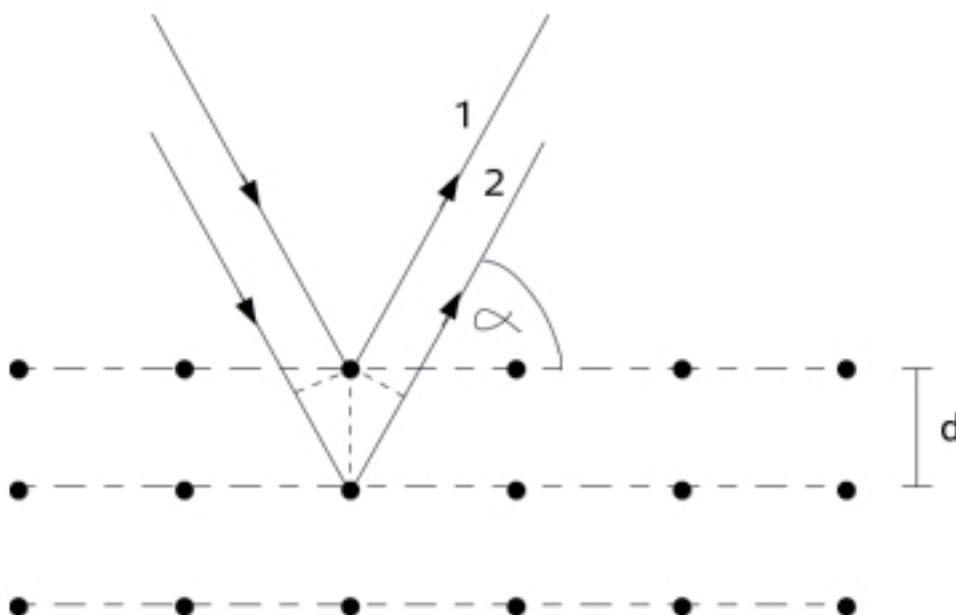


Abbildung 4.2: Zur Bragg-Reflexion

Ist  $d$  der Abstand der Kristallebenen, so gilt für den Gangunterschied  $\Delta$  zwischen Weg 1 und Weg 2

$$\Delta = 2 d \sin \alpha;$$

Damit erhält man für die Winkel, unter denen Maxima auftreten, die Bedingung

$$2 d \sin \alpha = n \lambda, \quad n \in \mathbb{N}.$$

Ist  $r$  der Radius des  $n$ -ten Beugungsrings und  $L$  der Abstand des Streuzentrums von der Wand der Glaskugel, dann folgt näherungsweise

$$d = \frac{n \lambda}{2 \sin (r/2L)}.$$

Bei Graphitkristallen findet man zwei Beugungsringe erster Ordnung. Vergleiche mit den Literaturwerten der Abstände der Kristallebenen von Graphit bestätigen die Hypothese der Bragg-Reflexion.

Es gibt viele weitere Beispiele für wellenähnliches Verhalten bei Teilchen, einige seien hier exemplarisch aufgeführt:

- Elektronenbeugung am Gitter [279]
- Neutronenbeugung an Silizium-Einkristall-Interferometern [422]
- Beugung von Helium-Atomen am Gitter [92], [93]
- Interferenz bei sich durchdringenden Bose-Einstein-Kondensaten [14]



- Beugung von  $C_{60}$  [18] oder  $C_{70}$  [381] am Gitter.
- Interferenz bei Tetraphenylporphyrinmolekülen ( $m = 614$  u) und bei Fluorofullerenen ( $C_{60}F_{48}$ ,  $m = 1632$  u) im Talbot-Lau-Interferometer [209]<sup>47</sup>.

Dabei bestätigt sich die Hypothese der de Broglie-Wellenlänge ausnahmslos. Das vierte, fünfte und sechste Beispiel zeigen überdies, daß die Grenze, bis zu der quantenhaftes Verhalten nachweisbar ist, langsam in den mesoskopischen Bereich hinausgeschoben wird.

#### 4.5.2.2 Wellen, die sich wie Teilchen zu verhalten scheinen

Die Schülerinnen und Schüler kommen nun im allgemeinen selbst darauf, daß man den Spieß auch umdrehen kann. Um zu demonstrieren, daß sich Licht, generell als der Inbegriff eines Wellenphänomens angesehen, gelegentlich auch teilchenähnlich verhalten kann, genügt es, photographische Aufnahmen ein und desselben Objekts mit sehr unterschiedlichen Belichtungszeiten vorzuführen. Bei sehr kurzer Belichtung zeigen sich auf der photographischen Schicht oder dem CCD-Chip des Photoapparats nur einzelne, scheinbar zufällig verteilte Punkte. Betrachtet man Aufnahmen mit immer längeren Belichtungszeiten, so fügen sich die Punkte erst nach und nach immer deutlicher zu einem Bild zusammen. Das legt die Vermutung nahe, daß auf dem Film einzelne, teilchenähnliche Objekte ankommen.

Damit könnte man es eigentlich bewenden lassen, insbesondere auch deshalb, weil man später auf diese Beobachtung zurückgreifen muß. Der Bildungsplan und das schriftliche Abitur zwingen uns jedoch dazu, an dieser Stelle auch den Photoeffekt zu berücksichtigen. Er wird nach wie vor verbreitet als nur im Teilchenbild deutbar dargestellt; das ist jedoch nicht der Fall, wie man schon seit den zwanziger Jahren weiß<sup>48</sup>.

Bei der Demonstration des Photoeffekts kommt eine Vakuum-Photozelle zum Einsatz, die eine Photokathode aus Metall und eine Anode in Form eines Drahttrings enthält. Man macht folgende Beobachtungen:

- Wird die Photozelle beleuchtet, so werden aus ihr Elektronen freigesetzt.
- Dies geschieht erst ab einer bestimmten Grenzfrequenz  $\nu_G$ , die vom Kathodenmaterial abhängt.

<sup>47</sup>Zur Funktionsweise eines Talbot-Lau-Interferometers siehe [76].

<sup>48</sup>Genauso falsch ist die Behauptung, Einstein habe in seiner Arbeit über die Lichtquantenhypothese von 1905 [136] im wesentlichen den Photoeffekt gedeutet. Dieser kommt in der besagten Arbeit nur ganz am Rande in einem einzigen Satz vor, als ein Beispiel, das mit dem Modell der Lichtquanten erklärt werden kann. Als weitere Beispiele erwähnt er ebenso knapp die Stokessche Regel der Photolumineszenz und die Gasionisation durch ultraviolettes Licht. Die Stokessche Regel besagt, daß die Frequenz des ausgestrahlten Lichts diejenige des eingestrahlten Lichts nicht überschreiten kann. Das läßt sich mit dem Energieerhaltungssatz und der Vorstellung erklären, daß bei solchen Prozessen Photonen der Energie  $h\nu_1$  in die photolumineszierende Substanz eindringen und dadurch angeregt dort Photonen der Energie  $h\nu_2$  mit  $\nu_1 \leq \nu_2$  abgegeben werden. Bei der Gasionisation stellt man eine vom Material abhängige maximale wirksame Wellenlänge des verwendeten Lichts fest, die durch die materialabhängige Ionisationsarbeit pro Gasmolekül in Verbindung mit der Energieübertragung von jeweils einem Photon an jeweils ein Gasmolekül verstanden werden kann. Der wesentliche Teil von Einsteins Arbeit besteht aus einer Herleitung der Lichtquantenhypothese über ein Entropie-Argument. Richardson [432], [433] sowie später Wentzel [518], [519] und Beck [46] zeigten 1912 beziehungsweise 1926 und 1927, daß man den Photoeffekt auch mit einer Beschreibung des Lichts ganz im Rahmen der klassischen Maxwell'schen Elektrodynamik, also im Wellenbild, erklären kann.

- Die Energie der Photoelektronen hängt nicht von der Intensität, sondern von der Frequenz des Lichts ab.
- Höhere Intensität verändert nur die Stromstärke des Photostroms.

Zur Messung der Photospannung verwenden wir die sogenannte Gegenfeldmethode (siehe Abbildung 4.3). Durch Anlegen einer Gegenspannung wird dabei der Photostrom auf Null heruntergeregelt. Dann ist die Gegenspannung vom gleichen Betrag wie die Photospannung. Man findet dabei, daß die Photospannung und damit die kinetische Energie der Elektronen mit wachsender Frequenz des Lichts linear zunimmt. Es ergibt sich eine Gerade mit der Gleichung

$$E = h\nu - E_A.$$

Auftragen von  $E$  über  $\nu$  liefert ein Schaubild, wie es in Abbildung 4.4 dargestellt ist. Die Steigung der Geraden ist die Plancksche Konstante  $h$ .

Besondere Bedeutung kommt bei solchen Diagrammen der Nullstelle und dem  $y$ -Achsenabschnitt zu. Wie gesagt treten Photoelektronen erst ab einer bestimmten *Grenzfrequenz*  $\nu_G$  auf. Der Grund dafür ist die Arbeit, welche die Metallelektronen verrichten müssen, um die Kathode überhaupt verlassen zu können. Diese Arbeit nennt man die *Austrittsarbeit*  $W_A = eU_A$  der Photokathode; sie ist nur vom Material abhängig, aus der die Kathode besteht und definiert gleichzeitig die Austrittsenergie  $E_A$ . Die Gerade im  $E$ - $\nu$ -Diagramm schneidet die  $\nu$ -Achse bei der Grenzfrequenz, ab der freigesetzte Photoelektronen auftreten. Der  $y$ -Achsenabschnitt entspricht der Austrittsarbeit des Kathodenmaterial. Eine mögliche Deutung des Photoeffekts ist die folgende:

*Die Energie elektromagnetischer Strahlung mit der Frequenz  $\nu$  wird in Quanten der Größe  $E = h\nu$  wirksam. Diese Quanten nennt man Photonen. Lichtenergie ist quantisiert.  $E$  hängt nicht von der Intensität ab. Helleres Licht hat mehr Quanten, nicht aber Quanten größerer Energie.*

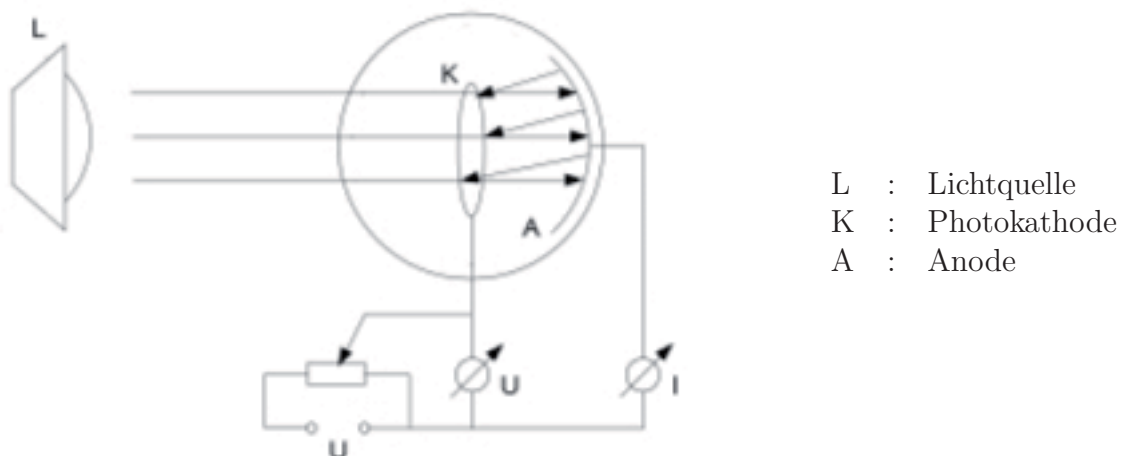
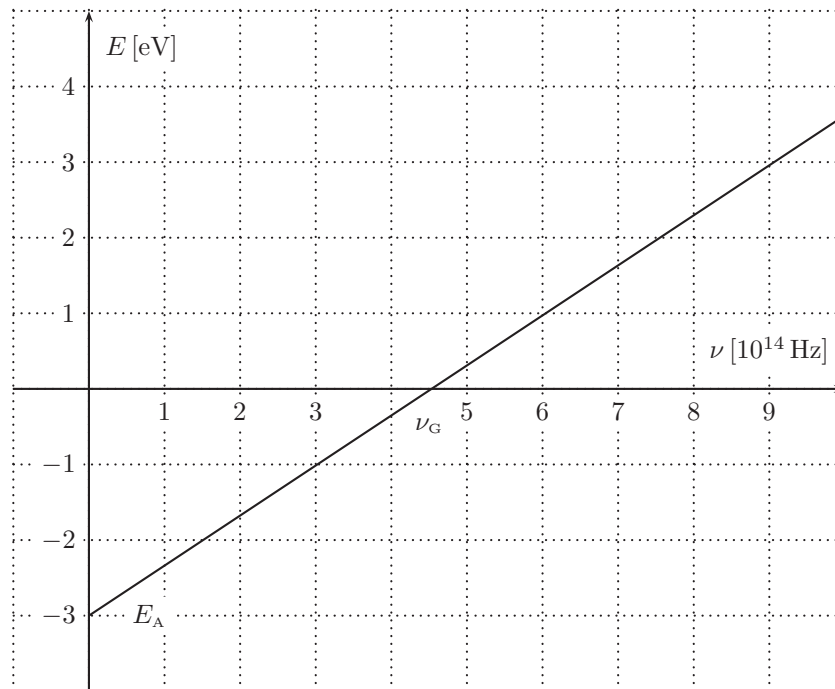


Abbildung 4.3: Messung der Photospannung mit der Gegenfeldmethode

Abbildung 4.4:  $E$ - $\nu$ -Diagramm beim Photoeffekt

Wir betonen nochmals, daß das nicht die einzig mögliche Deutung ist. Insbesondere darf der Photoeffekt keinesfalls als Nachweisexperiment für teilchenähnliches Verhalten elektromagnetischer Wellen verkauft werden.

## Aufgaben

1. Berechne die De Broglie-Wellenlänge für ein Elektron (Masse:  $9,11 \cdot 10^{-31}$  kg), ein Auto (Masse: 1 t), die Erde (Masse:  $6 \cdot 10^{24}$  kg) und die Sonne (Masse:  $2 \cdot 10^{30}$  kg) jeweils für eine charakteristische Geschwindigkeit.
2. Nimmt man die Quantenmechanik ernst, so müßten auch Elefanten (Masse jeweils circa 12 t) Interferenzerscheinungen zeigen. Wir betrachten eine Herde Elefanten, die mit  $v = 25$  m/s durch eine Baumreihe mit der Gitterkonstanten 4 m durchtrampelt und sich hinter dieser versammelt. Begründe mit Hilfe einer Rechnung, warum es sehr schwierig sein dürfte, hinter den Bäumen Häufigkeitsmaxima und -minima der Elefanten zu beobachten.
3. Elektronen mit der Geschwindigkeit  $3,25 \cdot 10^7$  m/s treffen auf einen Graphit-Kristall, wo sie gestreut werden und anschließend auf einen Leuchtschirm gelangen. Dort beobachtet man unter dem Streuwinkel  $\alpha_1 = 5,921^\circ$  beziehungsweise  $\alpha_2 = 10,081^\circ$  ringförmige konzentrische Maxima. Berechne daraus den Abstand der Kristallebenen des Graphitkristalls.

4. Bei einem Interferometrie-Experiment treten Helium-Atome ( $\lambda_{dB} = 1,03 \text{ \AA}$ ) durch einen Doppelspalt mit einem Spaltabstand von  $8 \mu\text{m}$ . Bei dem dabei auf einem  $64 \text{ cm}$  vom Gitter entfernten Detektor auftretenden Interferenzmuster mißt man einen mittleren Abstand der Maxima von  $7,7 \pm 0,5 \mu\text{m}$ . Bestimme die prozentuale Abweichung dieses Resultats vom theoretisch zu erwartenden Wert.
5. Fullerenmoleküle (Masse:  $1,2 \cdot 10^{-24} \text{ kg}$ ) treten einzeln nacheinander durch ein Gitter mit der Gitterkonstanten  $100 \text{ nm}$ . Auf einem  $1,25 \text{ m}$  entfernten Detektor tritt ein Interferenzmuster auf, dessen Maxima in Mittel  $50 \text{ nm}$  voneinander entfernt sind. Berechne die Geschwindigkeit der Fullerenmoleküle.
6. In einem Experiment wird weißes Licht durch ein Prisma zu einem Spektrum aufgefächert und sodann eine Caesium-Photozelle quer vom roten zum violetten Ende durch dieses Spektrum bewegt.
  - a) Beschreibe, was man, entsprechende experimentelle Ausrüstung vorausgesetzt, während der Bewegung der Photozelle bei diesem Versuch beobachten kann.
  - b) Welche Energie und Geschwindigkeit haben die schnellsten ausgelösten Photoelektronen, wenn die Austrittsarbeit bei Caesium  $1,94 \text{ eV}$  beträgt?
  - c) Bewegt man die Photozelle über das violette Ende des Spektrums hinaus, beobachtet man weiterhin Photoelektronen. Erläutere diesen Sachverhalt.
  - d) Kann man eine solche Beobachtung auch machen, wenn man die Photozelle über das rote Ende des Spektrums hinausbewegt? Begründe Deine Antwort.
  - e) Wie könnte man mit einer solchen Versuchsanordnung die Plancksche Konstante messen?
  - f) Was ist deren physikalische Bedeutung?
  - g) Inwiefern kann man Experimente dieser Art als Hinweis auf die Quantennatur des Lichts deuten?

### 4.5.3 Der Welle-Teilchen-Dualismus

Der vorige Abschnitt konfrontiert uns offensichtlich mit folgender Erfahrungstatsache: Quantenobjekte verhalten sich manchmal ähnlich wie Wellen, manchmal ähnlich wie Teilchen. Damit stellt sich die unvermeidliche Frage, was Quantenobjekte denn nun eigentlich sind. Um eine Antwort auf diese Frage zu finden, versuchen wir nun, Quantenobjekte auf die Probe zu stellen. Wir tun das mit einer Klasse von Experimenten, die aus der klassischen Wellenoptik wohlbekannt ist.

Wir betonen an dieser Stelle nochmals, daß es abwegig ist, einer „Überwindung“ des Welle-Teilchen-Dualismus das Wort zu reden, wie es stellenweise sowohl in der Fachdidaktik- wie auch in der Schulbuch-Literatur getan wird<sup>49</sup>. Eine solche Sprechweise beruht im günstigsten Fall auf einer fehlerhaften Verwendung dieses Begriffs. Der Welle-Teilchen-Dualismus, wie er hier beschrieben wird, ist eines der tragenden Grundprinzipien der Quantenmechanik und stellt damit eine fundamentale Eigenschaft der Natur dar. Wie fundamental – das ist überhaupt erst seit den neunziger Jahren bekannt.

---

<sup>49</sup>Siehe beispielsweise [37] oder [51].

### 4.5.3.1 Experimente mit Zwei-Wege-Interferometern

Die Grundidee des Zwei-Wege-Interferometers besteht darin, interferenzfähige Objekte auf ein Registriergerät zu lenken, wobei ihnen aus klassischer Sicht zwei verschiedene Wege zur Verfügung stehen. Kohärenz vorausgesetzt, ergibt sich je nach Gangunterschied der beiden Wege am Registriergerät Verstärkung oder Auslöschung, also ein Interferenzmuster. Beispiele für Zwei-Wege-Interferometer sind

- Der Doppelspalt,
- Das Michelson-Interferometer und
- Das Mach-Zehnder-Interferometer

Wir werden uns zunächst ausführlich mit dem dritten aus dieser Liste beschäftigen.

Bei einem Mach-Zehnder-Interferometer (siehe Abbildung 4.5) werden die von einer Quelle Q ausgesandten Photonen von zwei halbdurchlässigen Spiegeln S1 und S4 mit jeweils gleicher Wahrscheinlichkeit abgelenkt oder durchgelassen, sodaß sie die Detektoren D1 und D2 auf jeweils zwei Wegen erreichen können. Dort kommt es dann zu Interferenzerscheinungen. Dabei ist zu beachten, daß Licht ganz allgemein bei Reflexion an einem halbdurchlässigen Spiegel einen Phasensprung um  $\pi/2$  und bei Reflexion an einem vollverspiegelten Spiegel einen Phasensprung von  $\pi$  erfährt.

Was kommt nun an den beiden Detektoren an? Dazu betrachten wir jeweils die beiden klassisch möglichen Wege getrennt. Am Detektor D1 erfahren die Photonen auf dem Weg S1 S2 S4 D1 die Phasenverschiebung

$$\phi_1 = \pi + \frac{\pi}{2} = \frac{3\pi}{2}$$

und auf dem Weg S1 S3 S4 D1 die Phasenverschiebung

$$\phi_2 = \frac{\pi}{2} + \pi = \frac{3\pi}{2};$$

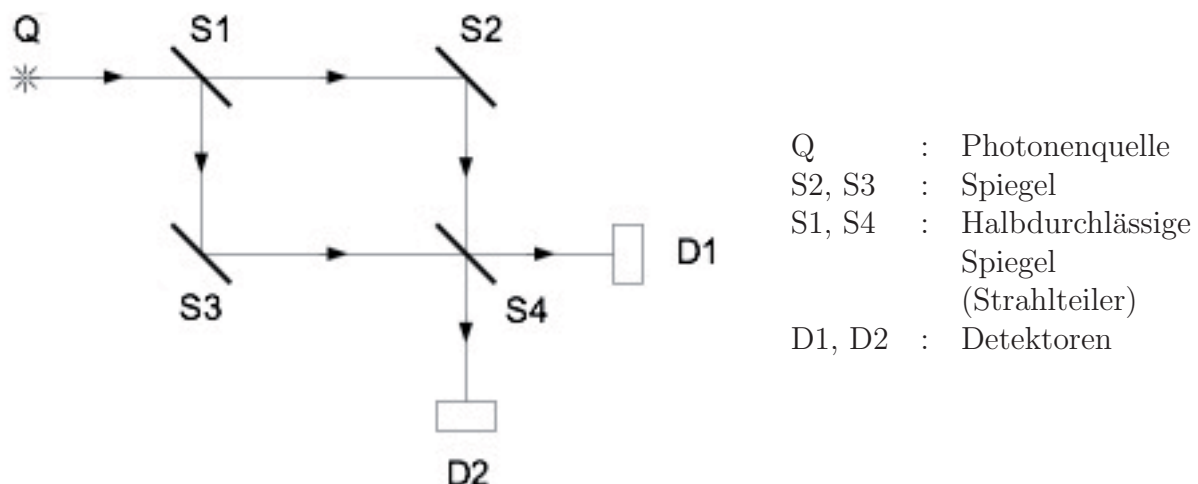


Abbildung 4.5: Aufbau eines Mach-Zehnder-Interferometers

das ergibt eine Phasendifferenz  $\Delta\phi = 0$  und damit *konstruktive Interferenz*, das heißt, bei D1 kommen ständig Photonen an. Am Detektor D2 liefert der Weg S1 S2 S4 D2 eine Phasenverschiebung von

$$\phi_1 = \pi$$

und der Weg S1 S3 S4 D2 eine Phasenverschiebung von

$$\phi_2 = \frac{\pi}{2} + \pi + \frac{\pi}{2} = 2\pi,$$

die Phasendifferenz beträgt somit  $\Delta\phi = \pi$ , und es tritt *destruktive Interferenz* auf; bei D2 kommt demnach nie ein Photon an. Genaugenommen gilt das natürlich nur für die exakte optische Achse des Interferometers. Da ein Laserstrahl keine mathematische Gerade ist, sondern eine gewisse Ausdehnung besitzt, durchläuft die Phasendifferenz radial nach außen kontinuierlich weitere Werte, und man erhält als Interferenzmuster Maxima und Minima in Form *konzentrischer* Kreise. Dabei tritt im Zentrum bei D1 ein Maximum und bei D2 ein Minimum auf.

In dieser Version ist der Versuch eigentlich langweilig, denn alle Interferenzphänomene mit vielen Photonen gleichzeitig sind auch rein klassisch im Rahmen der Wellenoptik erklärbar. Dasselbe gilt in eingeschränktem Ausmaß für alle Interferenzexperimente. Quantenmechanisch interessant wird es erst, wenn immer nur ein Quantenobjekt zur selben Zeit durch das Interferometer geht. Das ist inzwischen technisch in vielfältiger Weise möglich. Experimente mit Quantenobjekten in Zwei-Wege-Interferometern (z. B. Elektronen, Helium-Atome, Fullerenmoleküle am Doppelspalt, Photonen im Michelson-Interferometer oder Mach-Zehnder-Interferometer) liefern dabei folgende Resultate:

- Quantenobjekte zeigen Interferenzmuster. Das haben wir provisorisch mit wellenähnlichem Verhalten, De Broglie-Wellenlängen und so weiter zu erklären versucht.
- Wird die Intensität so weit heruntergedreht, daß sich stets nur höchstens ein Quantenobjekt gleichzeitig im Interferometer befindet, dann
  - ergibt sich zunächst zwar ein scheinbar regelloses Muster mit statistisch verteilten Treffern, aber
  - das Interferenzmuster taucht nach einiger Zeit trotzdem auf.

Hierfür scheint zunächst keine Erklärung möglich zu sein. Das ganze ist weder mit Teilchen- noch mit Wellenvorstellungen deutbar. Um der Sache auf den Grund zu gehen, betrachten wir weitere Experimente mit dem Mach-Zehnder-Interferometer. Natürlich sind Versuche mit einzelnen Quantenobjekten mit schulischen Mitteln nicht durchführbar. Es gibt jedoch leistungsfähige Simulationsprogramme, die hierfür einen adäquaten Ersatz dartellen<sup>50</sup>. Damit lassen sich die nachfolgend beschriebenen Versuche in allen

---

<sup>50</sup>Solche Simulationsprogramme lassen sich im Unterricht genauso mit Gewinn einsetzen, wie sie sich für interessierte Schülerinnen und Schüler selbst zum Experimentieren eignen, da sie völlig intuitiv bedienbar sind. Entsprechende Programme wurden am Lehrstuhl für Didaktik der Physik an der Universität München entwickelt [255] und sind für Windows unter [http://www.didaktik.physik.uni-muenchen.de/materialien/inhalt\\_materialien/doppelspalt/index.html](http://www.didaktik.physik.uni-muenchen.de/materialien/inhalt_materialien/doppelspalt/index.html) und [http://www.didaktik.physik.uni-muenchen.de/materialien/inhalt\\_materialien/interferometer/index.html](http://www.didaktik.physik.uni-muenchen.de/materialien/inhalt_materialien/interferometer/index.html) erhältlich. Für Unix-Systeme wie Linux oder MacOS gibt es leider noch nichts vergleichbares.

Einzelheiten durchexerzieren. Es sollte hierbei jedoch stets betont werden, daß es sich dabei um Simulationen von Experimenten handelt, die genauso real durchführbar sind und auch schon vielfach mit Erfolg durchgeführt wurden.

Im Versuch 1 werden einzelne Photonen nacheinander durch ein Mach-Zehnder-Interferometer geschickt, das zusätzlich auf jedem der beiden möglichen Wege einen drehbaren Polarisationsfilter enthält (siehe Abbildung 4.6). Wir beobachten folgendes:

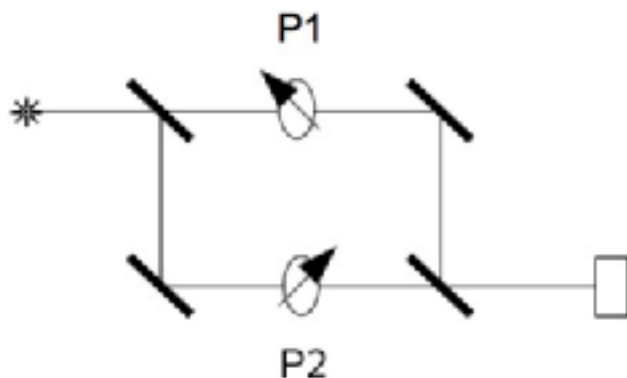
- a) Die Quelle liefert einzelne Photonen (technisch machbar); es erscheint ein Interferenzmuster.
- b) Wird Weg 1 oder Weg 2 versperrt, verschwindet das Interferenzmuster.
- c) Jetzt werden Polarisationsfilter P1 und P2 in Weg 1 und Weg 2 gebracht. Sind beide Polarisationsfilter parallel ausgerichtet, so findet man wieder ein Interferenzmuster.
- d) Werden die beiden Polarisationsfilter senkrecht zueinander ausgerichtet, verschwindet das Interferenzmuster. Wird der Winkel zwischen den Ausrichtungen der beiden Polarisationsfilter langsam von  $0^\circ$  nach  $90^\circ$  gedreht, so wird das anfangs deutliche Interferenzmuster immer undeutlicher und verschwindet schließlich.

Bei a) und c) sind die beiden möglichen Wege durch das Interferometer jeweils *ununterscheidbar*. Man sagt auch, es gibt keine Welcher-Weg-Information. Bei d) sind die beiden Wege unterschiedlich „markiert“, das heißt, die beiden Wege sind *unterscheidbar*. Es gibt also Welcher-Weg-Information. Man erhält offensichtlich nur dann ein Interferenzmuster, wenn die beiden Wege ununterscheidbar sind. Sobald die Wege unterscheidbar sind, verschwindet das Interferenzmuster.

Damit erhalten wir für das Verhalten von Quantenobjekten das folgende, vorläufig formulierte Prinzip:

*Entweder kann man Welcher-Weg-Information herausfinden, oder es entsteht ein Interferenzmuster.*

Dieses Prinzip muß nun präzisiert werden. Wir beginnen damit, indem wir Versuch 2 betrachten. Dazu wird in die beiden Armen des Mach-Zehnder-Interferometers jeweils ein



P1, P2: Polarisationsfilter mit drehbaren Polarisationsrichtungen

Abbildung 4.6: Mach-Zehnder-Interferometer mit Polarisationsfiltern



Polarisationsfilter P1 beziehungsweise P2 eingebaut und deren Richtungen gekreuzt, so daß zunächst kein Interferenzmuster entsteht. Jetzt kommt ein dritter Polarisationsfilter P3 *nach* dem zweiten Strahlteiler dazu, der zu den beiden anderen jeweils um  $45^\circ$  gedreht ist (siehe Abbildung 4.7). Die verblüffende Folge: Das Interferenzmuster taucht wieder auf. Offensichtlich macht der dritte Polarisationsfilter die eigentlich vorhandene Welcher-Weg-Information *nachträglich* wieder unkenntlich. Das Interferenzmuster kann durch nachträgliches Auslöschen der Welcher-Weg-Information wieder hergestellt werden. *Entscheidend ist die Möglichkeit* Welcher-Weg-Information zu bekommen, nicht aber ob sie überhaupt eingeholt wurde. Eine solche Anordnung nennt man *Quantenradierer*<sup>51</sup>.

Wir ziehen erste Folgerungen aus den Beobachtungen der bisher beschriebenen Versuche:

- Jedes Quantenobjekt interferiert jeweils mit sich selbst.
- Quantenobjekte scheinen zu wissen,
  - ob Weg 1 oder Weg 2 oder beide Wege offen sind und
  - ob sie beobachtet werden oder nicht.

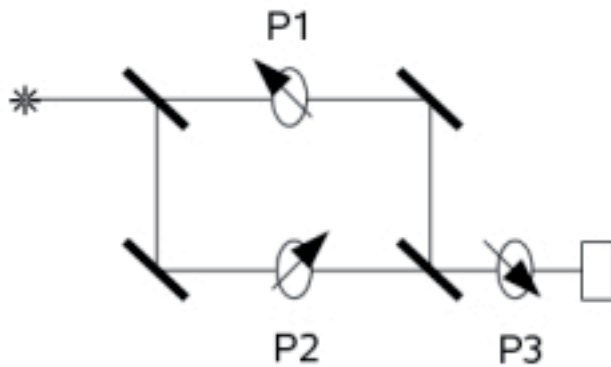
Damit stellen sich aber folgende Fragen:

- Bei Versuch 1a): Wie kann das Photon gleichzeitig beide Wege durchlaufen?
- Bei Versuch 1b): Wie kann ein Photon, das sich anscheinend in einem Weg des Interferometers befindet, wissen, ob der andere Weg offen ist oder nicht?
- Bei Versuch 1c) und d): Wie kann ein Photon, das anscheinend einen Polarisationsfilter durchläuft, wissen, wie der andere Polarisationsfilter eingestellt ist?
- Beim Quantenradierer und allgemein: Wie kann das Photon wissen, ob es beobachtet werden kann oder nicht?

Die erste der obigen Folgerungen provoziert zusätzlich eine weitere Frage: Teilt sich das Photon etwa auf? Hierzu betrachten wir Versuch 3, in welchem wir versuchen, nachzuprüfen, wo die Photonen „wirklich“ sind. Dazu werden in beiden Armen des Mach-Zehnder-Interferometers je ein Photonenzähler eingebaut, wodurch die beiden Wege gleichzeitig versperrt werden. Man beobachtet, daß von den beiden Photonenzählern jeweils stets nur einer klickt, aber niemals beide gleichzeitig. Die Photonen teilen sich nicht auf. *Wenn man nachsieht*, ist das Photon stets *entweder* in Weg 1 *oder* in Weg 2. Als vorläufiges Ergebnis finden wir somit folgendes:

---

<sup>51</sup>In diesem Zusammenhang werden gelegentlich Analogieversuche für Quantenradierer vorgeschlagen. Beispielsweise kann man einen „Quantenradierer“ mit einem Mach-Zehnder-Interferometer für Mikrowellen aufbauen [443]. Als Polarisationsfilter verwendet man Gitter mit Gitterkonstanten im cm-Bereich. Ein weiterer Modell-Quantenradierer, diesmal mit einem Laser als Quelle, besteht aus einem Doppelspaltversuch mit je einem drehbaren Polarisationsfilter vor beiden Spalten und einem gemeinsamen drehbaren Polarisationsfilter nach den Spalten [327]. Es handelt sich hierbei allerdings nur um Modellversuche, die auch im Rahmen der klassischen Wellenoptik problemlos erklärt werden können, da im Sculunterricht natürlich keine Versuche mit einzelnen Photonen möglich sind.



P1, P2, P3: Polarisationsfilter mit drehbaren Polarisationsrichtungen

Abbildung 4.7: Der Quantenradierer

- Das Auftreten von Interferenzmustern läßt sich mit der Wellenvorstellung, nicht aber mit der Teilchenvorstellung erklären.
- Das Auffinden der Quantenobjekte entweder im einen oder im anderen Weg des Interferometers läßt sich mit der Teilchenvorstellung, nicht aber mit der Wellenvorstellung erklären.
- Das statistische Auftreten einzelner Einschläge auf dem Schirm, das Auftreten von Interferenzmustern bei einzeln durchlaufenden Quantenobjekten sowie der Quantenradierer lassen sich weder mit der Teilchenvorstellung noch mit der Wellenvorstellung erklären.

#### 4.5.3.2 Formulierung des Prinzips vom Welle-Teilchen-Dualismus

Interferenzversuche mit einzelnen Quantenobjekten, wie sie die Versuche 1, 2 und 3 exemplarisch darstellen, lassen sich insgesamt offensichtlich weder mit einem Teilchenbild noch mit einem Wellenbild erklären. Quantenobjekte sind *weder* Teilchen *noch* Wellen, sondern irgendetwas ganz anderes, das wir uns nicht vorstellen können. Sie verhalten sich manchmal so ähnlich wie Teilchen und manchmal so ähnlich wie Wellen. Dies ist Inhalt des Prinzips vom *Welle-Teilchen-Dualismus*. Hierbei handelt es sich um einen Dualismus von teilchenähnlichem und wellenähnlichem Verhalten. Teilchenähnliches Verhalten manifestiert sich durch das Vorhandensein unterscheidbarer Wege für die Quantenobjekte, wellenähnliches Verhalten durch die Fähigkeit der Quantenobjekte, Interferenzmuster hervorzubringen. Damit hat man gleichzeitig einen Dualismus von Unterscheidbarkeit der Wege und Interferenzfähigkeit. Der Übergang ist, wie wir gesehen haben, allerdings fließend von perfekter bis völlig fehlender Unterscheidbarkeit der Wege einerseits und völlig fehlender bis perfekter Sichtbarkeit des Interferenzmusters andererseits. Damit erhalten wir eine Präzise qualitative Formulierung vom

*Welle-Teilchen-Dualismus:*

*Je besser die Interferenzfähigkeit von identischen Quantenobjekten ist, desto schlechter lassen sich verschiedene mögliche Wege von ihnen unterscheiden und umgekehrt.*

Gelegentlich findet man auch folgende alternative Formulierung: Je besser sich verschiedene mögliche Wege von identischen Quantenobjekten unterscheiden lassen, desto geringer ist die Sichtbarkeit von Interferenzerscheinungen. Hier ist jedoch Vorsicht geboten, denn in dieser Version ist die Umkehrung des Prinzips falsch. Beispielsweise führt selbst absolute Ununterscheidbarkeit der Wege bei speziellen Versuchsanordnungen im Falle destruktiver Interferenz nicht zum Auftreten von Interferenzerscheinungen<sup>52</sup>.

Wie genauere Analysen zeigen, läßt sich das ganze quantifizieren. Dabei definiert man in geeigneter Weise eine Größe  $\mathcal{D}$  als Maß für die Welcher-Weg-Information und eine Größe  $\mathcal{V}$  als Maß für die Sichtbarkeit von Interferenzmustern<sup>53</sup> und erhält dafür die *Dualitätsrelation*

$$\mathcal{D}^2 + \mathcal{V}^2 \leq 1.$$

Diese Ungleichung repräsentiert den Welle-Teilchen-Dualismus in allen Facetten, sowohl was die Extremfälle, als auch was die Zwischenstufen angeht, auf folgende Weise:

- Je besser die verschiedenen Wege unterscheidbar sind (je größer  $\mathcal{D}$ ), desto geringer die Interferenzfähigkeit (desto kleiner  $\mathcal{V}$ ).
- Je besser die Interferenzfähigkeit (je größer  $\mathcal{V}$ ), desto geringer die Unterscheidbarkeit der Wege.
- Maximale Unterscheidbarkeit der Wege ( $\mathcal{D} = 1$ ) bewirkt verschwindende Interferenzfähigkeit ( $\mathcal{V} = 0$ ).
- Maximale Interferenzfähigkeit ( $\mathcal{V} = 1$ ) setzt fehlende Unterscheidbarkeit der Wege voraus ( $\mathcal{D} = 0$ ).

Obige Dualitätsrelation ist in dieser Form exakt herleitbar<sup>54</sup> und wurde experimentell qualitativ und quantitativ in allen Einzelheiten bestätigt<sup>55</sup>.

#### 4.5.3.3 Anwendung: Wechselwirkungsfreie Messungen

Eine wichtige Anwendung des Welle-Teilchen-Dualismus stellen sogenannte wechselwirkungsfreie Messungen dar. Man sagt dazu auch zerstörungsfreie Messungen<sup>56</sup>. Diese erlauben es, ein typisches Problem quantenmechanischer Messungen zu umgehen: Da letztere das gemessene System im allgemeinen verändern, kann es passieren, daß der zu untersuchende Zustand durch die Messung zerstört wird. Das Grundprinzip zerstörungsfreier Messungen läßt sich anhand eines Gedankenexperiments erläutern, dem *Bombentest von Elitzur und Vaidman* [140]<sup>57</sup>. Man betrachtet dazu ein Mach-Zehnder-Interferometer, wie

<sup>52</sup>Ein weiterer Grund dafür, daß diese Fassung des Prinzips vom Welle-Teilchen-Dualismus nicht allgemein umkehrbar ist, ist in Abschnitt 4.3.3.3 beschrieben.

<sup>53</sup>Die Wahl der Variablen ergibt sich aus den englischen Ausdrücken *Distinguishability* für Unterscheidbarkeit und *Visibility* für Sichtbarkeit.

<sup>54</sup>Siehe auch hierzu Abschnitt 4.3.3.3.

<sup>55</sup>Das gelang 1998 S. Dürr, T. Nonn und G. Rempe [127], [128].

<sup>56</sup>Englisch Non Interaction Measurements beziehungsweise Nondemolition Measurements

<sup>57</sup>Vergleiche auch [513].

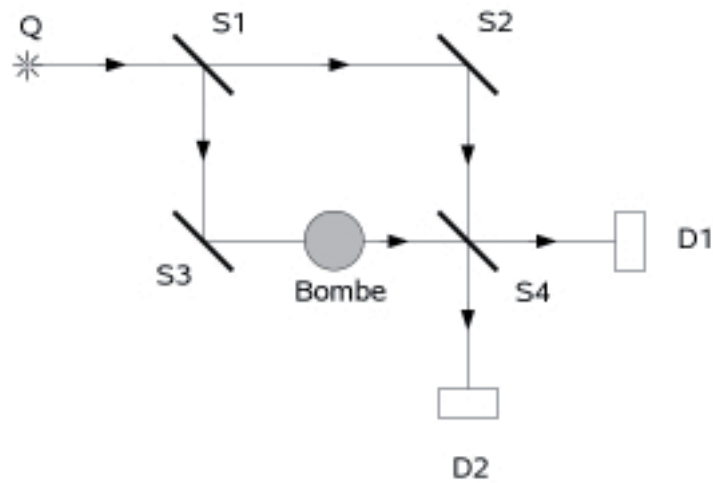


Abbildung 4.8: Das Bombentest-Experiment von Elizur und Vaidman

es weiter oben beschrieben wurde. Die Aufgabenstellung ist nun die folgende: Im unteren Arm des Interferometers könnte sich eine Bombe befinden, die schon explodiert, wenn sie von nur einem einzigen Photon getroffen wird. Wie kann man herausfinden, ob eine Bombe vorhanden ist, ohne daß sie explodiert? Nun gibt es zwei Möglichkeiten:

1. Keine Bombe im unteren Weg:

Es liegt keine Welcher-Weg-Information vor, daher tritt konstruktive Interferenz bei Detektor 1 und destruktive Interferenz bei Detektor 2 auf. Alle Photonen kommen bei Detektor 1 an, kein einziges bei Detektor 2.

2. Bombe im unteren Weg:

Es liegt Welche-Weg-Information vor, daher tritt keine Interferenz auf.

- Bei S1 nimmt das Photon mit einer Wahrscheinlichkeit von 50 % den Weg nach unten, und die Bombe explodiert. Kein Detektor tickt.
- Mit einer Wahrscheinlichkeit von 50 % nimmt es bei S1 den Weg nach rechts.
- Bei S4 nimmt es mit einer Wahrscheinlichkeit von 50 % den Weg nach rechts, und Detektor 1 tickt.
- Mit einer Wahrscheinlichkeit von 50 % nimmt es bei S4 den Weg nach unten, und Detektor 2 tickt.

Da Detektor 2 ohne Bombe nie tickt, kann man die Bombe mit einer Wahrscheinlichkeit von 25 % nachweisen, ohne sie zu zerstören. Damit erhält man in diesen Fällen Informationen über den unteren Weg, ohne dort hingeschaut zu haben. Mit einer raffiniert erweiterten Version des Versuchs kann man die Nachweiswahrscheinlichkeit auf ca. 90 % erhöhen. Anordnungen nach obigem Prinzip ermöglichen es, Messungen an quantenmechanischen Systemen zu machen, ohne diese zu beeinflussen.

#### 4.5.3.4 Vorüberlegungen zum quantenmechanischen Zustandsbegriff

Bevor wir uns nun daranmachen, die in Abschnitt 4.2 in mathematisch exakter Form aufgelisteten Axiome der elementaren Quantenmechanik in einer der gymnasialen Oberstufe angemessenen vereinfachten Form vorzustellen, sind noch ein paar Bemerkungen zum Begriff des Zustands eines physikalischen Systems angebracht, da sich die klassische und die quantenmechanische Fassung dieses Begriffs sehr stark unterscheiden. Das läßt sich aus den bisherigen Resultaten unseres Unterrichtsganges ableiten.

In der klassischen Physik ist der Zustand eines Systems beispielsweise durch dessen Bahn beschreibbar, etwa, wenn es sich dabei um einen Massenpunkt in einem Kraftfeld handelt. Ein sehr instruktives Beispiel ist der schiefe Wurf<sup>58</sup>. Der Zustand des Systems „Punktteilchen beim schiefen Wurf“ wird zu jeder Zeit vollständig durch dessen Bahnkurve, also einer Funktion  $\vec{r} = \vec{r}(t)$ , beschrieben. Für die Quantenmechanik zeigen Experimente wie zum Beispiel die Versuche 1, 2 und 3 aus Abschnitt 4.5.3.1, daß dieses Konzept hier nicht funktioniert. Der Bahnbegriff macht häufig keinen Sinn. Genauer gesagt beweisen die Versuche 1, 2 und 3, daß dem Photon im Falle der Sichtbarkeit eines Interferenzmusters die Eigenschaft, eine bestimmte Bahn zu beschreiben, *nicht zugeschrieben werden kann*. Allgemeiner gilt folgendes:

*In der Quantenmechanik kann es vorkommen, daß ein System sich in einem Zustand befindet, in dem es bestimmte, in der klassischen Physik wohlbestimmte Eigenschaften nicht besitzt. (I)*

Dies gilt natürlich auch für Eigenschaften, die kein klassisches Analogon haben. Die Aussage (I) lautet eigentlich allgemeiner: In der Quantenmechanik kann es vorkommen, daß ein System sich in einem Zustand befindet, in dem es bestimmte Eigenschaften nicht besitzt. Da quantenmechanische Eigenschaften ohne klassische Analogien jedoch in der gymnasialen Oberstufe nicht oder höchstens am Rand von Bedeutung sind, empfiehlt sich die erste Formulierung, die überdies den Unterschied zur klassischen Physik für Neulinge des Gegenstands deutlicher zum Ausdruck bringt. – Als Folgerung ergibt sich unmittelbar:

*Quantenmechanische Systeme können sich in Zuständen befinden, die in der klassischen Physik nicht möglich wären. (II)*

Wir betrachten als Beispiel wieder Photonen im Mach-Zehnder-Interferometer, und genauer die Versuche 1a), c) und 2 aus Abschnitt 4.5.3.1.

Zu (I): Ist ein Interferenzmuster sichtbar, so können einzelnen Photonen im Interferometer die Eigenschaften, einen Ort zu besitzen, eine Bahn zu beschreiben, u. a., nicht zugeschrieben werden.

Zu (II): Ist ein Interferenzmuster sichtbar, so befindet sich jedes einzelne Photon im Interferometer in einem Zustand, in dem es sich gleichzeitig in beiden Armen des Interferometers aufhält.

Das läßt sich mit dem oben beschriebenen klassischen Zustandskonzept in keiner Weise vereinbaren. Damit ziehen wir zwei weitere Folgerungen:

<sup>58</sup>Da dieser üblicherweise im Gymnasium nicht mehr behandelt wird, behelfe man sich mit dem waagrecht Wurf, an dem alle wesentlichen, hier diskutierten Aspekte genauso abgelesen werden können.

- Zur Beschreibung des Zustandes physikalischer Systeme ist in der Quantenmechanik ein *grundlegend anderes Konzept* erforderlich als in der klassischen Physik.
- In der Quantenmechanik gibt es für physikalische Systeme *viel mehr mögliche Zustände* als in der klassischen Physik.

Die Bereitstellung eines neuen, quantenmechanischen Zustandskonzepts und die Beantwortung, was für Zustände in diesem Sinn physikalische Systeme einnehmen können, ist zwar nicht das einzige, aber ein wesentliches Ziel der folgenden Abschnitte.

## Aufgaben

- Elektronen werden in einem elektrischen Feld beschleunigt und treffen senkrecht auf einen Doppelspalt, dessen Spaltmittenabstand  $1,0 \mu\text{m}$  beträgt. Danach werden sie auf einem  $34 \text{ cm}$  vom Doppelspalt entfernten Leuchtschirm nachgewiesen. Auf diesem ist ein Interferenzmuster mit mehreren Intensitätsmaxima zu erkennen. Die beiden Maxima erster Ordnung haben einen Abstand von  $150 \mu\text{m}$  voneinander.
  - Mit welcher Geschwindigkeit sind die Elektronen durch den Doppelspalt getreten?
  - Die Elektronen treten *einzelnd nacheinander* durch den Doppelspalt. Beschreibe genau, was man auf dem Schirm
    - nach kurzer Zeit und wenigen durchgelaufenen Elektronen und
    - nach längerer Zeit und sehr vielen durchgelaufenen Elektronen
 beobachten kann. Inwiefern ist dieses Verhalten weder mit einer klassischen Teilchen- noch mit einer klassischen Wellenvorstellung der Elektronen zu erklären?
- Neutronen werden an einem Kristall aus dem Kohlenstoff-Isotop  $^{13}\text{C}$  gestreut. Dabei beobachtet man Bragg-Reflexion mit mehreren ringförmigen Maxima und Minima.
  - Erkläre das Auftreten des Bragg-Interferenzmusters mit Hilfe des Welle-Teilchen-Dualismus.
  - Ein kleiner Teil der Neutronen verursacht bei der Streuung ein Umklappen der Spins der Kohlenstoffatome, an denen sie gestreut wurden, wobei auch der Spin dieser Neutronen umgeklappt wird. Solche Spin-Änderungen sind mit geeigneten experimentellen Mitteln meßbar, die hier jedoch nicht zum Einsatz kommen. Inwiefern machen sich solche Prozesse bei der genauen Struktur des Interferenzmusters bemerkbar?
- Erläutere das Prinzip des Welle-Teilchen-Dualismus anhand eines Michelson-Interferometers (Abbildung 4.9).

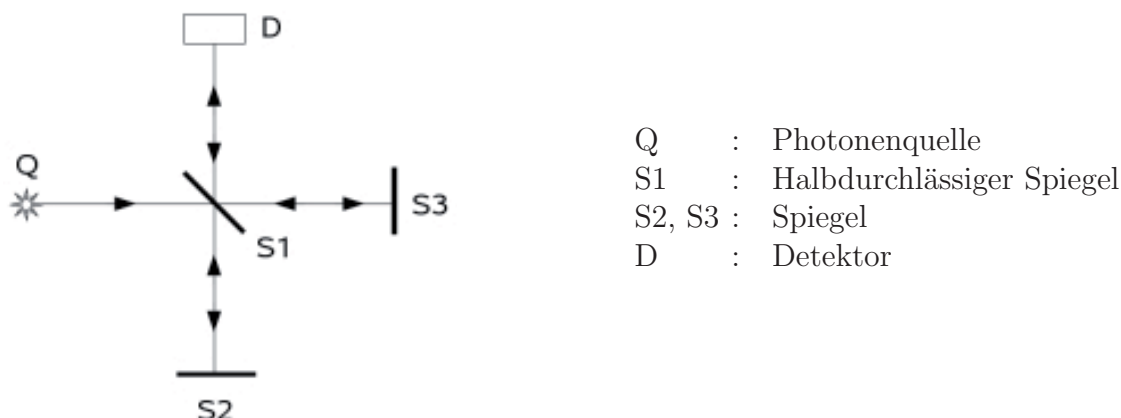


Abbildung 4.9: Schematischer Aufbau eines Michelson-Interferometers

4. Wir betrachten ein Mach-Zehnder-Interferometer, in das zusätzlich zwei nichtlineare Kristalle K1 und K2 sowie zwei weitere Detektoren D1 und D2 eingebaut wurden (Abbildung 4.10). Ein solcher Kristall macht aus jedem Photon zwei Photonen mit jeweils etwa der halben Frequenz des ursprünglichen Photons. Dabei wird das eine Photon (das *Idler-Photon*) jeweils rechtwinklig abgelenkt, so daß es zum Detektor D1 beziehungsweise D2 gelangt, das andere Photon (das *Signal-Photon*) läuft jeweils zunächst geradeaus weiter, verbleibt damit im Interferometer und trägt damit zum Resultat am Detektor D bei.

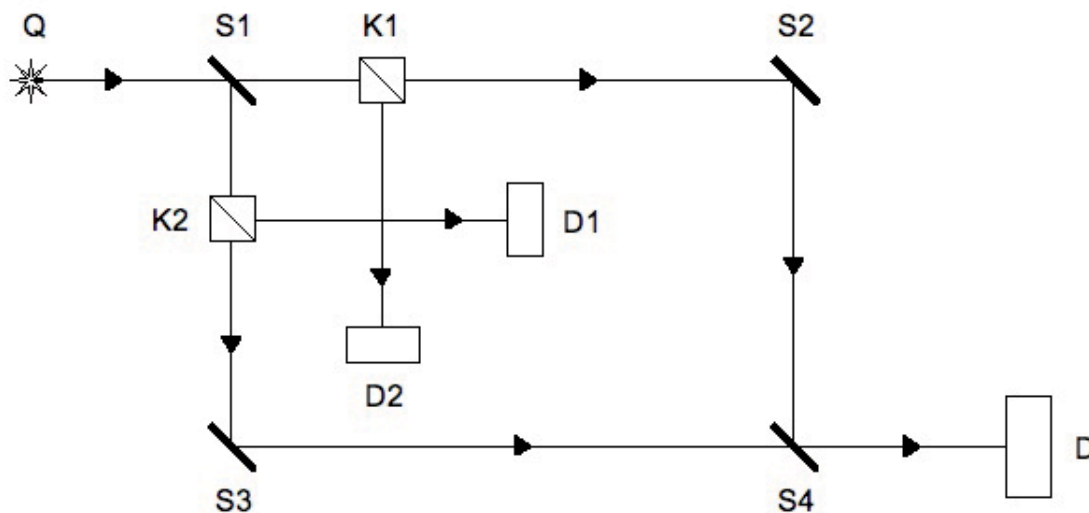


Abbildung 4.10: Mach-Zehnder-Interferometer mit nichtlinearen Kristallen

- a) Was beobachtet man bei dieser Anordnung am Detektor D? Begründe Deine Antwort.



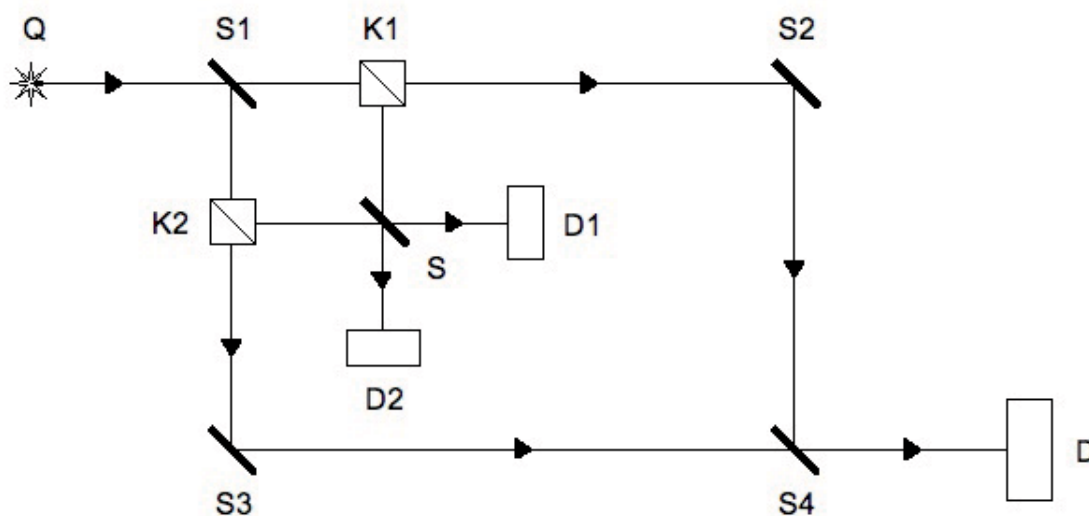


Abbildung 4.11: Quantenradierer, zweite Variante

- b) Nun wird ein zusätzlicher Strahlteiler S in das Interferometer eingebaut (Abbildung 4.11). Was ändert sich nun am Detektor D? Begründe auch hier Deine Antwort.

#### 4.5.4 Elementare Axiomatik der Quantenmechanik

Die Quantenmechanik baut, wie jedes Theoriengebäude der Physik, auf einem System von Axiomen, also unbeweisbaren Grundregeln auf, die postuliert werden müssen, also vom Himmel fallen. Sie können lediglich einerseits auf ihre Widerspruchsfreiheit und andererseits auf ihre Tauglichkeit hinsichtlich der Vorhersage von Meßresultaten überprüft werden – und bewähren sich dabei aus beiden Perspektiven glänzend.

In Abschnitt 4.2 wurden diese Axiome in einer mathematisch präzisen, aber damit auch sehr formalen Sprache dargestellt; in dieser Form erlauben sie die systematische Ableitung des gesamten Apparates der nichtrelativistischen Quantenmechanik, sind aber gleichzeitig natürlich weit davon entfernt, in der Schule unterrichtbar zu sein<sup>59</sup>. Wir beschreiben nun, wie man die Axiome in einer stark elementarisierten Form so im Physikunterricht des Gymnasiums präsentieren kann, daß einerseits die zentralen Aussagen erhalten bleiben<sup>60</sup>, andererseits aber die Schulmathematik zu ihrer Diskussion bei weitem

<sup>59</sup>Selbst in Kursvorlesungen über Quantenmechanik an den Universitäten werden vereinfachte Versionen der beschriebenen Axiome vorgetragen, die dann jedoch nicht mehr die selbe mathematische Strenge aufweisen. Insbesondere die Spektraltheorie steht in denjenigen Semestern, in welchen im allgemeinen einführende Quantenmechanik-Vorlesungen stattfinden, normalerweise noch nicht zur Verfügung und ist ganz generell kein typischer Bestandteil des Stoffkanons im Physikstudium.

<sup>60</sup>Da die komplexen Zahlen im Pflichtprogramm der Oberstufenmathematik bis auf weiteres außen vor bleiben, muß ein wesentlicher Charakterzug der Quantenmechanik hier unberücksichtigt bleiben. Die komplexen Zahlen sind, anders als in der klassischen Physik, wo sie rein rechentechnische Hilfsmittel darstellen, für die Quantenmechanik ein integraler, unvermeidlicher Bestandteil. Die gesamte, hier beschriebene Unterrichtseinheit beschränkt sich, den Vorgaben des Mathematikunterrichts folgend, auf rein reelle Größen. Das ist, wie wir sehen werden, jedoch verkraftbar, und stellt kein Hindernis dar, wesentliche Besonderheiten quantenmechanischen Verhaltens auch schon in der Schule kennenzulernen.

ausreicht. Wie oben werden wir auch hier die Axiome zunächst gemeinsam auflisten und erst anschließend im Detail betrachten. Das hat den erwünschten Nebeneffekt, die Aussage zu unterstreichen, daß es sich hier um unbeweisbare, als gegeben zu akzeptierende Grundtatsachen handelt.

#### 4.5.4.1 Die Grundregeln der Quantenmechanik

1. Der Zustand eines quantenmechanischen Systems wird durch die Wellenfunktion  $\psi$  des Systems vollständig beschrieben.
2. Es ist möglich, daß sich quantenmechanische Systeme in Zuständen befinden, in denen sie gewisse in der klassischen Physik wohlbestimmte Eigenschaften nicht besitzen.
3. Wird an einem quantenmechanischen System eine physikalische Größe gemessen, so erhält man zwar stets ein Meßresultat, das heißt jedoch nicht, daß das System die gemessene Eigenschaft vor der Messung überhaupt besessen hat.
4. Nach der Messung befindet sich das System in einem Eigenzustand der gemessenen Größe, das heißt, eine sofortige Wiederholung der Messung dieser Größe liefert wieder das selbe Ergebnis.
5. Die Wahrscheinlichkeit, das System in einem Gebiet  $G$  anzutreffen, ist gegeben durch

$$\mathfrak{P}(G) = \int_G |\psi(x)|^2 dx.$$

6. Die Wellenfunktion eines quantenmechanischen Systems ist gegeben als Lösung der Schrödingergleichung

$$-\frac{\hbar^2}{2m} \psi'' + V \psi = E \psi$$

des Systems; dabei ist  $m$  die Masse des Systems,  $V$  das Potential, in dem es sich befindet, und  $E$  seine Gesamtenergie.

#### 4.5.4.2 Interpretation und Bemerkungen

Wir betrachten nun die Axiome im einzelnen. Dabei kann es sicher nicht schaden, jeweils auch darauf hinzuweisen, wo und andeutungsweise auch inwiefern hier Vereinfachungen im Interesse der Schulphysik erforderlich sind.

Zu Grundregel 1:

- Jedem physikalischen System ist eine Wellenfunktion  $\psi$  zugeordnet.
- $\psi$  ist eine vom Ort und von der Zeit abhängige Funktion:

$$\psi : \mathbb{R}^3 \times \mathbb{R} \mapsto \mathbb{C}, \quad \psi = \psi(x, y, z, t).$$

Wir beschränken uns hier auf reellwertige und im wesentlichen auf zeitunabhängige Wellenfunktionen sowie auf eine räumliche Variable. Die zuletzt genannte Vereinfachung bedarf gegenüber den Schülerinnen und Schülern einer Rechtfertigung, da diese naturgemäß mit der Vorstellung eindimensionaler Systeme nichts anfangen können. Es genügt dabei, darauf hinzuweisen, daß es einerseits Systeme gibt, bei denen nur eine Raumrichtung wesentlich ist, etwa langgestreckte Kettenmoleküle oder kugelsymmetrische Orbitale, und daß andererseits bei drei Raumvariablen alles im Prinzip gleich verläuft und nur die Rechnungen sehr viel mühsamer sind.

- Die Wellenfunktion enthält sämtliche Informationen, die man über das zugehörige System prinzipiell wissen kann und beschreibt dieses daher vollständig. Folglich können wir aus physikalischer Sicht nicht etwa deshalb nicht mehr über ein System wissen, als uns dessen Wellenfunktion sagen kann, weil wir zu doof sind, sondern weil es *prinzipiell nicht mehr darüber zu wissen gibt*. Die Art und Weise dieser Beschreibung wird durch die weiteren Grundregeln 2 bis 6 geregelt.
- Gibt es für ein physikalisches System mehrere verschiedene möglichen Zustände, so gibt es für dieses System auch mehrere Wellenfunktionen. Zu jedem Zustand gibt es eine eigene Wellenfunktion. Die Schrödingergleichung besitzt in solchen Fällen mehrere verschiedene Lösungen. Häufig gibt es für physikalische Systeme unendlich viele mögliche Zustände und entsprechend unendlich viele verschiedene Wellenfunktionen.

Zu Grundregel 2:

Eine erste wesentliche Folgerung aus diesem Axiom ist die Feststellung, daß die Beschreibung quantenmechanischer Systeme mit klassischen Begriffen im allgemeinen grundsätzlich nicht möglich ist. Näheres dazu folgt bei Grundregel 4.

Zu Grundregel 3:

Im dritten Axiom zeigt sich der fundamentale Unterschied zwischen klassischen und quantenmechanischen Meßprozessen. In der klassischen Physik geht man ganz selbstverständlich davon aus, daß ein System, an dem man gerade eine Größe gemessen und dabei einen bestimmten Meßwert erhalten hat, diese gemessene Eigenschaft unmittelbar vor der Messung bereits aufwies. In der Quantenmechanik gilt jedoch: Auch wenn ein System eine bestimmte Eigenschaft nicht besitzt, kann diese an ihm gemessen werden. Das Meßresultat läßt dann jedoch im allgemeinen keine Rückschlüsse über diese Eigenschaft vor der Messung zu.

Zu Grundregel 4:

Hier taucht erstmals ein besonders wichtiger quantenmechanischer Begriff auf, nämlich der des *Eigenzustands* einer physikalischen Größe. Ein System befindet sich genau dann in einem Eigenzustand einer bestimmten physikalischen Größe, wenn wiederholte Messungen dieser Größe an identischen solchen System stets den gleichen Meßwert erbringen. Befindet sich das System nicht in einem Eigenzustand der gemessenen Größe, so muß bei wiederholten Messungen auch dann mit unterschiedlichen Meßergebnissen gerechnet werden, wenn diese Messungen allesamt an identischen Systemen vorgenommen werden.

Gleichzeitig erfahren die merkwürdigen Eigenschaften quantenmechanischer Messungen ihre Fortsetzung: Physikalische Zustände befinden sich nach einer Messung in einem Eigenzustand der gemessenen Größe, vor der Messung aber im allgemeinen nicht. Das bedeutet: Die gemessene Eigenschaft kann dem System unmittelbar nach der Messung stets zugeschrieben werden, vor der Messung aber im allgemeinen nicht. Eine fundamentale Konsequenz aus den Grundregeln 3 und 4 ist daher folgende Aussage:

*Durch Messungen werden quantenmechanische Systeme im allgemeinen verändert.*

Hier muß jedoch unbedingt darauf hingewiesen werden, daß das *nichts* mit irgendwelchen „Störungen“ oder sonstigen Beeinflussungen des Systems durch Wechselwirkungen welcher Art auch immer zu tun hat<sup>61</sup>. Damit ist diese Beeinflussung auch durch noch so raffinierte technische Tricks nicht einzudämmen; sie ist grundsätzlicher Art und tritt immer auf. In der klassischen Physik geht man dagegen davon aus, daß Messungen zumindest im Prinzip möglich sind, ohne das gemessene System zu beeinflussen.

Zu Grundregel 5:

In den Axiomen 2, 3 und 4 deutet sich eine Besonderheit der Quantenmechanik bereits an, die nun im Axiom 5 festgeschrieben wird. Wenn quantenmechanische Messungen das System, an dem sie vorgenommen werden, verändern, darf man nicht unbedingt erwarten, daß Meßreihen schön vorhersehbare Ergebnisse liefern. Daraus läßt sich bereits folgern, daß die Quantenmechanik im allgemeinen keine deterministischen Voraussagen erlaubt, sondern nur Wahrscheinlichkeitsaussagen macht. Wie das genau vor sich geht, wird im Axiom 5 beschrieben, wobei wir uns hier natürlich wieder auf den allereinfachsten, stark idealisierten Fall beschränken; die Grundidee der Wahrscheinlichkeitsinterpretation ist dabei gleichwohl klar erkennbar.

Die Wellenfunktion  $\psi$  eines Systems läßt sich nicht ohne weiteres direkt physikalisch interpretieren. Axiom 5 sagt uns, daß  $\psi(x)$  eine *Wahrscheinlichkeitsamplitude* ist, das heißt,  $|\psi(x)|^2$  eine *Wahrscheinlichkeitsdichte*<sup>62</sup>. Das ist die zentrale Aussage der *statistischen Deutung der Wellenfunktion*.

Zum Begriff der Wahrscheinlichkeitsdichte sind einige zusätzliche Bemerkungen angebracht. Eine Dichte muß erst mit etwas multipliziert werden, damit die dadurch beschriebene Größe herauskommt; man vergleiche das mit Größen wie Massendichte, Ladungsdichte, und dergleichen mehr, die mit einem Volumen multipliziert werden müssen, damit Masse, Ladung, und so weiter dabei herauskommen. Hier gilt: Die Wahrscheinlichkeit, ein durch  $\psi$  beschriebenes System in dem *kleinen* Bereich  $\Delta x$  um  $x$  herum anzutreffen,

<sup>61</sup>Die Fehlinterpretationen, auf die man hier in der physikalischen Literatur stößt, sind unüberschaubar, angefangen bei den Veröffentlichungen der Pioniere der Quantenmechanik selbst über ungezählte gutgemeinte, aber abwegige populäre Darstellungen bis zu weitverbreitetem Unsinn in vielen Schulbüchern. Das neueste Beispiel ist ein jüngst im Selbstverlag und zusätzlich in einem Sonderdruck speziell für die Gymnasien in Baden Württemberg verlegtes Heft von A. Wünschmann [531], in dem nahezu sämtliche Fehldeutungen, die ihren Begründern in der Anfangsphase der neueren Quantenmechanik unterliefen, konsequent nachgeturnt werden. Siehe zu dieser Problematik auch Aufgabe 3.

<sup>62</sup>Die Betragsstriche sind erforderlich, da  $\psi$  im allgemeinen komplex ist. Hier kommen ausschließlich reelle  $\psi$ 's vor, sodaß wir die Betragsstriche weglassen könnten. Wir schreiben sie aber trotzdem hin, um den auftretenden Formeln die übliche Gestalt zu verleihen.

beträgt

$$\Delta\mathfrak{P}(x) = |\psi(x)|^2 \Delta x.$$

Für größere Gebiete  $G$  muß man die einzelnen Wahrscheinlichkeiten addieren:

$$\mathfrak{P}(G) = \sum |\psi(x)|^2 \Delta x.$$

Da  $\psi$  von  $x$  abhängt, muß man genaugenommen zum Integral übergehen:

$$\mathfrak{P}(G) = \int_G |\psi(x)|^2 dx.$$

Behält man die Beschränkung auf eine Raumdimension bei, wie es hier durchgängig geschieht, ist das Gebiet  $G$  einfach ein Intervall  $[a, b]$ , und das Integral bekommt die Gestalt

$$\mathfrak{P}([a, b]) = \int_a^b |\psi(x)|^2 dx.$$

Wir werden das später anhand konkreter Beispiele illustrieren.

Die Wahrscheinlichkeit, das betrachtete Quantenobjekt überhaupt irgendwo zu finden, muß 1 sein. Das führt auf die Forderung

$$\int_{-\infty}^{+\infty} |\psi(x)|^2 dx = 1.$$

Man nennt das *Normierung der Wellenfunktion*. Diese Forderung ist an jede Wellenfunktion zu stellen, was durch geeignete Wahl eines Vorfaktors auch stets erfüllbar ist.

Einige Anmerkungen, die normalerweise nicht erwähnt werden sollten, für besonders interessierte Schülerinnen und Schüler jedoch sinnvolle Ergänzungen darstellen, sind die folgenden:

- Das Normierungsintegral ist ein Spezialfall des *Skalarprodukts* für Wellenfunktionen,

$$(\varphi, \psi) = \int_{-\infty}^{\infty} \varphi(x) \psi(x) dx,$$

und schreibt sich dann einfach

$$(\psi, \psi) = 1.$$

- Im allgemeinen hat man es mit komplexen Wellenfunktionen zu tun; das Skalarprodukt schreibt sich dann

$$(\varphi, \psi) := \int_{-\infty}^{\infty} \varphi^*(x) \psi(x) dx,$$

wobei der Stern für komplexes Konjugieren steht. An Normierungsintegralen ändert sich dadurch wegen  $\psi^* \psi = |\psi|^2$  nichts.

- Im allgemeinen hat man es außerdem mit dreidimensionalen Systemen zu tun. Das Skalarprodukt lautet dann

$$(\varphi, \psi) := \int \int \int_{\mathbb{R}^3} \varphi^*(x, y, z, t) \psi(x, y, z, t) dx dy dz = \int \int \int_{\mathbb{R}^3} \varphi^*(x, y, z, t) \psi(x, y, z, t) dV$$

und die Normierungsbedingung

$$\int \int \int_{\mathbb{R}^3} |\psi(x, y, z, t)|^2 dx dy dz = 1.$$

Zu Grundregel 6:

Dynamische Größen, die physikalische Systeme beschreiben, erhält man in der Regel als Lösungen von bestimmten Differentialgleichungen (man vergleiche zum Beispiel die Differentialgleichung der harmonischen Schwingungen). Die Differentialgleichung, die als ihre Lösungen die möglichen Wellenfunktionen eines physikalischen Systems liefert, ist die Schrödingergleichung dieses Systems. Dabei ist  $\hbar$  die Plancksche Konstante,  $\hbar = h/2\pi$ ,  $m$  ist die Masse des betrachteten Systems, und  $V$  ist das *Potential* – ein Begriff, der einer näheren Erklärung bedarf. Für  $V$  ist derjenige Term einzusetzen, der die potentielle Energie des Systems beschreiben würde, wenn es sich um ein klassisches System handeln würde.  $V$  darf physikalisch jedoch im allgemeinen *nicht* als potentielle Energie interpretiert werden.  $E$  ist die Gesamtenergie des Systems. Die besondere Bedeutung dieser Größe liegt auf der Hand, denn der Energiesatz gilt bei abgeschlossenen Systemen auch in der Quantenmechanik.

Die Schrödingergleichung muß postuliert werden, das heißt, sie fällt vom Himmel. Der axiomatische Charakter dieser Gleichung wird am besten verdeutlicht, wenn sie ohne weitere Begründung als gegeben zu akzeptieren ist. Die Schrödingergleichung kann man nicht herleiten, und das ganz einfach deshalb, *weil es keine Herleitung gibt*. „Plausibelmachen“ ihrer Form führt zu Fehlinterpretationen und sollte vermieden werden<sup>63</sup>.

Wir betrachten hier natürlich eine „abgespeckte“ Spezialform der Schrödingergleichung, die sich als Sonderfall der allgemeinen Schrödingergleichung für eindimensionale, stationäre (zeitunabhängige) Systeme erweist. Die allgemeine Form der Schrödingergleichung lautet

$$-\frac{\hbar^2}{2m} \left( \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2} \right) \psi + V\psi = i\hbar \frac{\partial \psi}{\partial t}$$

und kann im schulischen Physikunterricht nicht näher diskutiert werden. Man sollte sie dennoch einmal anschreiben und auch erwähnen, daß es auch noch weiter verallgemeinerte Formen der Schrödingergleichung sowie auch Verallgemeinerungen der Schrödingergleichung selbst gibt.

Aus den Grundregeln 3, 4 und 6 folgt das *Problem des quantenmechanischen Meßprozesses*: Messungen bewirken eine abrupte, also un stetige Veränderung des Zustands des System, an dem gemessen wird. Die Schrödingergleichung liefert als Differentialgleichung jedoch generell stetige Lösungen. Folglich gilt:

<sup>63</sup>Leider halten sich im allgemeinen weder Schulbücher noch Fortbildungen an diesen Rat.

*Die Schrödingergleichung kann den Vorgang des Meßprozesses nicht beschreiben.*

Genauer gesagt bedeutet das: Vor der Messung befindet sich das System normalerweise nicht in einem Eigenzustand der gemessenen Größe. Die Messung bewirkt einen abrupten Übergang in einen solchen Eigenzustand, das heißt, die Wellenfunktion, die sich gemäß der Schrödingergleichung stetig ändern sollte, wird plötzlich unstetig verändert. Man nennt das den Kollaps oder auch die Reduktion der Wellenfunktion. Es scheint damit zwei unterschiedliche Dynamiken in der Quantenmechanik zu geben, eine für ungestörte Systeme und eine für den quantenmechanischen Meßprozeß. Wie diese beiden Dynamiken unter einen Hut zu bringen sind, das heißt also, wann genau diese Reduktion stattfindet, wodurch sie ausgelöst wird und was genau dabei passiert, ist bis jetzt eine offene Frage.

Häufig wird behauptet, die Grundregeln 4 und 5 bedeuteten die Aufgabe des Determinismus in der Quantenmechanik. Das ist nicht richtig, wie Grundregel 6 zeigt: Die zeitliche Entwicklung von Wellenfunktionen und damit auch von allen daraus folgenden Größen wie etwa quantenmechanischen Wahrscheinlichkeitsdichten erfolgt *streng deterministisch* gemäß der Schrödingergleichung, denn wenn die Rand- und Anfangswerte der Wellenfunktionen und ihrer ersten Ableitungen bekannt sind, kann die Wellenfunktion eines System zumindest im Prinzip zu jeder Zeit und an jedem Ort berechnet werden<sup>64</sup>. Was nicht vorausgesagt werden kann, sind die Meßwerte bei quantenmechanischen Messungen, es sei denn, das System befindet sich bereits vor der Messung in einem Eigenzustand der gemessenen Größe.

Eine weitere und wahrscheinlich die wichtigste Folgerung aus Grundregel 6 ist das quantenmechanische Superpositionsprinzip. Dies wird später in einem eigenen Abschnitt diskutiert.

#### 4.5.4.3 Beispiele

Wir betrachten einige konkrete Systeme. Da eine exakte Behandlung zu kompliziert wäre, begnügen wir uns hier mit einer qualitativen Diskussion. Quantitative Überlegungen zu einfachen Systemen folgen im übernächsten Abschnitt.

a) Ebene Wellen; Teilchen mit definiertem Impuls

$\psi$  sei eine sich ausbreitende ebene Welle mit fester Wellenlänge  $\lambda$ <sup>65</sup>. Für materielle Teilchen der Masse  $m$  ist  $\lambda = h/p$  die de Broglie-Wellenlänge. Folglich beschreibt  $\psi$  ein Teilchen mit *fest definiertem Impuls*. Die Funktion  $\psi$  ist dabei auf den gesamten Raum verteilt, das heißt, ein durch die Wellenfunktion  $\psi$  beschriebenes Teilchen *hat keinen Ort*. Das entspricht der Tatsache, daß einem quantenmechanischen Teilchen die Eigenschaft,

<sup>64</sup>Zeitabhängige Wellenfunktionen werden in diesem Kurs wie gesagt quantitativ nicht betrachtet.

<sup>65</sup>Formal ist das die Wellenfunktion

$$\psi(x) = \frac{e^{2\pi i \vec{x} \vec{n} / \lambda}}{(2\pi)^{3/2}},$$

wobei  $\vec{n}$  ein Einheitsvektor in Ausbreitungsrichtung der Welle ist. Solche Zustände sind jedoch nicht nur für den Physikunterricht zu schwierig, sondern auch aus der Sicht der mathematischen Quantenmechanik nicht trivial, da sie ein Beispiel für Wellenfunktionen darstellen, die nicht normierbar sind. Dieser Sachverhalt ist erst durch eine Erweiterung des quantenmechanischen Formalismus auf die Theorie der Distributionen angemessen handhabbar; siehe dazu auch die Abschnitte 2.6.3 und 4.2.2.



einen definierten Impuls zu besitzen und die Eigenschaft, einen definierten Ort zu besitzen, nicht gleichzeitig zugeschrieben werden kann – ein Sachverhalt, der uns im nächsten Abschnitt ausführlich beschäftigen wird.

b) Elektronen am Doppelspalt

Ist  $\psi$  die Wellenfunktion eines Elektrons in einem Interferometer, so ist gemäß Axiom 5 der Zahlenwert von  $|\psi(x)|^2$  ein Maß für die Wahrscheinlichkeit, einen Punkt am Ort  $x$  auf dem Leuchtschirm zu finden. Genauer gesagt ist  $\mathfrak{P}([a, b]) = \int_a^b |\psi(x)|^2 dx$  die Wahrscheinlichkeit, ein Elektron zwischen  $x = a$  und  $x = b$  zu finden. Nun sei bei einem Doppelspaltinterferometer

- (i)  $\psi_1$  die Wellenfunktion des Elektrons, wenn nur der linke Spalt offen ist und
- (ii)  $\psi_2$  die Wellenfunktion des Elektrons, wenn nur der rechte Spalt offen ist.

Wenn beide Spalte offen sind, erhält man als Wellenfunktion des Elektrons folglich den Ausdruck

$$\psi = \psi_1 + \psi_2.$$

Klassisch würde man für die Wahrscheinlichkeitsdichte, das Elektron an einer bestimmten Stelle auf dem Schirm anzutreffen, die Summe der Einzelwahrscheinlichkeitsdichten erwarten, also  $|\psi_1|^2 + |\psi_2|^2$ . Das überprüfen wir jetzt, indem wir diese Wahrscheinlichkeitsdichte quantenmechanisch durch Bilden des Betragsquadrates der Wellenfunktion  $\psi$  berechnen:

$$|\psi|^2 = |\psi_1 + \psi_2|^2 = |\psi_1|^2 + |\psi_2|^2 + 2\psi_1\psi_2.$$

Das unterscheidet sich vom klassischen Resultat durch einen zusätzlichen gemischten Term auf der rechten Seite<sup>66</sup>.

Die physikalische Interpretation dieser bis hierher rein mathematischen Resultate liegt nun auf der Hand: Ist der linke Spalt offen, erhält man auf dem Schirm ein in der Verlängerung des linken Spaltes zentriertes Interferenzbild des Einfachspaltes, ist der rechte Spalt offen, erhält man das selbe in der Verlängerung des rechten Spaltes. Klassisch müßte man nun erwarten, daß man, wenn beide Spalte offen sind, auf dem Schirm die Summe dieser beiden Bilder erhält, also zwei Hauptmaxima jeweils in Verlängerung der beiden Spalte. Dies entspricht der Wahrscheinlichkeitsdichte

$$w = |\psi_1|^2 + |\psi_2|^2.$$

<sup>66</sup>Eigentlich müßte die Berechnung mit komplexwertigen Wellenfunktionen durchgeführt werden und lautet dann wie folgt:

$$\begin{aligned} |\psi|^2 &= |\psi_1 + \psi_2|^2 = (\psi_1 + \psi_2)(\psi_1 + \psi_2)^* \\ &= |\psi_1|^2 + |\psi_2|^2 + \psi_1^*\psi_2 + \psi_1\psi_2^* = |\psi_1|^2 + |\psi_2|^2 + 2\Re(\psi_1^*\psi_2). \end{aligned}$$

Für die hier verfolgten Zwecke liefert das jedoch nichts neues, weil die Eigenschaft quantenmechanischer Wellenfunktionen, im allgemeinen komplexwertig zu sein, zwar in der Tat fundamental ist, das jedoch anders als vielfach behauptet nicht die Ursache für die Interferenzfähigkeit quantenmechanischer Superpositionen darstellt; diese Ursache ist vielmehr die Linearität der Schrödingergleichung.

Die Quantenmechanik liefert jedoch ein anderes Resultat, und das bestätigt sich auch experimentell: Es entsteht das bekannte Interferenzmuster des Doppelspalt mit dem Hauptmaximum in der Mitte und symmetrischen Nebenmaxima auf beiden Seiten davon. Dies entspricht der Wahrscheinlichkeitsdichte

$$w = |\psi_1|^2 + |\psi_2|^2 + 2 \psi_1 \psi_2.$$

Der zusätzliche gemischte Term zeigt, daß ein Interferenzmuster auftritt. Dieser Term heißt daher *Interferenzterm*. Eine genauere Berechnung unter Verwendung geeigneter Modell-Wellenfunktionen, die mit schulischen Mitteln nicht durchführbar ist, liefert eine Intensitätsverteilung, die gerade dem beobachteten Interferenzmuster entspricht<sup>67</sup>. Dieses wird offensichtlich vom quantenmechanischen Formalismus korrekt vorausgesagt. Die Abbildungen 4.12 und 4.13 veranschaulichen den beschriebenen Unterschied<sup>68</sup>.

Nach diesen theoretischen Vorüberlegungen wäre es nun angebracht, dieselben experimentell zu überprüfen. Da das wiederum mit den Mitteln der Schulphysik nicht durchführbar ist, behelfen wir uns auch hier mit einer Computersimulation des Doppelspaltexperimentes für Elektronen<sup>69</sup>. Dabei soll die Elektronenquelle wieder so schwach sein, daß sie einzelne Elektronen liefert und sich stets nur höchstens ein Elektron in der Apparatur befindet. Bei dieser Gelegenheit können wir das in Abschnitt 4.5.3 eingeführte Konzept des Welle-Teilchen-Dualismus exemplarisch wiederholen. Wir führen vier verschiedene Experimente durch:

- a) Der linke Spalt ist offen, der rechte geschlossen. Auf dem Schirm erscheint nach und nach eine Anhäufung von Einschlägen in Verlängerung des linken Spaltes.
- b) Der rechte Spalt ist offen, der linke geschlossen. Auf dem Schirm erscheint nach und nach eine Anhäufung von Einschlägen in Verlängerung des rechten Spaltes.
- c) Beide Spalte sind offen. Auf dem Schirm erscheint nach und nach das bekannte Interferenzmuster des Doppelspalt mit hellen und dunklen Streifen.
- d) Jetzt wird einer der beiden Spalte beleuchtet (das heißt mit Photonen beschossen). Das Interferenzmuster verschwindet. Stattdessen erscheint eine Überlagerung der Muster der beiden Einzelspalte.

Interferenzeffekte der Einzelspalte werden hier vernachlässigt, da diese die Situation verkomplizieren, aber prinzipiell nichts ändern. Bei genügend engen Spalten ist diese Näherung problemlos möglich. Nun zur Erklärung:

- Bei den Versuchen a) und b) gibt es jeweils nur einen möglichen Weg und in der hier verwendeten Näherung folglich keine Interferenzmuster.
- Bei Versuch c) gibt es *keine Welcher-Weg-Information* und damit *maximale Sichtbarkeit des Interferenzmusters*, denn aus  $\mathcal{D} = 0$  und  $\mathcal{D}^2 + \mathcal{V}^2 = 1$  folgt  $\mathcal{V} = 1$ .

<sup>67</sup>Genauer zur Berechnung der Intensitätsverteilung bei Interferenzmustern steht in [487].

<sup>68</sup>Die Wellenfunktion eines Elektrons in einem Interferometer ist natürlich eine durchaus nichttriviale Angelegenheit. Als Näherung werden meist Gaußsche Wellenpakete herangezogen, wobei auch hier gilt, daß damit nicht nur aus schulischer Sicht nichts anzufangen ist, sondern auch aus der Sicht der mathematischen Physik in die Trickkiste der Distributionen gegriffen werden muß.

<sup>69</sup>Sieh Fußnote 50 in Abschnitt 4.5.3.1.

- Bei Versuch d) kann man durch Beobachtung der gestreuten Photonen herausfinden, durch welchen Spalt die Elektronen gegangen sind (Für Versuch c) machen solche Aussagen keinen Sinn: Jedes Elektron geht irgendwie durch beide Spalte gleichzeitig). Daher gibt es (im Idealfall) *maximale Welcher-Weg-Information* und damit *kein Interferenzmuster*. Denn aus  $\mathcal{D} = 1$  und  $\mathcal{D}^2 + \mathcal{V}^2 = 1$  folgt  $\mathcal{V} = 0$ .

Die beobachteten Ergebnisse entsprechen genau den oben mit Hilfe der Wellenfunktionen berechneten Ergebnissen: Mit den dort verwendeten Bezeichnungen lautet die Wellenfunktion für Versuch a)  $\psi_1$  und für Versuch b)  $\psi_2$ . Die Wahrscheinlichkeitsdichte ist folglich bei Versuch c)  $w = |\psi_1|^2 + |\psi_2|^2 + 2\psi_1\psi_2$ , mit Interferenzterm, und für Versuch d)  $w = |\psi_1|^2 + |\psi_2|^2$ , ohne Interferenzterm. Damit können wir aus dem Welle-Teilchen-Dualismus eine Rechenregel für Wellenfunktionen ableiten:

*Sind bei einem Zwei-Wege-Interferometer  $\psi_1$  und  $\psi_2$  die beiden Wellenfunktionen für Teilchen in Arm 1 beziehungsweise Teilchen in Arm 2, dann gelten für die Berechnung der Wahrscheinlichkeitsdichten auf dem Schirm folgende Rechenregeln:*

(i) *Mit Welcher-Weg-Information: Erst quadrieren, dann addieren*

(ii) *Ohne Welcher-Weg-Information: Erst addieren, dann quadrieren*

Diese Rechenregel berücksichtigt zwar nur die beiden Extremfälle vollständiger Welcher-Weg-Information beziehungsweise vollständig fehlender Welcher-Weg-Information, für diese Fälle gilt sie dafür ganz allgemein. Wir wenden die Rechenregel nun auf das Doppelspalt-Experiment an. Im Fall (i) erhält man

$$w = |\psi_1|^2 + |\psi_2|^2,$$

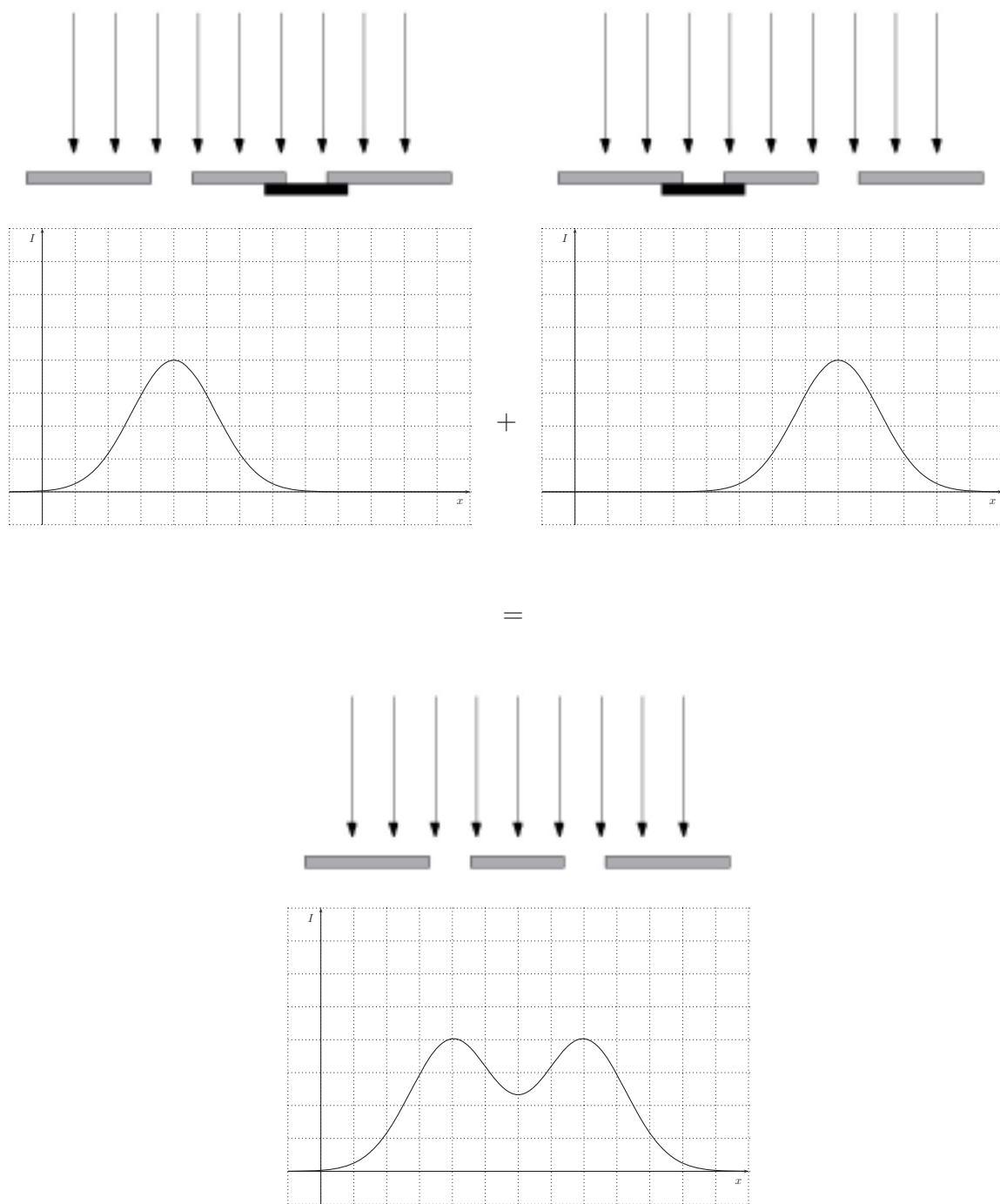
also keinen Interferenzterm und damit kein Interferenzmuster, im Fall (ii) dagegen

$$w = |\psi_1 + \psi_2|^2 = |\psi_1|^2 + |\psi_2|^2 + 2\psi_1\psi_2,$$

also einen Interferenzterm und damit auch ein Interferenzmuster, in völliger Übereinstimmung mit den Beobachtungen<sup>70</sup>.

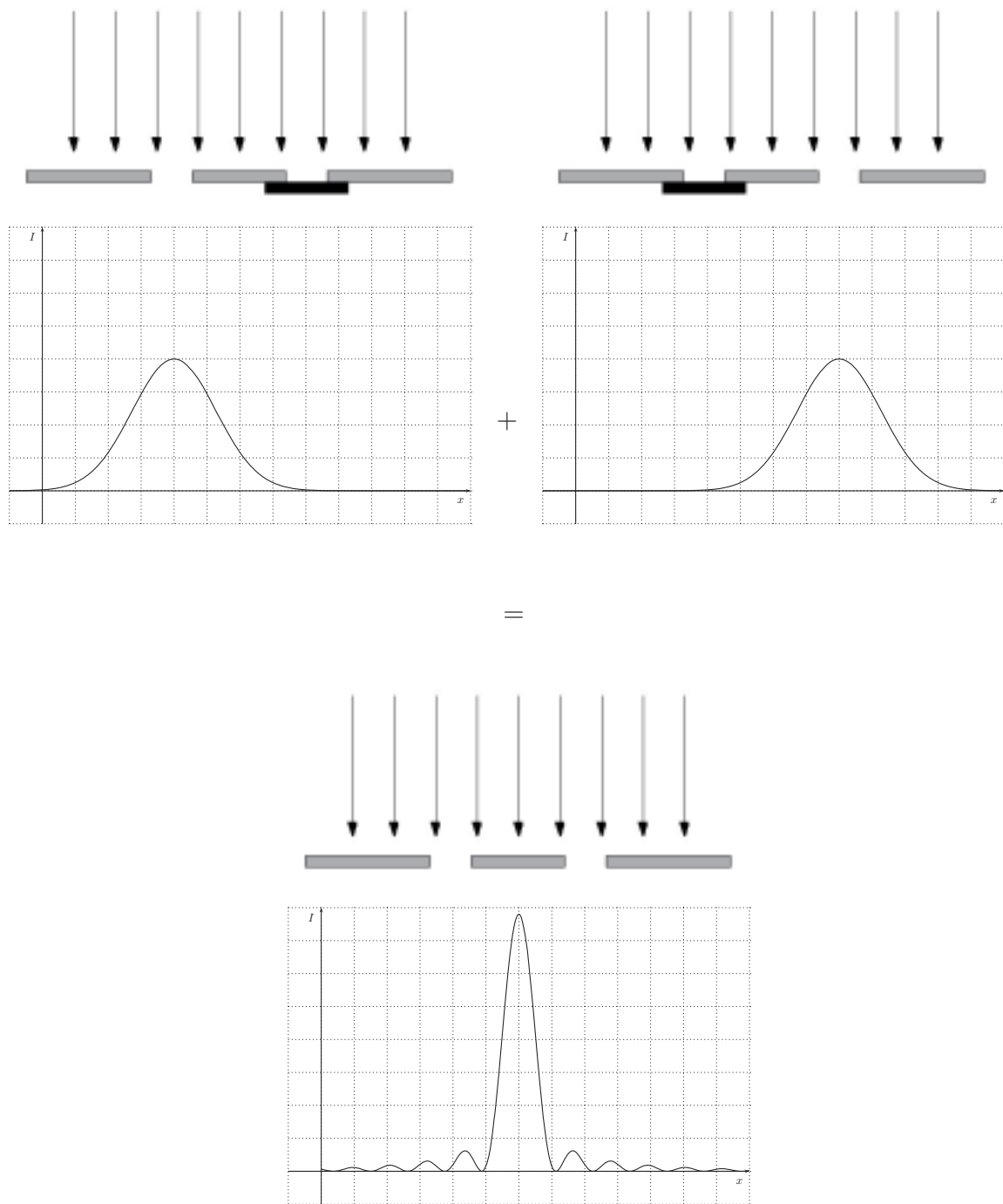
---

<sup>70</sup>Wieder müsste eigentlich wie oben beschrieben mit komplexen Wellenfunktionen gerechnet werden. Dabei ändert sich jedoch auch hier nichts grundlegendes.



$I$  = Intensität des Musters auf dem Schirm

Abbildung 4.12: Klassische Beschreibung des Doppelspaltexperiments



$I$  = Intensität des Musters auf dem Schirm

Abbildung 4.13: Quantenmechanische Beschreibung des Doppelspaltexperiments

## Aufgaben

1. Veranschauliche die quantenmechanischen Grundregeln 2, 3 und 4 anhand eines Interferenzexperiments mit einem Mach-Zehnder-Interferometer.
2. Formuliere zu jeder der sechs quantenmechanischen Grundregeln eine entsprechende Regel für die klassische Physik.
3. In der Anfangsphase der Entwicklung der Quantenmechanik versuchte man, die Veränderung des Zustands physikalischer Systeme durch quantenmechanische Messungen zu erklären, indem man davon ausging, daß das Meßgerät das System durch direkte oder indirekte Wirkung von Kräften stört. Warum ist diese Erklärung unhaltbar?

### 4.5.5 Die Heisenbergsche Unschärferelation

Wir kommen nun zu einem Thema, bei dem man wieder darüber streiten könnte, ob es überhaupt sinnvoll ist, im Physikunterricht des Gymnasiums überhaupt ausführlich darüber zu sprechen, da eine Herleitung mit den Mitteln der Schulphysik nicht durchführbar ist. Abgesehen davon, daß sich ein solcher Streit allein schon aufgrund der Präsenz des Gegenstands im Pflichtteil der Bildungspläne erübrigt<sup>71</sup>, gibt es einen weiteren triftigen Grund, sich damit zu beschäftigen, denn die Heisenbergsche Unschärferelation wird in der Schulbuchliteratur fast durchweg höchst unpräzise und sehr häufig sogar falsch dargestellt<sup>72</sup>. Schon deshalb ist es wichtig, hier für Klarheit zu sorgen.

Wie wir gesehen haben, ist es in der Quantenmechanik möglich, daß sich physikalische Systeme in Zuständen befinden, in denen sie gewisse vermeintlich wohldefinierte Eigenschaften nicht besitzen. Dies kann in gewisser Weise quantitativ beschrieben werden durch sogenannte Unschärferelationen (oder auch Unbestimmtheitsrelationen). Dazu formulieren wir zunächst das folgende Prinzip:

- Ein eindeutiger Hinweis darauf, daß sich ein System in einem Zustand befindet, in dem es eine bestimmte Eigenschaft besitzt, ist die Feststellung, daß bei vielen Messungen dieser Eigenschaft an identischen solchen Systemen (bis auf Meßfehler) immer derselbe, vorausberechenbare Meßwert herauskommt.
- Ein eindeutiger Hinweis darauf, daß sich ein System in einem Zustand befindet, in dem es eine bestimmte Eigenschaft nicht besitzt, ist die Feststellung, daß bei vielen Messungen dieser Eigenschaft an identischen solchen Systemen die Meßwerte streuen (Das hat nichts mit ungenauen Messungen zu tun!).

Daraus folgt: Die Streuung der Meßwerte beim Messen einer physikalischen Eigenschaft ist für ein quantenmechanisches System ein Maß für das Besitzen oder Nichtbesitzen dieser Eigenschaft.

<sup>71</sup>Fragen dazu tauchen beispielsweise im Abitur in Baden-Württemberg regelmäßig auf.

<sup>72</sup>Ausnahmen sind [74] und [330].

#### 4.5.5.1 Streuungen von Meßwerten

Offensichtlich ist eine präzise Definition des Begriffs der Streuung und damit ein Überblick über einige elementare statistische Grundlagen für ein Verständnis des vorliegenden Abschnitts unerlässlich. Wir beginnen dazu mit der Feststellung, daß in der Physik Meßreihen üblich sind, das heißt, man macht sehr viele Messungen an identischen Systemen. Nun seien

- $S$  ein physikalisches System,
- $B$  eine Größe, die man daran messen kann,
- $b$  eine Zufallsvariable, deren Wertemenge die möglichen Meßwerte enthält, und
- $\langle b \rangle$  der Mittelwert oder Erwartungswert bei sehr vielen solchen Messungen.

Wie kann die Streuung der Meßwerte quantitativ erfaßt werden? Naheliegender wäre die Verwendung des Mittelwerts der Größe  $b - \langle b \rangle$ , also dem Mittelwert der Abweichungen der Meßwerte vom Mittelwert. Das ist jedoch unbrauchbar, denn hierfür gilt

$$\langle b - \langle b \rangle \rangle = 0;$$

der Mittelwert verschwindet, weil die Meßwerte gleichmäßig nach oben und unten streuen. Besser ist die Betrachtung des Mittelwerts der Größe  $(b - \langle b \rangle)^2$ , also dem Mittelwert der quadratischen Abweichungen der Meßwerte vom Mittelwert, da jetzt nur noch der Betrag und nicht mehr die Richtung der Abweichungen von Bedeutung ist. Das führt auf folgende Definition:

$\mathfrak{V} = \langle (b - \langle b \rangle)^2 \rangle$  heißt *mittlere quadratische Abweichung oder Varianz der Meßwerte von B*.

Die Definitionsgleichung von  $\mathfrak{V}$  kann in eine handlichere Form gebracht werden:

$$\begin{aligned} \mathfrak{V} &= \langle (b - \langle b \rangle)^2 \rangle = \langle b^2 - 2b\langle b \rangle + \langle b \rangle^2 \rangle = \langle b^2 \rangle - \langle 2b\langle b \rangle \rangle + \langle \langle b \rangle^2 \rangle \\ &= \langle b^2 \rangle - 2\langle b \rangle^2 + \langle b \rangle^2 = \langle b^2 \rangle - \langle b \rangle^2, \end{aligned}$$

es gilt also auch

$$\mathfrak{V} = \langle b^2 \rangle - \langle b \rangle^2 \quad (4.44)$$

Die Einheit der Varianz einer Meßreihe ist natürlich das Quadrat der Einheit der Meßwerte. Da man jedoch die Streuung im allgemeinen mit der Größe der Meßwerte vergleichen will, sollte beide dieselbe Einheit haben, und man definiert weiter:

$\Delta b = \sqrt{\mathfrak{V}} = \sqrt{\langle (b - \langle b \rangle)^2 \rangle} = \sqrt{\langle b^2 \rangle - \langle b \rangle^2}$  heißt *Standardabweichung oder Streuung oder Unschärfe von B*.

Die Unschärfe einer physikalischen Größe ist die Quadratwurzel aus dem Mittelwert des Quadrats der Abweichungen der Meßwerte von deren Mittelwert bei vielen Messungen dieser Größe an identischen Systemen. Statt Unschärfe sagt man häufig auch Unbestimmtheit. Unschärfen sind Größen, die sich nur auf viele Messungen an identischen Systemen anwenden lassen; es ist sinnlos, von der Unschärfe einer Einzelmessung zu sprechen. Unschärfen sind ein Maß für Streuungen von Meßwerten.



Um den Begriff der quantenmechanischen Unschärfe einer Größe zu verdeutlichen, nehmen wir wieder einmal einen Vergleich zwischen klassischer Physik und Quantenmechanik vor.

- klassische Physik:
  - Bei Messungen können die Unschärfen der gemessenen Größen, das heißt die Streuungen der Meßwerte, im Prinzip stets beliebig klein gemacht werden.
  - Streuungen der Meßwerte sind stets Folgen von technischer Unzulänglichkeit oder mangelnder Kenntnis der Systemeigenschaften.
- Quantenmechanik:
  - Es kann vorkommen, daß bei Messungen die Unschärfen der gemessenen Größen prinzipiell nicht beliebig klein gemacht werden können, das heißt, daß die Meßwerte grundsätzlich streuen.
  - Quantenmechanische Unschärfen haben nichts mit technischen Angelegenheiten wie ungenauen Messungen oder mit teilweiser Unkenntnis des Systems zu tun, sondern sie sind eine grundlegende, unvermeidliche Eigenschaft der Natur.
  - Wenn quantenmechanischen Systemen eine bestimmte physikalische Größe nicht zugeschrieben werden kann, führen Messungen dieser Größe an solchen Systemen grundsätzlich zu Streuungen. Ein charakteristisches Beispiel hierfür sind Ortsmessungen an Elektronen in der Atomhülle, wie wir später sehen werden.

#### 4.5.5.2 Unschärferelation für Ort und Impuls

Das Auftreten quantenmechanischer Unschärfen ist für sich genommen bereits ein Sachverhalt, der sich deutlich von klassisch-physikalischem Verhalten unterscheidet. Das ist jedoch nicht alles; in der Quantenmechanik gibt es stets Paare von physikalischen Größen, deren Unschärfen sich gegenseitig beeinflussen, wenn an ein und demselben System, das heißt also an identischen Exemplaren dieses Systems, Meßreihen dieser Größen aufgenommen werden. Das bei weitem prominenteste Beispiel für ein Paar solcher Größen sind Ort und Impuls; diese waren es auch, für die Heisenberg 1927 die erste Unschärferelation gefunden hat [219]. Die allgemeine Form der Unschärferelation für beliebige Größen wurde 1929 von Robertson entdeckt [434]. Beide Herleitungen sind viel zu kompliziert für die gymnasiale Physik, weshalb wir uns hier mit einer Diskussion der Unschärferelation für Ort und Impuls begnügen und diese ohne Beweis, dafür aber in ihrer exakten Form angeben:<sup>73</sup>.

<sup>73</sup>Die folgende, von Heisenberg selbst ursprünglich erdachte „Herleitung“ der Unschärferelation für Ort und Impuls ist weit verbreitet: Man betrachtet Elektronen, die an einem Einzelspalt gebeugt werden. Läßt man viele Elektronen durch den Spalt, so beobachtet man ein Interferenzmuster. Eine Abschätzung der Ortsunschärfe am Spalt erfolgt über den Spaltdurchmesser  $d$ , das heißt, man setzt  $\Delta x \lesssim d$ . Für den Impuls des Elektrons gilt  $\vec{p} = \vec{p}_x + \vec{p}_y$  mit  $|\vec{p}| = p = h/\lambda$ . Nun wird  $\Delta p$  durch den Querimpuls abgeschätzt, der klassisch zu einem Auftreffpunkt im ersten Minimum führen würde. Ist  $\alpha$  der Winkel,

*Heisenbergsche Unschärferelation für Ort und Impuls:*

*Es ist unmöglich, ein physikalisches System in einen Zustand zu bringen, so daß bei vielen Messungen einer Ortskomponente  $x$  und vielen Messungen der zugehörigen Impulskomponente  $p_x$  an identischen solchen Systemen folgende Ungleichung verletzt wird:*

$$\Delta x \Delta p_x \geq \frac{\hbar}{2}.$$

Zunächst zwei Anmerkungen dazu:

- Die Unschärferelation ist in dieser Form mathematisch exakt herleitbar.
- Es gibt weitere Größen, für die Unschärferelationen exakt hergeleitet werden können. Entgegen einer weit verbreiteten Auffassung gilt das jedoch *nicht für Energie und Zeit*. Sämtliche in der Literatur herumgeisternde mehr oder weniger heuristische Abschätzungen für diese beiden Größen, die wie Unschärferelationen aussehen, *haben nichts mit Unschärferelationen zu tun*.<sup>74</sup>

Daraus ergeben sich wichtige Folgerungen:

- Gilt für zwei Größen eine Unschärferelation, so gehören diese Größen zu physikalischen Eigenschaften, die dem betrachteten quantenmechanischen System nicht gleichzeitig zugeschrieben werden können. Solche Größen nennt man komplementäre Größen.
- Je kleiner die Unschärfe der einen Größe ist, desto größer ist die Unschärfe der anderen Größe und umgekehrt.

Wir betrachten zwei Beispiele:

- Das erste Beispiel handelt von lokalisierten Quantenobjekten. Unter diesem Begriff sollen hier Quantenobjekte verstanden werden, die aus irgendeinem Grund gezwungen sind, sich innerhalb eines bestimmten räumlichen Bereiches aufzuhalten, z. B.

---

unter welchem dieses Minimum auftritt, so folgt mit  $\sin \alpha = \lambda/d$

$$\Delta p \approx |\vec{p}_x| = p_x = p \sin \alpha = \frac{h}{\lambda} \cdot \frac{\lambda}{d} = \frac{h}{d}$$

und damit weiter

$$\Delta x \Delta p \gtrsim h.$$

Diese Herleitung ist jedoch *physikalisch nicht haltbar*. Abgesehen vom numerisch falschen Ergebnis, das dabei herauskommt, ist die Erklärung des Interferenzmusters durch streuende Querimpulse nicht richtig. Ort, Impuls und Bahn sind Begriffe, die in dieser Situation nicht anwendbar sind. Insbesondere ist die Unschärferelation *nicht* als Folge der Störung eines Einzelobjekts durch dessen Umgebung (etwa durch Impulsübertrag) auffaß- und herleitbar, worauf wir gleich wieder zurückkommen werden.

<sup>74</sup>Die allgemeine, von Robertson bewiesene Form der Unschärferelation ist immer dann auf zwei Größen anwendbar, wenn der Kommutator der beiden Operatoren, welche die Größen repräsentieren, nicht verschwindet. Allein schon die Tatsache, daß es in der Quantenmechanik keinen Zeitoperator gibt, zeigt, daß es eine Energie-Zeit-Unschärferelation nicht geben kann.

Elektronen am Beugungsgitter. Sei z. B. die Spaltbreite des Gitters  $d = 100$  nm, so erhält man als obere Schranke für die Ortsunschärfe  $\Delta x \lesssim 100$  nm und damit

$$\Delta p \gtrsim \frac{\hbar}{2\Delta x} = 6,62 \cdot 10^{-27} \frac{\text{m kg}}{\text{s}}$$

Zur Veranschaulichung: Das entspricht einer Streuung bei Messungen der Quergeschwindigkeit unmittelbar nach dem Spalt von  $\Delta v \approx 7275$  m/s. Aus der Impulsunschärfe kann man die Unschärfe der kinetischen Energie berechnen:

$$\Delta E_{\text{kin}} = \frac{1}{2} m \Delta v \gtrsim \frac{\hbar^2}{8m(\Delta x)^2}.$$

Dafür hat die Schulphysik den Begriff „Lokalisierungsenergie“ erfunden. In obigem Beispiel ergibt das quantitativ, daß bei sehr vielen Messungen der kinetischen Energie eine Streuung der Meßwerte von  $\Delta E_{\text{kin}} \gtrsim 6,02 \cdot 10^{-23}$  J vorliegt.

- Beim zweiten Beispiel geht es um ein Auto an einer Toröffnung. Die untere Schranke  $\Delta x \lesssim 2$  m für die Ortsunschärfe liefert

$$\Delta p \gtrsim \frac{\hbar}{2\Delta x} = 6,62 \cdot 10^{-34} \frac{\text{m kg}}{\text{s}},$$

das heißt, es ergibt sich unmittelbar nach dem Tor bei vielen identischen durch das Tor durchtretenden Autos der Masse  $m = 1000$  kg eine Streuung der Quergeschwindigkeit von  $\Delta v \approx 3,31 \cdot 10^{-37}$  m/s.

Zur Unschärferelation ist eine zusätzliche Bemerkung angebracht, die eine weitere, verbreitet auftretende Fehlinterpretation betrifft. Es wird immer wieder die Meinung vertreten, die Ursache für das Auftreten von Unschärferelationen läge in der unkontrollierbaren Störung, die jede Messung am gemessenen System verursacht. Das ist nicht der Fall, was man schon an der Formulierung der Unschärferelationen selbst erkennen kann. Störungen des Systems durch die Messungen können keine Erklärung sein, weil entweder *das System immer wieder neu präpariert wird*, oder die Messungen an *verschiedenen identisch präparierten Systemen* vorgenommen werden. Solche unkontrollierbaren Störungen wären überdies äußerst merkwürdige Vorgänge. Sie müßten entweder auf irgendwelchen noch unbekanntem nichtklassischen Wechselwirkungen beruhen, wodurch man aber das Problem nur auf ein anderes ungeklärtes Phänomen abgewälzt hätte, oder auf irgendwelchen klassischen Wechselwirkungen, und diese sind, zumindest im Prinzip, nicht unkontrollierbar<sup>75</sup>. Die klassische Variante scheidet auch aus erkenntnistheoretischen Erwägungen aus. Denn die Quantenmechanik ist ein über die klassische Physik hinausgehendes und diese als Grenzfall enthaltendes Theoriegebäude. Klassische Argumentationen können daher keinesfalls dazu herangezogen werden, typische nichtklassische Resultate wie die Unschärferelation zu erklären.

#### 4.5.5.3 Das Komplementaritätsprinzip

Welle-Teilchen-Dualismus einerseits und Unschärferelationen andererseits haben nichts miteinander zu tun. Sie sind voneinander unabhängige quantenmechanische Naturgesetze. Keines von beiden läßt sich aus dem jeweils anderen herleiten. Beide sind jedoch

<sup>75</sup>Eine ausführliche Diskussion hierzu findet man in [357].

Beispiele für ein die gesamte Quantenmechanik durchziehendes erkenntnistheoretisches Prinzip, das

*Komplementaritätsprinzip:*

*Die vollständige Beschreibung der Ergebnisse von Beobachtungen an quantenmechanischen Systemen kann nicht in einem begrifflich einheitlichen Bild zusammengefaßt werden. Erst die Gesamtheit der sich aus klassischer Sicht im allgemeinen gegenseitig widersprechenden Beobachtungsergebnisse liefert die vollständigen möglichen Informationen über quantenmechanische Objekte.*

Entsprechend nennt man zwei physikalische Begriffe *komplementär*, wenn sie einander einerseits aus klassischer Sicht ausschließen, sie aber andererseits zur vollständigen Beschreibung des betrachteten Systems beide zusammen erforderlich sind.

- Beispiel Welle-Teilchen-Dualismus:

Wellenähnliches und teilchenähnliches Verhalten schließen einander aus klassischer Sicht gegenseitig aus, dennoch sind beide gemeinsam erforderlich, um beispielsweise Experimente an Zwei-Wege-Interferometern deuten zu können.

- Beispiel Unschärferelation:

Die Ergebnisse von an identisch präparierten Objekten vorgenommenen Ortsmessungen einerseits und Impulsmessungen andererseits schließen sich gegenseitig aus, dennoch sind beide gemeinsam für ein Verständnis des betrachteten Systems erforderlich. Häufig nennt man daher auch zwei physikalische Größen, für die eine nicht verschwindende Unschärferelation formuliert werden kann (wie z.B. Ort und Impuls), komplementäre Größen.

Hier taucht jedoch ein Problem auf: Das Komplementaritätsprinzip versucht, verschiedene Dinge, die nichts miteinander zu tun haben, unter einen gemeinsamen Hut zu bringen. Daran ist erkennbar, daß es sich dabei *nicht* um ein physikalisches Gesetz handeln kann. Das Komplementaritätsprinzip ist lediglich eine erkenntnistheoretische Aussage:

*Das Komplementaritätsprinzip bringt die prinzipielle Unmöglichkeit zum Ausdruck, quantenmechanische Phänomene auf eine der klassischen Physik entsprechende Art und Weise zu beschreiben.*

Das ist eigentlich nichts Neues und wiederholt nur unsere bereits getroffene Aussage: Eine Rückkehr zur klassischen Physik ist nicht möglich. Die notwendigerweise unanschaulichen Methoden der Quantenmechanik sind zur Beschreibung quantenmechanischer Systeme unvermeidlich. Es sei erneut darauf hingewiesen, daß dies nichts mit technischen oder sonstigen Unzulänglichkeiten zu tun hat. Es handelt sich dabei vielmehr um eine prinzipielle, nicht überschreitbare Grenze unserer Erkenntnisfähigkeit.

Besonders wichtig ist folgende Zusatzbemerkung: Jegliche Analogien zu angeblich „komplementären“ Begriffspaaren außerhalb der Quantenmechanik (wie z.B. „Komplementarität“ von gefühlsmäßigem und analytischem Denken, Individualität und Gemeinschaftlichkeit oder ähnlichem Unsinn) sind *abwegig* und *unbedingt zu vermeiden*. Sie

tauchen sehr häufig auf und sind keineswegs nur in populärwissenschaftlichen Veröffentlichungen zu finden. Die Verwirrungen, die mit dem Begriff der Komplementarität seit seinem Auftauchen einhergehen, sind durch seine uneinheitliche, unklare und zum Teil auch esoterische Handhabung verursacht. Es gibt sehr wohl auch eine physikalisch exakte Definition des Begriffs der Komplementarität für physikalische Größen. Zwei Größen heißen komplementär, wenn bei jedem System, das die streuungsfreie Messung der einen Größe gestattet, die Messung der anderen Größe zu beliebiger Streuung führt und umgekehrt. Das Komplementaritätsprinzip besagt in diesem Zusammenhang, daß es bei jedem physikalischen System zu jedem Freiheitsgrad ein Paar komplementärer Größen gibt<sup>76</sup>. Taucht der Begriff der Komplementarität in der Literatur auf, ist damit jedoch sehr häufig nicht diese exakte Definition, sondern irgend etwas anderes diffuses, nicht so recht definiertes gemeint.

## Aufgaben

1. Elfmeterschießen kann als Ortsmessung an Fußbällen aufgefaßt werden, wobei im Falle eines verwandelten Elfmeters die Ortsunschärfe in Querrichtung durch die Breite eines Fußballtores von 7,32 m begrenzt wird. Zeige, daß sich der Torwart bei ständig nichtgehaltenen Elfmetern nicht quantenmechanisch mit der Unschärfe der Geschwindigkeitskomponenten der Fußbälle in Querrichtung und der angeblich daraus resultierenden Unberechenbarkeit ihrer Flugbahnen herausreden kann.
2. Berechne die aufgrund des typischen Radius eines Atoms von ca.  $10^{-10}$  m zu erwartende Streuung der Ergebnisse bei Geschwindigkeitsmessungen an Elektronen in der Atomhülle.
3. Warum ist es nicht möglich, mit Hilfe der Heisenbergschen Unschärferelation eine Aussage über die zu erwartende Genauigkeit einer einzelnen Ortsmessung an einem Elektron zu machen?
4. Man liest oder hört häufig folgende Aussage:  
 „Die Heisenbergsche Unschärferelation besagt, daß es nicht möglich ist, den Ort und den Impuls eines quantenmechanischen Systems gleichzeitig beliebig scharf zu messen.“  
 Beurteile diese Aussage hinsichtlich ihrer physikalischen Korrektheit und verbessere sie gegebenenfalls.

### 4.5.6 Einfache quantenmechanische Systeme

Nachdem wir uns nun seit vier Abschnitten im wesentlichen mit formalen Sachverhalten beschäftigen, ist es an der Zeit, das Gelernte auch auf konkrete physikalische Systeme anzuwenden. Natürlich muß man sich in der Schulphysik auf die allereinfachsten Fälle beschränken. Das heißt jedoch nicht, daß dabei nicht wesentliche Aspekte der quantenmechanischen Naturbeschreibung deutlich würden.

---

<sup>76</sup>Siehe Abschnitt 4.3.1.

### 4.5.6.1 Beschreibung quantenmechanischer Systeme

Auf der Ebene der Physik in der Oberstufe muß man sich damit begnügen, zur Beschreibung quantenmechanischer Systeme die stationäre Schrödingergleichung in ihrer einfachsten, eindimensionalen Form zu betrachten. Selbst das führt sehr schnell auf rechnerisch sehr aufwendige Probleme, sodaß man dabei nochmals eine sehr restriktive Auswahl treffen muß.

Die Arbeitsschritte sind beim hier beschriebenen Zugang stets dieselben:

1. Aufstellen der Schrödingergleichung
2. Berechnen der Wellenfunktionen durch Lösen der Schrödingergleichung
3. Heraussortieren der brauchbaren Lösungen und Bestimmen der Energieeigenwerte durch Randbedingungen und Normieren der Wellenfunktionen

Dazu sind wohl jeweils einige Anmerkungen erforderlich. Das Aufstellen der Schrödingergleichung

$$-\frac{\hbar^2}{2m} \psi'' + V \psi = E \psi;$$

ist häufig dann einfach, wenn das zu beschreibende System ein klassisches Analogon besitzt – und das ist bei den in der Schule zugänglichen Beispielen durchweg der Fall. Die Aufgabe besteht lediglich darin, das richtige Potential  $V$  zu finden. Hierbei gilt folgende Regel:

*Für  $V$  ist derjenige Term einzusetzen, der die potentielle Energie des Systems liefern würde, wenn man es nicht quantenmechanisch, sondern klassisch beschreiben würde.*

Gibt es kein klassisches Analogon, so ist man auf geschicktes Raten angewiesen.

Beim Lösen der nunmehr aufgestellten Schrödingergleichung hat man es mit den üblichen Anforderungen beim Lösen von Differentialgleichungen zu tun. Wir werden das nur im einfachsten der hier diskutierten Fälle wirklich vom Anfang bis zum Schluß durchexerzieren und uns bei den weiteren Beispielen darauf beschränken, vorgegebene Lösungen darauf zu überprüfen, ob sie tatsächlich welche sind.

Im allgemeinen gibt es unendlich viele mathematisch möglichen Lösungen der betrachteten Schrödingergleichung; da jedoch nicht alle davon auch physikalische, dem betrachteten Problem angemessene Lösungen sind, muß man letztere aus den ersten herausuchen. Das liefert gleichzeitig eine Relation, mit der die möglichen Werte der Gesamtenergie des Systems festgelegt werden. Dabei treten bereits in den einfachsten Fällen Integrale auf, die mit den inzwischen stark eingeschränkten Möglichkeiten der Oberstufenmathematik nicht mehr berechnet werden können. Es obliegt der pädagogischen Freiheit der Fachlehrerin oder des Fachlehrers, zu entscheiden, ob hier ein mathematischer Exkurs lohnend ist, oder ob man sich damit begnügt, die erforderlichen Resultate nur anzugeben.

### 4.5.6.2 Der lineare Potentialtopf

Wir beginnen mit dem einfachsten nichttrivialen quantenmechanischen System überhaupt. Das Ziel ist die Beschreibung eines Teilchens, das auf einer Geraden auf den Bereich  $0 \leq x \leq a$  beschränkt ist. Diese Aufgabenstellung ist nicht nur von rein akademischem Interesse; eine Anwendung ist beispielsweise die modellmäßige Beschreibung von Valenzelektronen in langgestreckten Kettenmolekülen. Das Teilchen soll den erlaubten Bereich nicht verlassen können. Das Potential ist folglich von der Form

$$V(x) = \begin{cases} 0 & \text{für } 0 \leq x \leq a, \\ \infty & \text{für } x < 0 \text{ oder } x > a. \end{cases}$$

Man kann sich dabei anschaulich vorstellen, das Teilchen sei in einem Kasten der Breite  $a$  mit undurchdringlichen Wänden eingesperrt (Abbildung 4.14). Für  $x < 0$  und für  $x > a$  gilt  $\psi = 0$ . Für  $0 \leq x \leq a$  lautet die Schrödingergleichung

$$-\frac{\hbar^2}{2m} \psi'' = E \psi$$

oder

$$\psi'' = -\frac{2mE}{\hbar^2} \psi.$$

Mit der Abkürzung  $k^2 = 2mE/\hbar^2$  schreiben wir dafür

$$\psi'' = -k^2 \psi.$$

Die Schrödingergleichung des linearen Potentialtopfs mit unendlich hohen Wänden ist äquivalent zur Differentialgleichung des klassischen harmonischen Oszillators. Die Lösung kann daher formal völlig analog dazu erfolgen. Als Ansatz schreiben wir entsprechend

$$\psi(x) = A \sin(kx + \delta).$$

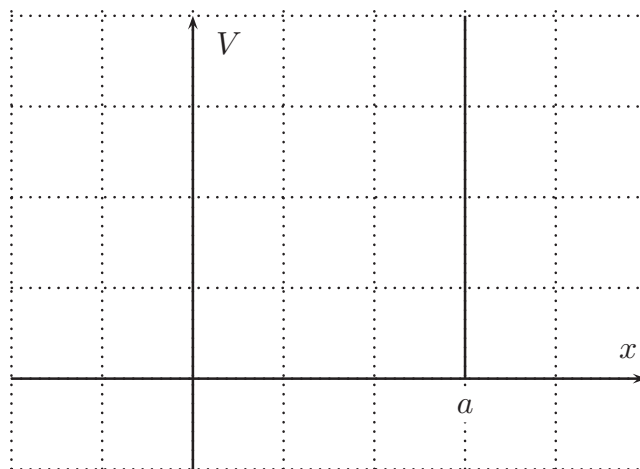


Abbildung 4.14: Der lineare Potentialtopf mit unendlich hohen Wänden



Als nächstes bestimmen wir die genaue Form der brauchbaren Lösungen. Außerhalb des erlaubten Bereiches gilt  $V = \infty$ . Damit erhalten wir die Randbedingungen  $\psi(0) = 0$  und  $\psi(a) = 0$ . Aus der ersten Randbedingung folgt  $\delta = 0$ , aus der zweiten Randbedingung  $k = n\pi/a$  mit  $n \in \mathbb{N} \setminus \{0\}$ . Den hier auftretenden Parameter  $n$  nennt man *Quantenzahl*. Die Wellenfunktionen lauten damit

$$\psi_n(x) = \begin{cases} A \sin\left(\frac{n\pi x}{a}\right) & \text{für } 0 \leq x \leq a, \\ 0 & \text{für } x < 0 \text{ oder } x > a. \end{cases}$$

mit einer noch zu bestimmenden Normierungskonstante  $A$ , und die Gesamtenergie ist

$$E = \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2 \pi^2 n^2}{2m a^2}, \quad n \in \mathbb{N} \setminus \{0\}.$$

Man nennt diese Gleichung die *Eigenwertgleichung der Gesamtenergie*. Damit haben wir einen Sachverhalt gefunden, der exemplarisch für alle gebundenen quantenmechanischen Systeme ist: Die Gesamtenergie ist *quantisiert*. Es gibt keine kontinuierliche Energieverteilung, sondern diskrete Energieniveaus, festgelegt durch die Quantenzahl  $n$ .

Nun kommen wir zur Normierung der Wellenfunktionen  $\psi_n$ . Die Forderung

$$\int_{-\infty}^{+\infty} |\psi_n|^2 dx = 1$$

führt hier speziell auf

$$\int_0^a |\psi_n|^2 dx = \int_0^a |A|^2 \sin^2 \frac{n\pi x}{a} dx = 1$$

Die Berechnung des hier auftretenden Integrals erledigen wir durch dreifache partielle Integration; das liefert zunächst

$$\begin{aligned} \int \sin^2(bx) dx &= -\frac{1}{b} \sin(bx) \cos(bx) - \int \cos^2(bx) dx \\ &= -\frac{1}{b} \sin(bx) \cos(bx) - \int [1 - \sin^2(bx)] dx \\ &= -\frac{1}{b} \sin(bx) \cos(bx) + \int dx - \int \sin^2(bx) dx, \end{aligned}$$

damit also

$$2 \int \sin^2(bx) dx = \int dx - \frac{1}{b} \sin(bx) \cos(bx),$$

und folglich

$$\begin{aligned} \int \sin^2(bx) dx &= \frac{1}{2} \int dx - \frac{1}{2b} \sin(bx) \cos(bx) \\ &= \frac{x}{2} - \frac{1}{4b} [\sin(bx) \cos(bx) + \sin(bx) \cos(bx)]. \end{aligned}$$

Anwendung eines der Additioinstheoreme der trigonometrischen Funktionen, nämlich

$$\sin \alpha \cos \beta + \sin \beta \cos \alpha = \sin (\alpha + \beta),$$

liefert schließlich

$$\int \sin^2(bx) dx = \frac{x}{2} - \frac{1}{4b} \sin(2bx)$$

Für das bestimmte Integral ergibt sich somit

$$\int_0^a |\psi_n|^2 dx = |A|^2 \left[ \frac{x}{2} - \frac{n\pi}{4a} \sin \left( \frac{2n\pi x}{a} \right) \right]_0^a = |A|^2 \frac{a}{2}$$

und für die Normierungsbedingung

$$|A|^2 \frac{a}{2} = 1.$$

Für die Normierungskonstante erhalten wir folglich

$$A = \sqrt{\frac{2}{a}}$$

und für die Wellenfunktionen

$$\psi_n(x) = \begin{cases} \sqrt{\frac{2}{a}} \sin \left( \frac{n\pi x}{a} \right) & \text{für } 0 \leq x \leq a, \\ 0 & \text{für } x < 0 \text{ oder } x > a. \end{cases}$$

Die Wahrscheinlichkeitsdichte für das Antreffen des Teilchens im Potentialtopf ist somit

$$|\psi_n(x)|^2 = \frac{2}{a} \sin^2 \left( \frac{n\pi x}{a} \right).$$

Der tiefste Energieeigenzustand heißt *Grundzustand*, die höheren Zustände nennt man *angeregte Zustände*.

Das hier beschriebene System weist trotz seiner Einfachheit typisch quantenmechanisches Verhalten auf. Insbesondere ist die immer wieder vertretene Auffassung, man könne sich dabei ein im Kasten hin- und herschwingendes Teilchen vorstellen, völlig unhaltbar. Im Detail gilt folgendes:

- Es gibt Stellen, wo die Wahrscheinlichkeit, das Teilchen anzutreffen, besonders groß ist, aber auch Stellen, wo diese Wahrscheinlichkeit Null ist. Das widerspricht in beliebigem Maß jeglicher Vorstellung von Bewegung irgendwelcher Art.
- Dem Teilchen im Potentialtopf kann man die Eigenschaften, einen bestimmten Ort und einen bestimmten Impuls zu besitzen, nicht zuschreiben.
- Entsprechend kann man ihm auch die Eigenschaft, potentielle und kinetische Energie zu besitzen, nicht zuschreiben.

- Das Teilchen hat eine wohlbestimmte Gesamtenergie. Deshalb sagt man, es befindet sich in einem Energieeigenzustand.

Da die Energieniveaus im Potentialtopf nur feste, diskrete Werte annehmen können, gilt das selbe auch für die Energiedifferenzen zwischen beliebigen solchen Niveaus. Eine wichtige Konsequenz der diskreten Energieniveaus ist daher die Möglichkeit von *Energieübergängen*. Ein Teilchen im Potentialtopf kann Energie genau dann absorbieren, wenn sie exakt in einer Menge ankommt, die einer Differenz zwischen zwei Energieniveaus entspricht. Für diese Energiedifferenzen gilt die Relation

$$\Delta E_{n_1, n_2} = E_{n_2} - E_{n_1} = \frac{\hbar^2 \pi^2}{2m a^2} (n_2^2 - n_1^2).$$

Anschließend kann das Teilchen diese Energie wieder in Form von elektromagnetischer Strahlung abgeben; wegen  $E = h\nu$  hat diese die Frequenz

$$\nu_{n_1, n_2} = \frac{\Delta E_{n_1, n_2}}{h} = \frac{\hbar \pi}{4m a^2} (n_2^2 - n_1^2).$$

Eine wichtige Anwendung dieses Modells ist die Beschreibung der Verwendung bestimmter Kettenmoleküle als *Farbstoffe*. Diese Moleküle absorbieren vermöge eines Ener-

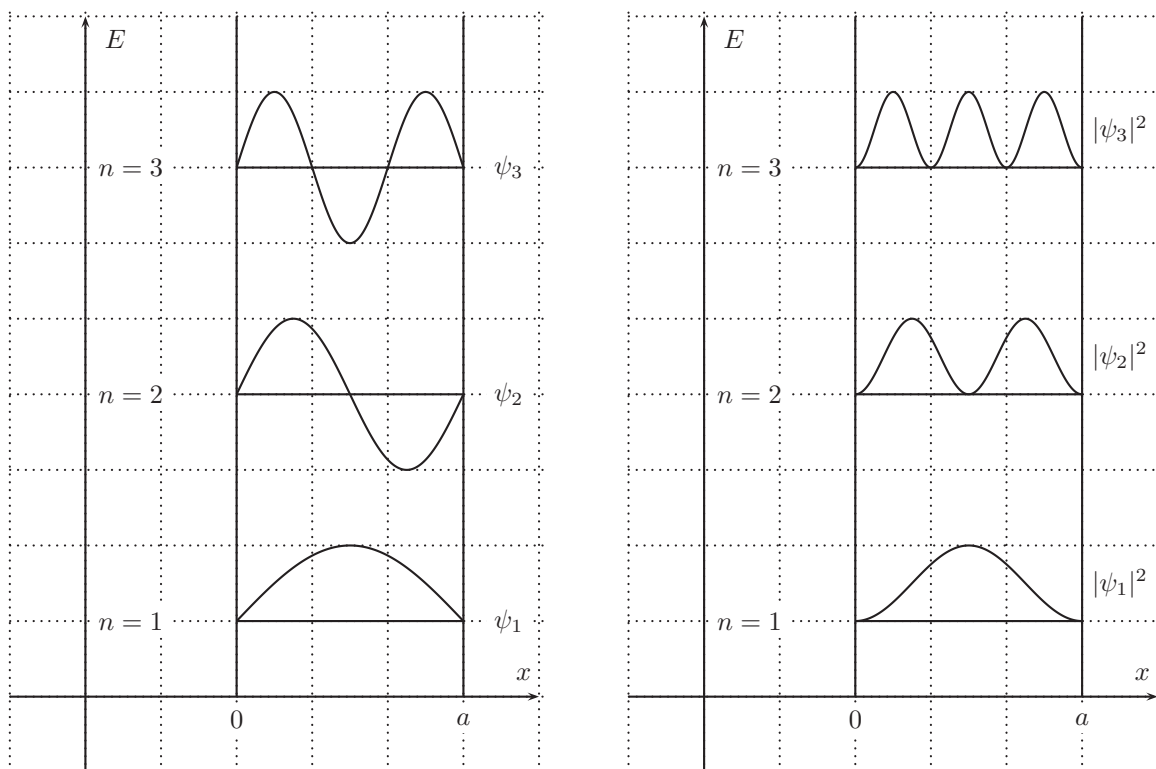


Abbildung 4.15: Wellenfunktionen (links) und Aufenthaltswahrscheinlichkeitsdichten (rechts) für die ersten drei Eigenzustände beim linearen Potentialtopf mit unendlich hohen Wänden

gieübergangs einen Teil des sichtbaren Lichts und reflektieren den Rest. Dieses Licht erscheint dann in der Komplementärfarbe des absorbierten Lichts<sup>77</sup>.

#### 4.5.6.3 Das Wasserstoffatom

Nun kommen wir zum wahrscheinlich prominentesten Beispielsystem der elementaren Quantenmechanik. Allerdings läßt sich dieses im Gegensatz zum vorigen Beispiel mit den Mitteln der Schulphysik nicht mehr komplett durchrechnen. Dennoch kann man dabei auch in der Schule bereits zu wesentlichen quantitativen Resultaten gelangen.

Wir beginnen mit einem historischen Überblick über *Atommodelle*:

- 400 v. Chr.: Atommodell von Demokrit  
Dieses Modell geht eigentlich auf Demokrits Lehrer Leukipp zurück. Danach besteht alle Materie aus nicht weiter zerlegbaren winzigen Bausteinen, den Atomen.
- 1800: Gesetz der konstanten und multiplen Proportionen von J. Dalton  
Bei chemischen Reaktionen stehen die Massenverhältnisse zweier Elemente, die sich zu verschiedenen chemischen Verbindungen vereinigen, stets im Verhältnis ganzer Zahlen zueinander. Das ist durch die Atomhypothese erklärbar.
- 1897: Rosinenkuchenmodell von J. J. Thomson  
Das Atom wird als kontinuierlich verteilte Wolke positiver Ladung mit darin verteilten punktförmigen negativ geladenen Elektronen betrachtet.
- 1908: Planetensystemmodell von E. Rutherford  
Fast die gesamte Masse der Atome befindet sich in ihren positiv geladenen Atomkernen, der Rest verteilt sich auf die Elektronen. Die Elektronen kreisen um den Atomkern wie Planeten um einen Stern.

Rutherford konnte mit seinem berühmten Streuexperiment unter anderem die Größenverhältnisse in den Atomen klären: Der typische Atomdurchmesser beträgt einige  $10^{-10}$  m, der typische Atomkerndurchmesser einige  $10^{-15}$  m.

Das Rutherfordsche Atommodell führt jedoch auf ein unlösbares Problem, denn Elektronen auf Kreisbahnen sind beschleunigte elektrische Ladungen und würden ständig elektromagnetische Wellen und damit Energie abgeben. Als Folge müßten die Elektronen in den Kern stürzen. Das ist ein Widerspruch zur offensichtlichen Stabilität der Atome. Außerdem zeigen Versuche mit Atomen, daß Elektronen in der Atomhülle nicht beliebige, sondern stets nur bestimmte, diskret verteilte, von der Atomsorte abhängige Energiewerte besitzen. Beispielsweise absorbiert jedes Element nur Licht ganz bestimmter, materialabhängiger Frequenzen und läßt alle anderen durch. Davon kann man sich mit dem *Franck-Hertz-Versuch* überzeugen [175], [176]. Bei diesem beobachtet man die Anregung von Quecksilberatomen durch Elektronenbeschuß (siehe Abbildung 4.16). Man stellt fest, daß die Stromstärke zwischen Gitteranode und Kathode bei  $U_B = 4,9\text{ V}, 9,8\text{ V}, 14,7\text{ V}, \dots$  regelmäßig stark ansteigt und dazwischen jeweils deutlich abfällt (siehe Abbildung 4.17).

---

<sup>77</sup>Vergleiche Aufgabe 2.

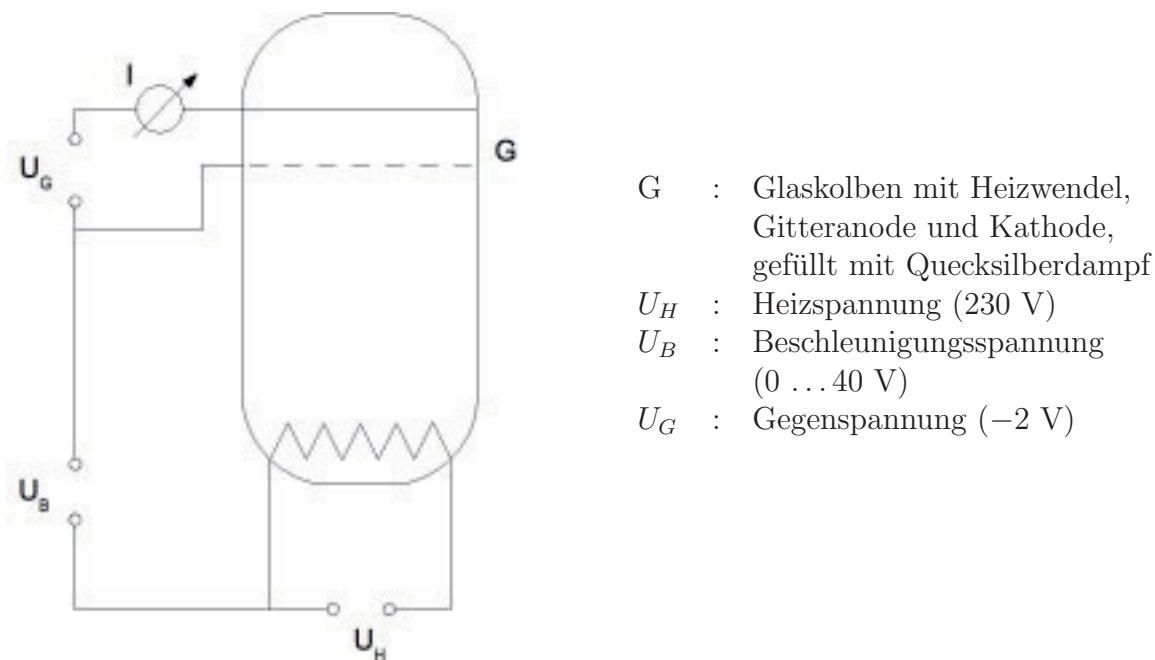


Abbildung 4.16: Der Franck-Hertz-Versuch

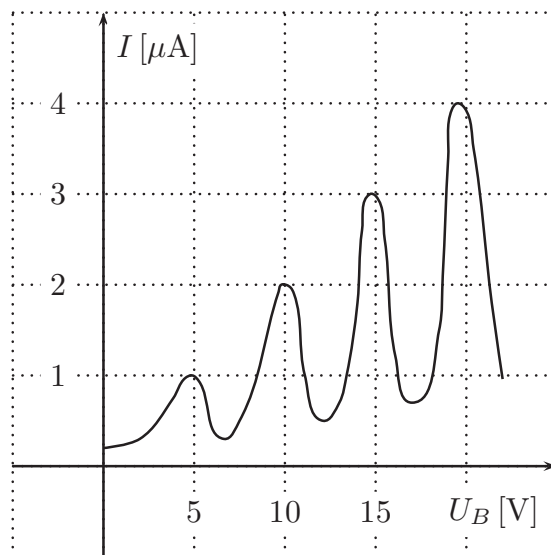


Abbildung 4.17: Strom-Spannungs-Verlauf beim Franck-Hertz-Versuch

Das läßt sich folgendermaßen erklären: Bei genügend großer Beschleunigungsspannung können von der Glühkathode abgegebene, durch die Gitteranode hindurchtretende Elektronen die Gegenspannung zunächst überwinden. Ab einer Beschleunigungsspannung von mehr als 4,9 V können sie ein- oder mehrmals Energie an Elektronen der Quecksilberatome abgeben, aber stets nur in ganzzahligen Vielfachen von 4,9 V.

Wir schließen daraus, daß die Elektronen in der Atomhülle nur ganz bestimmte, diskrete Energieniveaus besetzen können. Bei Energiezufuhr können die Elektronen in einen energetisch höheren Zustand übergehen, dabei nehmen sie jedoch nur genau die entsprechende Energiedifferenz der beiden Zustände auf. Kehren sie in einen energetisch tieferen Zustand zurück, geben sie genau die entsprechende Energie wieder ab, z. B. durch Aussenden eines Photons.

1913 erschien eine neue Beschreibung des Wasserstoffatoms von Niels Bohr [64]<sup>78</sup>. Das Bohrsche Atommodell „erklärt“ durch Ad-Hoc-Annahmen die Stabilität der Atome und erlaubt die Berechnung der Energieniveaus der Wasserstoffatomhülle. Das ist jedoch keine Erklärung, denn die Schwierigkeiten werden einfach „per Erlaß“ ausgeblendet. Außerdem verwendet das Bohrsche Atommodell in willkürlicher Weise klassische Physik und deren Außerkräftsetzen gleichzeitig. Das größte Problem daran ist jedoch das Festhalten an Elektronenbahnen. Wie wir gesehen haben, kann man Quantenobjekten, die sich in Energieeigenzuständen befinden, die Eigenschaft, Orte und Geschwindigkeiten zu besitzen, nicht zuschreiben. *Die Vorstellung von Elektronenbahnen in der Atomhülle ist daher abwegig*<sup>79</sup>. Damit stellt sich die Frage: Wie kann man die Stabilität der Atome und die diskreten Energieniveaus physikalisch sauber erklären? Man erreicht das durch eine *quantenmechanische Beschreibung des Wasserstoffatoms*<sup>80</sup>. Dazu müssen wir wieder das oben beschriebene Drei-Punkte-Programm abspulen.

Wir stellen zunächst die Schrödingergleichung auf. Klassisch ist die potentielle Energie einer Ladung  $e$  im Coulomb-Feld einer zweiten, gleich großen Ladung  $e$  mit umgekehrtem Vorzeichen, die sich im Abstand  $r$  von der ersten befindet, durch

$$E_{\text{pot}} = -\frac{e^2}{4\pi\epsilon_0 r},$$

gegeben, folglich wählen wir als Potential den Ausdruck

$$V = -\frac{e^2}{4\pi\epsilon_0 r}$$

(Abbildung 4.18). Da es sich um ein räumliches System handelt, hat man es mit einer dreidimensionalen Schrödingergleichung zu tun; diese lautet

$$-\frac{\hbar^2}{2m} \left( \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2} \right) \psi - \frac{e^2}{4\pi\epsilon_0 r} \psi = E \psi,$$

<sup>78</sup>Bohr publizierte noch im selben Jahr umfangreiche Erweiterungen seiner Theorie [65], [66].

<sup>79</sup>Selbst ohne dieses moderne quantenmechanische Argument ist das Bohrsche Atommodell unhaltbar, denn Keplerbahnen liegen stets in einer Ebene, so daß man das absurde Bild scheibenförmiger Atome erhält. Ein weiteres, für die Schulphysik irrelevantes, aber deshalb nicht minder schweres Problem ist die Tatsache, daß das Bohrsche Atommodell nur bei wasserstoffähnlichen Atomen quantitativ brauchbare Ergebnisse liefert und auf andere Atome nicht übertragbar ist. Sommerfeld entwickelte 1916 eine relativistische Verallgemeinerung, die das Modell um elliptische und räumlich geneigte Bahnen erweiterte [485], [486]. Das Sommerfeldsche Atommodell liefert zwar spektakulär genau quantitative Resultate, die konzeptionellen und praktischen Probleme seines Vorläufers bleiben dabei jedoch erhalten. Die bleibende Bedeutung der Atommodelle von Bohr und Sommerfeld ist wesentlich eine historische, da sie die ersten nichtklassischen Versuche darstellten und damit der unumgänglichen Abkehr von der klassischen Physik den Weg bahnten. Beim Sommerfeldschen Atommodell kommt noch dessen Subtilität hinzu, die es aus mathematischer Sicht auch heutzutage noch überaus interessant erscheinen lassen.

<sup>80</sup>Eine solche gelang zuerst 1926 Erwin Schrödinger [454] – [458].

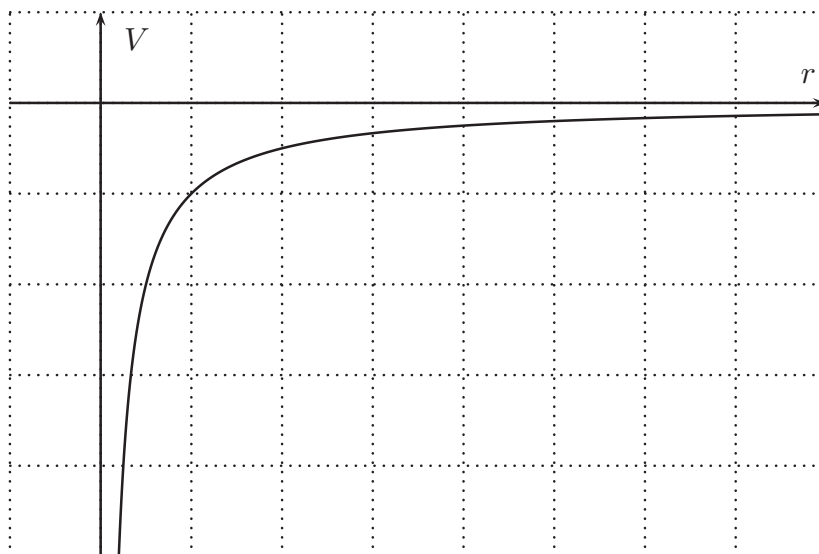


Abbildung 4.18: Potentialverlauf beim Wasserstoffatom

mit  $r = \sqrt{x_1^2 + x_2^2 + x_3^2}$ . Außerdem ist das Potential kugelsymmetrisch, das heißt, der Wert von  $V$  hängt nur vom Abstand zur felderzeugenden Ladung ab. Daher empfiehlt sich der Übergang von kartesischen Koordinaten  $x_1, x_2, x_3$  zu Kugelkoordinaten  $r, \theta, \phi$  (Abbildung 4.19).

Die allgemeine Lösung der Schrödingergleichung für das Wasserstoffatom ist mit schulischen Mitteln nicht machbar. Man kann jedoch zeigen, daß alle Lösungen die Form eines Produkts aus einer nur von  $r$ , einer nur von  $\theta$  und einer nur von  $\phi$  abhängigen Funktion haben, sodaß diese jeweils getrennt betrachtet werden können. Insbesondere gibt es spezielle Zustände, die selbst auch kugelsymmetrisch sind (sogenannte  $s$ -Zustände). Die Wellenfunktion  $\psi$  hängt für solche Zustände nur von  $r$  ab. Das bedeutet, daß  $\psi$  auf jeder Kugelschale konstant ist. Wir beschränken uns auf solche kugelsymmetrischen Zustände.

Der von  $r$  abhängige Anteil der zweiten Ableitungen von  $\psi$  lautet  $\frac{d^2\psi}{dr^2} + \frac{2}{r} \frac{d\psi}{dr}$ . (Auch hier ist der Nachweis mit schulischen Mitteln nicht erbringbar). Damit erhält man für den hier ausschließlich betrachteten radialen Anteil der Schrödingergleichung

$$-\frac{\hbar^2}{2m} \left( \psi''(r) + \frac{2}{r} \psi'(r) \right) - \frac{e^2}{4\pi\epsilon_0 r} \psi(r) = E \psi(r)$$

Als Ansatz für den Grundzustand wählen wir

$$\psi(r) = B e^{-r/a}$$

(mit schulischen Mitteln ebenfalls nicht ohne weiteres begründbar). Die ersten beiden Ableitungen lauten

$$\psi'(r) = -\frac{B}{a} e^{-r/a} = -\frac{\psi(r)}{a}, \quad \psi''(r) = \frac{B}{a^2} e^{-r/a} = \frac{\psi(r)}{a^2}.$$



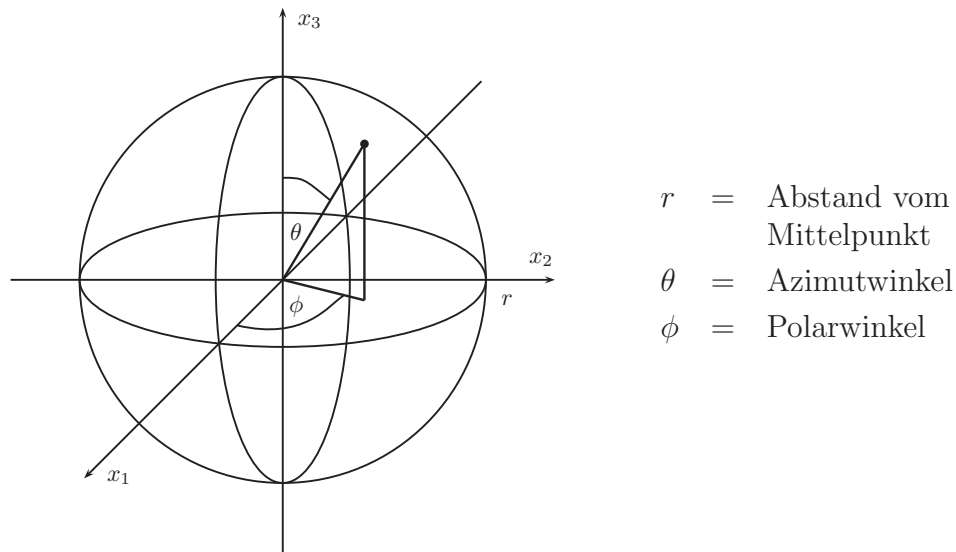


Abbildung 4.19: Zur Einführung von Kugelkoordinaten

Einsetzen in Schrödingergleichung liefert

$$-\frac{\hbar^2}{2m} \left( \frac{\psi(r)}{a^2} - \frac{2}{r} \frac{\psi(r)}{a} \right) - \frac{e^2}{4\pi\epsilon_0 r} \psi(r) = E \psi(r)$$

und beiderseitiges Dividieren durch  $\psi(r) \neq 0$  ergibt

$$\frac{2}{ar} - \frac{1}{a^2} = \frac{2m}{\hbar^2} \left( E + \frac{e^2}{4\pi\epsilon_0 r} \right)$$

Aus dieser Gleichung lassen sich durch Koeffizientenvergleich erste Resultate gewinnen: Der Ansatz kann nur stimmen, wenn die Terme mit  $1/r$  und die Terme ohne  $1/r$  jeweils gleich sind.

(i) Vergleichen der Summanden mit  $1/r$  liefert

$$\frac{2}{ar} = \frac{2me^2}{4\pi\hbar^2\epsilon_0 r}$$

und damit

$$a = \frac{4\pi\epsilon_0\hbar^2}{m_e e^2} = 0,053 \text{ nm.}$$

$a$  heißt Bohrscher Radius; diese Konstante ist ein Maß für typische Längenskalen in der Atomphysik.

(ii) Vergleichen der Summanden ohne  $1/r$  liefert

$$\frac{1}{a^2} = \frac{2mE}{\hbar^2},$$

was auf

$$E = \frac{\hbar^2}{2ma^2} = -\frac{m_e e^4}{32\pi^2\epsilon_0^2\hbar^2} = -13,6 \text{ eV}$$

führt. Damit haben wir die Grundzustandsenergie des Wasserstoffatoms hergeleitet. Der Zahlenwert ist experimentell sehr gut bestätigt.  $E_1 = -13,6$  eV ist diejenige Energie, die man benötigt, um ein Elektron im Grundzustand aus der Atomhülle des Wasserstoffatoms zu befreien, das heißt, um das Atom zu ionisieren.  $E_1$  heißt deshalb auch Ionisierungsenergie des Wasserstoffatoms.

Als nächstes kommen wir zur Normierung der Wellenfunktion. Da es sich um ein dreidimensionales System handelt, lautet die Normierungsbedingung hier

$$\int |\psi(r)|^2 dV = 1.$$

Das ist ein Raumintegral; so etwas kann man normalerweise mit Schulmathematik nicht berechnen. Aufgrund der Kugelsymmetrie kann man das Raumintegral in diesem Fall jedoch auf ein einfaches Integral zurückführen. Dazu wählen wir für die Raumelemente  $dV$  Kugelschalen der Dicke  $dr$  und der Oberfläche  $4\pi r^2$ , also

$$dV = 4\pi r^2 dr.$$

Die Normierungsbedingung erhält damit die Gestalt

$$4\pi \int_0^{\infty} |\psi(r)|^2 r^2 dr = 1.$$

Wir werten das Integral mit Hilfe von partieller Integration aus und erhalten

$$\begin{aligned} \int_0^{\infty} |\psi(r)|^2 r^2 dr &= \int_0^{\infty} |B e^{-r/a}|^2 r^2 dr = B^2 \int_0^{\infty} r^2 e^{-2r/a} dr \\ &= -\frac{1}{2} a B^2 \int_0^{\infty} r e^{-r/a} dr = \frac{1}{4} a^2 B^2 \int_0^{\infty} e^{-r/a} dr = \frac{1}{4} a^3 B^2. \end{aligned}$$

Mit der Normierungsbedingung ergibt sich hieraus

$$\pi a^3 B^2 = 1$$

und folglich für den Vorfaktor

$$B = \frac{1}{\sqrt{\pi a^3}}.$$

Die Wellenfunktion für den Grundzustand des Wasserstoffatoms (1s-Zustand) hat damit die Gestalt

$$\psi_1(r) = \frac{1}{\sqrt{\pi a^3}} e^{-r/a}.$$

Die ausführliche quantenmechanische Rechnung liefert für die Wellenfunktion des ersten angeregten kugelsymmetrischen Zustands (2s-Zustand)

$$\psi_2(r) = \frac{1}{\sqrt{8\pi a^3}} \left(1 - \frac{r}{2a}\right) e^{-r/2a}$$

und für die Wellenfunktion des zweiten angeregten kugelsymmetrischen Zustands (3s-Zustand)

$$\psi_3(r) = \frac{1}{\sqrt{27\pi a^3}} \left( 1 - \frac{2r}{3a} + \frac{2r^2}{27a^2} \right) e^{-r/3a}.$$

Einsetzen in die Schrödingergleichung liefert analog wie oben (nur mit ein wenig mehr Rechenaufwand<sup>81</sup>) für die Energie des ersten und zweiten angeregten Zustandes

$$E_2 = -\frac{m_e e^4}{128\pi^2 \epsilon_0^2 \hbar^2} = -\frac{1}{4} \frac{m_e e^4}{32\pi^2 \epsilon_0^2 \hbar^2}$$

$$E_3 = -\frac{m_e e^4}{288\pi^2 \epsilon_0^2 \hbar^2} = -\frac{1}{9} \frac{m_e e^4}{32\pi^2 \epsilon_0^2 \hbar^2}$$

Hier ist eine Regelmäßigkeit bereits klar erkennbar: Allgemein gilt für die Energieniveaus im Wasserstoffatom

$$E_n = -\frac{m_e e^4}{32\pi^2 \epsilon_0^2 \hbar^2 n^2} = -\frac{13,6 \text{ eV}}{n^2}, \quad n \in \mathbb{N} \setminus \{0\}.$$

$n$  heißt Hauptquantenzahl. Diese Relation ist die Energieeigenwert-Gleichung des Wasserstoffatoms. Sie heißt auch *Rydberg-Formel*.

Die Situation ist analog zu derjenigen des linearen Potentialtopfs, mit gewissen Unterschieden im Detail. Genauer gilt folgendes:

- Die Energieniveaus sind quantisiert, das heißt, es sind wieder nur spezielle, diskrete Werte der Gesamtenergie möglich.
- Der Abstand der Energieniveaus wird mit steigender Hauptquantenzahl immer kleiner, das heißt die Energieniveaus werden für  $n \rightarrow \infty$  näherungsweise kontinuierlich („Grenzkontinuum“).
- Für  $n \rightarrow \infty$  geht  $E_n$  gegen  $E = 0$ .

Das Wasserstoffatom verfügt über definierte Werte der Gesamtenergie. Es handelt sich folglich um Energieeigenzustände. Dem Elektron in der Atomhülle können daher die Eigenschaften, einen bestimmten Ort, einen bestimmten Impuls sowie potentielle und kinetische Energie zu besitzen, *nicht zugeschrieben werden*.

Zwei weitere Aspekte können im Unterricht zwar nicht näher begründet werden, man sollte sie jedoch dennoch erwähnen:

- Die vollständige Lösung der Schrödingergleichung für das Wasserstoffatom einschließlich der winkelabhängigen Anteile zeigt unter anderem, daß die Energie  $E_n$  von den Winkeln  $\theta$  und  $\phi$  unabhängig ist. Unterschiedliche Wasserstoff-Wellenfunktionen mit derselben Quantenzahl  $n$  gehören somit zur selben Energie  $E_n$ . Man spricht hier auch von *Energie-Entartung*.

<sup>81</sup>Siehe Aufgabe 4.

- Genauere Rechnungen zeigen, daß diese Energie-Entartung nur näherungsweise gegeben ist, das heißt, Wasserstoff-Zustände mit gleichem  $n$ , aber unterschiedlicher Winkelabhängigkeit haben unterschiedliche Energien. Im Rahmen der relativistischen Quantenmechanik kann man diese sogenannte Feinstruktur der Wasserstoff-Energieniveaus exakt berechnen<sup>82</sup>.

Aufgrund des diskreten Energiespektrums sind auch in beim Wasserstoffatom *Energieübergänge* möglich: Bei geeigneter Energiezufuhr kann das Elektron des Wasserstoffatoms von einem tieferen in ein höheres Energieniveau übergehen. Beim Übergang von einem höheren in ein tieferes Energieniveau gibt es die entsprechende Energie wieder ab, meist unter Aussendung eines Photons. Für die Übergangsenergien gilt

$$\Delta E_{n_1, n_2} = E_{n_2} - E_{n_1} = \frac{m_e e^4}{32\pi^2 \epsilon_0^2 \hbar^2} \left( \frac{1}{n_1^2} - \frac{1}{n_2^2} \right)$$

Elektronen, die vom Niveau  $n_2$  ins Niveau  $n_1$  herunterpurzeln, geben folglich elektromagnetische Strahlung mit charakteristischen, diskreten Frequenzen ab. Diese nennt man *Spektrallinien*. Für die Frequenzen gilt:

$$\nu_{n_1, n_2} = \frac{E_{n_2} - E_{n_1}}{h} = \frac{m_e e^4}{64\pi^3 \epsilon_0^2 \hbar^3} \left( \frac{1}{n_1^2} - \frac{1}{n_2^2} \right) = \nu_R \left( \frac{1}{n_1^2} - \frac{1}{n_2^2} \right).$$

$$\nu_R = \frac{m_e e^4}{64\pi^3 \epsilon_0^2 \hbar^3} = 3,29 \cdot 10^{15} \text{ Hz heißt Rydbergfrequenz.}$$

Dieses Resultat kann auch mit dem Bohrschen Atommodell erzielt werden. Während letzteres jedoch mit willkürlichen Annahmen und der völlig unphysikalischen Vorstellung von Elektronenbahnen arbeitet, ist obiges Modell physikalisch hieb- und stichfest.

Für festes  $n_1$  bekommen die jeweiligen Übergänge eigene Namen (siehe auch Abbildung 4.20):

1.  $n_1 = 1, n_2 = 2, 3, 4, \dots$ : Lymanserie (ultraviolett)
2.  $n_1 = 2, n_2 = 3, 4, 5, \dots$ : Balmerserie (teilweise sichtbar)
3.  $n_1 = 3, n_2 = 4, 5, 6, \dots$ : Paschenserie (infrarot)
4.  $n_1 = 4, n_2 = 5, 6, 7, \dots$ : Bracketserie
5.  $n_1 = 5, n_2 = 6, 7, 8, \dots$ : Pfundserie
6. ...

Die einzelnen Linien werden mit griechischen Buchstaben durchnummeriert:

Bei der Lyman-Serie  $L_\alpha, L_\beta, L_\gamma, \dots$  usw.

Bei der Balmer-Serie  $H_\alpha(656,3 \text{ nm}), H_\beta(486,1 \text{ nm}), H_\gamma(434,0 \text{ nm}), \dots$  usw.

<sup>82</sup>Hier kommt anstelle der Schrödingergleichung deren relativistische Version zum Einsatz, die *Dirac-Gleichung*.

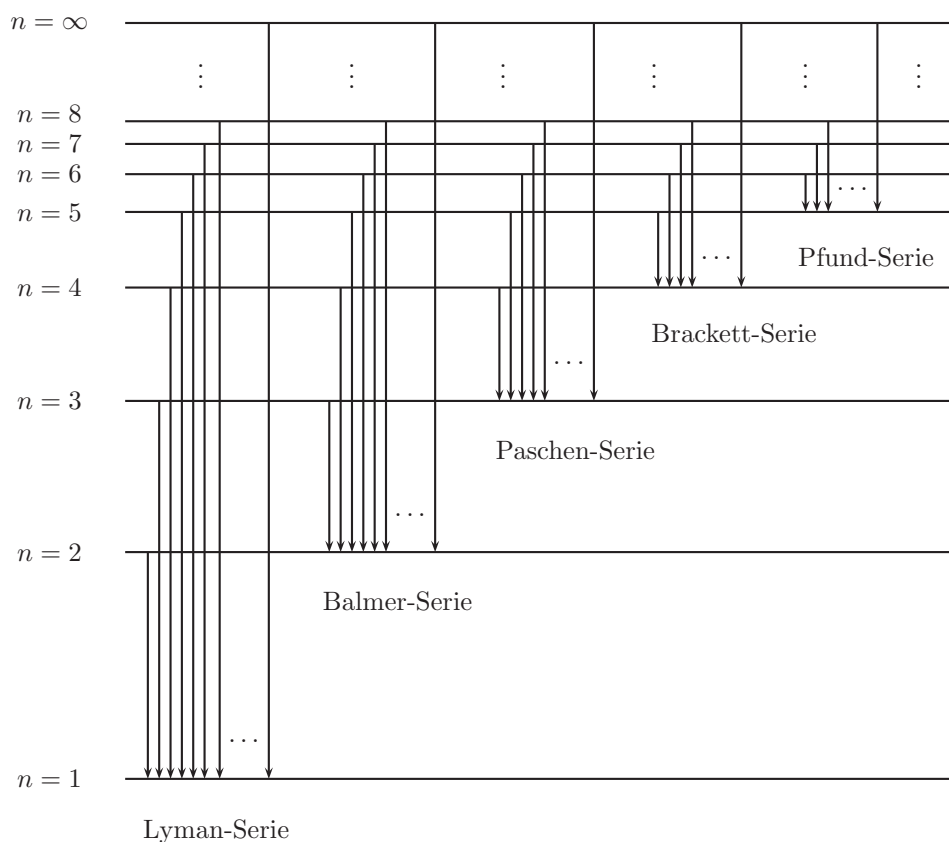


Abbildung 4.20: Graphische Darstellung der Übergänge beim Wasserstoffatom

Wird dem Elektron in der Hülle des Wasserstoffatoms mehr Energie als  $E_1 = -13,6$  eV zugeführt, so verläßt es die Atomhülle und wird zu einem freien Elektron. Man sagt, das Wasserstoffatom wird ionisiert. Die Lösung der Schrödingergleichung für freie Teilchen zeigt, daß deren Energie *nicht quantisiert ist*, sondern beliebig dicht liegende Werte hat; es gilt  $E > 0$ . Übergänge aus diesem Kontinuum in definierte Energieniveaus der Atomhülle mit  $E < 0$  erzeugen ein *kontinuierliches Emissionsspektrum*.

Die Wasserstoffwellenfunktionen ermöglichen die Berechnung der Aufenthaltswahrscheinlichkeitsdichten des Elektrons in den unterschiedlichen Zuständen. Da wir nur kugelsymmetrische Zustände betrachten und folglich stillschweigend über den Azimut- und den Polarwinkel integrieren, liefert

$$w_n(r) = 4\pi r^2 |\psi_n(r)|^2$$

die Aufenthaltswahrscheinlichkeitsdichte für das Elektron in Abhängigkeit vom Abstand  $r$  vom Kernmittelpunkt (siehe Abbildung 4.21). Die Wahrscheinlichkeit, das Elektron bei einer Ortsmessung in einem Abstand  $r$  mit  $r_1 \leq r \leq r_2$  vom Kern anzutreffen, beträgt damit

$$\mathfrak{P}_n[r_1, r_2] = 4\pi \int_{r_1}^{r_2} |\psi_n(r)|^2 r^2 dr.$$

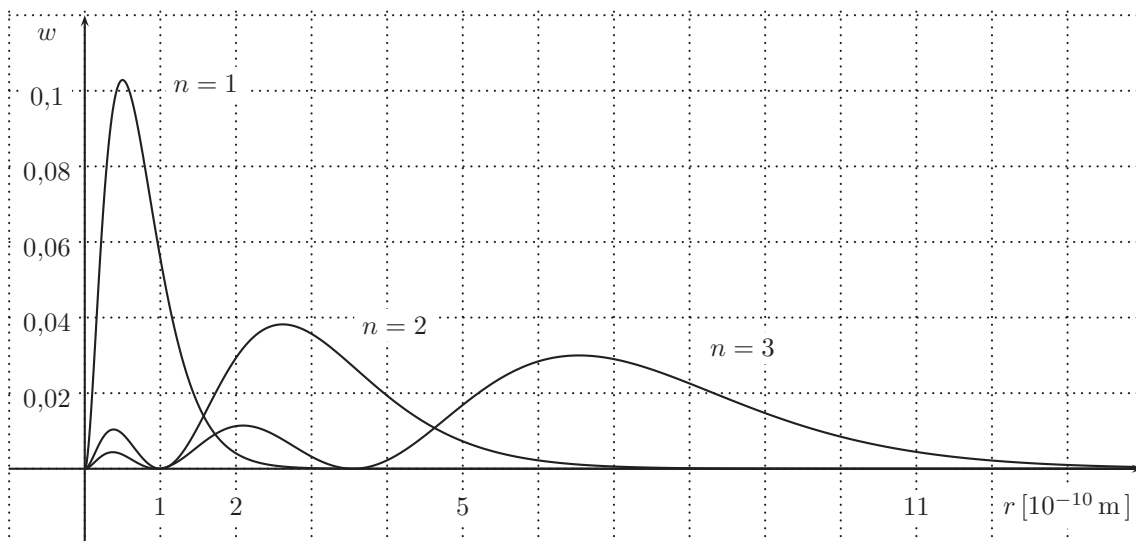


Abbildung 4.21: Radiale Aufenthaltswahrscheinlichkeitsdichten beim Wasserstoffatom für die ersten drei  $s$ -Zustände.

Für  $n > 1$  besitzen die Funktionen  $w_n(r)$  Nullstellen, folglich gibt es für  $n > 1$  Radien, an denen die Wahrscheinlichkeit, das Elektron anzutreffen, Null ist. Diese Wahrscheinlichkeit ist überall sonst ungleich Null, was die Tatsache illustriert, daß die Vorstellung von Elektronenbahnen in der Atomhülle abwegig ist.

Die  $w_n(r)$  veranschaulichen für  $n = 1, 2, 3, \dots$  die kugelsymmetrischen  $s$ -Orbitale. Berücksichtigt man bei der Lösung der Schrödingergleichung auch winkelabhängige Anteile (rechnerisch sehr aufwendig), so erhält man Wellenfunktionen  $\psi_{nlm}(r, \theta, \phi)$ , die von den drei räumlichen Koordinaten  $r$  (Abstand vom Kernmittelpunkt),  $\theta$  (Azimutwinkel) und  $\phi$  (Polarwinkel) sowie von mehreren Quantenzahlen abhängen.  $l$  heißt Drehimpulsquantenzahl,  $m$  heißt magnetische Quantenzahl.

$$w_{nlm}(r, \theta, \phi) = |\psi_{nlm}(r, \theta, \phi)|^2$$

ist dann die (im allgemeinen nicht mehr kugelsymmetrische) Aufenthaltswahrscheinlichkeitsdichte in Abhängigkeit vom Ort  $(r, \theta, \phi)$ . Die  $w_{nlm}$  veranschaulichen je nach den Werten der Quantenzahlen  $n, l$ , und  $m$  die kugelsymmetrischen  $s$ -Orbitale und die nicht kugelsymmetrischen  $p$ -,  $d$ - und  $f$ -Orbitale.  $l$  beschreibt die geometrische Form der Orbitale. Folgende Bezeichnungen sind in der Spektroskopie üblich (die Namen kommen vom Erscheinungsbild der zugehörigen Spektrallinien):

- $l = 0$  :  $s$ -Orbitale („sharp“)
- $l = 1$  :  $p$ -Orbitale („principal“)
- $l = 2$  :  $d$ -Orbitale („diffuse“)
- $l = 3$  :  $f$ -Orbitale („fundamental“)
- $l = 4$  :  $g$ -Orbitale
- $l = 5$  :  $h$ -Orbitale
- $\vdots$  :  $\vdots$

$m$  beschreibt im wesentlichen die räumliche Orientierung der Orbitale. In der Natur kommen nur Atome bis  $l = 3$  vor.

In Kugelkoordinaten  $r, \theta, \phi$  gilt für das Volumenelement

$$dV = dx dy dz = r^2 \sin \theta dr d\theta d\phi.$$

Folglich erhält man für die Wahrscheinlichkeit, bei einer Ortsmessung ein Elektron im Zustand  $(nlm)$  in dem durch  $r_1 \leq r \leq r_2$ ,  $\theta_1 \leq \theta \leq \theta_2$ ,  $\phi_1 \leq \phi \leq \phi_2$  definierten räumlichen Gebiet  $G$  anzutreffen, den Ausdruck

$$\mathfrak{P}_{nlm}(G) = \int_G |\psi_{nlm}|^2 dV = \int_{\phi_1}^{\phi_2} \int_{\theta_1}^{\theta_2} \int_{r_1}^{r_2} |\psi_{nlm}(r, \theta, \phi)|^2 r^2 \sin \theta dr d\theta d\phi.$$

Genaugenommen kommt bei Elektronen in der Atomhülle noch eine vierte Quantenzahl hinzu. Elektronen besitzen wie viele andere Elementarteilchen auch einen weiteren inneren Freiheitsgrad, den sogenannten Spin. Dieser hat die Dimension eines Drehimpulses und wird behelfsmäßig mit der Vorstellung einer Eigenrotation des Elektrons im („Spin down“) oder gegen den Uhrzeigersinn („Spin up“) veranschaulicht, was jedoch nicht wörtlich genommen werden darf. Der Spin ist eine rein quantenmechanische Größe und nicht klassisch beschreibbar. In der Atomhülle wird das durch die Spin-Quantenzahl  $s$  beschrieben:  $s = 1/2$  steht für Spin aufwärts,  $s = -1/2$  für Spin abwärts. Der Spin ist ein Effekt der relativistischen Quantenmechanik und zusammen mit dem Bahndrehimpuls des Elektrons für die erwähnte Feinstruktur der Energieniveaus beim Wasserstoffatom verantwortlich<sup>83</sup>.

#### 4.5.6.4 Quantenmechanik schwererer Atome

Das oben vereinfacht beschriebene Verfahren funktioniert auch für sogenannte wasserstoffähnliche Atome, also solche, die bis auf ein einziges Elektron ionisiert sind. Man hat dabei nur die Kernladungszahl  $Z$  des betrachteten Atoms zu ergänzen. Beispielsweise gilt dann für die Energieniveaus

$$E_n = -\frac{m_e Z^2 e^4}{32\pi^2 \epsilon_0^2 \hbar^2 n^2}$$

und für die Übergangsfrequenzen entsprechend

$$\nu = \frac{m_e Z^2 e^4}{64\pi^3 \epsilon_0^2 \hbar^3} \left( \frac{1}{n_1^2} - \frac{1}{n_2^2} \right).$$

Ist das betrachtete Atom nur soweit ionisiert, daß die äußere Schale noch genau ein Elektron enthält, kann man die Rydbergformel weiterhin näherungsweise verwenden, wenn man die Kernladungszahl  $Z$  durch eine effektive Kernladungszahl  $Z_{\text{eff}} < Z$  ersetzt, die entweder experimentell oder durch (nichttriviale) Näherungsberechnungen bestimmt werden muß. Sie berücksichtigt, daß die positive Ladung des Kerns durch die negativen

<sup>83</sup>Bei der in Anmerkung 82 erwähnten Dirac-Gleichung ist der Spin von vorn herein mit enthalten.



Ladungen der inneren Elektronen teilweise abgeschirmt wird. Sobald es auch in der äußersten Schale mehr Elektronen sind, müssen zur Berechnung der Wellenfunktionen und Energieniveaus aufwendigere Näherungsverfahren herangezogen werden.

Zwei weitere Aspekte seien noch gesondert erwähnt:

- Schnelle Elektronen können in den Atomhüllen von schweren Atomen (Ordnungszahl  $Z \gg 1$ ) ein Elektron mit  $n = 1$  herausschlagen. Rückt aus einem höheren Energieniveau ein Elektron in die Lücke nach, so gibt es ein Photon, dessen Frequenz näherungsweise berechnet werden kann, indem man wieder die Abschirmung der Kernladung betrachtet. Für diese ist hier nur das innerste Elektron verantwortlich, daher ersetzt man in der Rydbergformel die Kernladungszahl  $Z$  durch  $Z - 1$ , erhält so für die Frequenz des emittierten Photons

$$\nu = (Z - 1)^2 \nu_R \left( \frac{1}{1^2} - \frac{1}{n^2} \right),$$

und damit

$$\nu = (Z - 1)^2 \nu_R \left( 1 - \frac{1}{n^2} \right).$$

Diese Relation nennt man auch das *Moseleysche Gesetz* [374] – [375].

- Diskrete Energiespektren führen nicht nur zur Emission von Licht mit charakteristischen Frequenzen, sondern auch zum umgekehrten Effekt. Das sieht man beispielsweise an hand von folgendem Versuch: Weißes Licht wird durch ein Geradsichtprisma in ein kontinuierliches Spektrum aufgefächert. In den Lichtkegel wird Natriumdampf gebracht. Im gelben Teil des Spektrums erscheint eine dunkle Linie. Natriumdampf absorbiert Photonen, aber nur diejenigen, die genau die richtige Energie haben, um Elektronen in den Atomhüllen der Natriumatome von einem bestimmten Energieniveau auf ein bestimmtes höheres anzuregen. Man nennt diesen Effekt die *Umkehrung der Natrium-Linie*. Er tritt bei jedem System mit diskreten Energieniveaus auf; man spricht dann allgemein von der *Umkehrung von Spektrallinien*. Spektren mit solchen dunklen Absorptionslinien nennt man *Absorptionsspektren*. Eine wichtige Anwendung hiervon ist das Verfahren der *Spektralanalyse*: Jeder Stoff ist gekennzeichnet durch die ganz speziellen Energieniveaus seiner Atome oder Moleküle. Durch Untersuchung der Linien in Absorptionsspektren kann man analysieren, welche Stoffe diese Linien verursacht haben.

## Aufgaben

- Bestimme die Wellenfunktionen und Energieeigenwerte des „angehobenen“ Potentialtopfs mit unendlich hohen Wänden mit dem Potential

$$V(x) = \begin{cases} V_0 & \text{für } 0 \leq x \leq a, \\ \infty & \text{für } x < 0 \text{ oder } x > a. \end{cases}$$

- $\beta$ -Carotin ist eine organische Verbindung, die aus langgestreckten Kettenmolekülen besteht (siehe Abbildung 4.22), in denen sich jeweils 22 Elektronen praktisch frei bewegen, das Molekül dabei jedoch nicht verlassen können. Solche Moleküle können näherungsweise mit dem Modell eines linearen Potentialtopfs mit unendlich hohen Wänden beschrieben werden, wobei man annimmt, daß sich maximal je zwei Elektronen im selben Energiezustand aufhalten können. Im Absorptionsspektrum von  $\beta$ -Carotin findet man eine Linie bei 451 nm. Berechne die Länge eines solchen Moleküls.

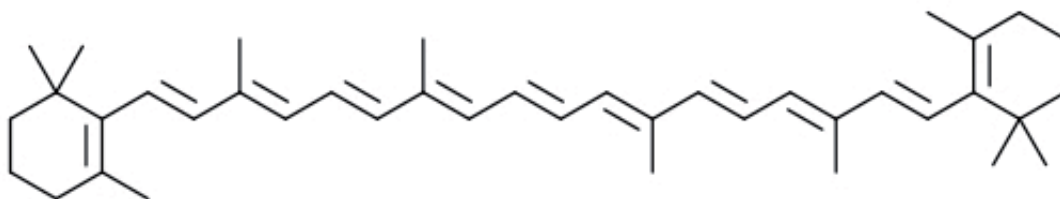


Abbildung 4.22: Strukturformel von  $\beta$ -Carotin

- Ein eindimensionaler quantenmechanischer harmonischer Oszillator mit der Masse  $m$  wird durch das Potential

$$V = \frac{1}{2} k x^2$$

beschrieben; dabei ist  $x$  der Abstand des Systems von der Gleichgewichtslage, und für die Kreisfrequenz gilt  $\omega = \sqrt{k/m}$ .

- Begründe die Form dieses Potentials.
- Bestimme die Wellenfunktion des Grundzustandes mit Hilfe des Ansatzes

$$\psi_0(x) = B e^{-a x^2}$$

und der Formel

$$\int_{-\infty}^{\infty} e^{-2 a x^2} dx = \sqrt{\frac{\pi}{2a}}.$$

- Bestimme die Grundzustandsenergie des harmonischen Oszillators.

- d) Berechne die Energie des ersten angeregten Zustandes. Verwende dazu den Ansatz

$$\psi_1(x) = B x e^{-ax^2}.$$

- e) Wie lautet der Ausdruck für die Energie des zweiten angeregten Zustandes? Ansatz:

$$\psi_2(x) = B (2ax^2 - 1) e^{-ax^2/2}.$$

- f) Welche allgemeine Formel für die Energieeigenwerte des harmonischen Oszillators zu beliebigem  $n \in \mathbb{N}$  kann man daraus ablesen?
- h) Welche Aussagen lassen sich über die potentielle und die kinetische Energie eines solchen Systems machen? Begründe Deine Antwort.
4. Bestimme das zweite und das dritte Energieniveau des Elektrons in der Hülle des Wasserstoffatoms. Verwende dazu die Wellenfunktionen des ersten und zweiten angeregten kugelsymmetrischen Zustandes.
5. Rydberg-Atome sind Atome, die sich in sehr hoch angeregten Zuständen befinden.
- a) Berechne die Wellenlänge der elektromagnetischen Strahlung, die emittiert wird, wenn ein in einem Rydberg-Zustand vorliegendes Wasserstoffatom vom Zustand mit  $n = 200$  in das benachbarte Niveau übergeht.
- b) Wie groß ist die Ionisierungsenergie eines Wasserstoffatoms im Rydberg-Zustand mit  $n = 400$ ?
6. Myonen sind negativ geladene Teilchen, welche die 207-fache Masse eines Elektrons und ansonsten dieselben physikalischen Eigenschaften haben. Atomkerne können durch Einfangen solcher Myonen sogenannte myonische Atome bilden. Berechne die minimale Wellenlänge der Photonen, die abgegeben werden, wenn aus Myonen mit vernachlässigbarer kinetischer Energie und vollständig ionisierten Aluminiumatomen myonische Aluminiumionen entstehen.
7. Betrachte ein Teilchen, das sich aus nicht zu großer Höhe über der Erdoberfläche im freien Fall befindet.
- a) Skizziere den Verlauf der potentiellen Energie des Teilchens in Abhängigkeit von der Höhe aus klassischer Sicht.
- b) Wie lautet die Schrödingergleichung für dieses System?

### 4.5.7 Superpositionsprinzip und verschränkte Zustände

Das abschließende Kapitel des vorliegenden Unterrichtsganges tanzt wie bereits gesagt ein wenig aus der Reihe, denn die hier diskutierten Themen können im Gymnasium aufgrund des mathematischen Aufwands einer exakten Beschreibung nur oberflächlich angerissen werden. Andererseits ist der Begriff der verschränkten Zustände<sup>84</sup> neben dem quantenmechanischen Superpositionsprinzip, mit dem er, wie wir sehen werden, auf das engste verwandt ist, der wichtigste Unterschied zwischen Quantenmechanik und klassischer Physik. Ob man den Schülerinnen und Schülern die Mühe, sich im Unterricht damit zu beschäftigen, zumuten will, muß wohl jede und jeder selbst entscheiden. Für das vorliegende Kapitel ist die Kenntnis der in den bisherigen Kapiteln entwickelten elementaren Sachverhalte in dem Sinn ausreichend, als keine weiteren Vorkenntnisse vorausgesetzt werden; ob auf der Ebene der gymnasialen Oberstufe tatsächlich ein erstes Verständnis dazu erzielt werden kann, ist jedoch eine ganz andere Frage.

Die Themen dieses Kapitels werden oft auch unter dem Begriff „Nichtlokalität der Quantenmechanik“ zusammengefaßt. Das ist problematisch, da dieser Begriff in der Physik sehr unterschiedliche Verwendung findet und bei falschem Gebrauch völliger Unsinn herauskommen kann (und in der Literatur leider auch verbreitet herauskommt); es herrscht hier eine ähnliche Begriffs- und Inhaltsverwirrung wie bei der Komplementarität.

#### 4.5.7.1 Lokalitätseigenschaften der Physik

Die klassische Physik (in dem Sinn aufgefaßt, wie sie in Abschnitt 4.5.1 charakterisiert wurde) weist verschiedene starke *Lokalitätseigenschaften* auf:

- L1 Die Lichtgeschwindigkeit ist für alle Signalübertragungen die höchste mögliche Geschwindigkeit.
- L2 Wenn sich räumlich getrennte Objekte vermöge irgendwelcher physikalischer Wechselwirkungen gegenseitig beeinflussen, kann sich diese Beeinflussung höchstens mit Lichtgeschwindigkeit ausbreiten.
- L3 Räumlich getrennte Objekte besitzen alle ihre physikalischen Eigenschaften unabhängig voneinander.

Die beiden Aussagen L1 und L2 werden zusammengenommen auch als Prinzip der *Mikrokausalität* bezeichnet. Entgegen verbreitet anzutreffenden Aussagen selbst in der Fachliteratur gilt nun folgendes *Lokalitätsprinzip*:

*In der Quantenmechanik ist das Prinzip der Mikrokausalität uneingeschränkt gültig.*

Wir werden sehen, daß die Aussage L3 in der Quantenmechanik maximal verletzt ist. In diesem (und nur in diesem!) Sinn ist die Quantenmechanik nichtlokal. Zur Vermeidung der erwähnten Verwirrungen verwendet die Fachsprache im hier vorliegenden Zusammenhang häufig anstelle von „Nichtlokalität“ den Begriff „Nichtseparierbarkeit“.

<sup>84</sup>Dieser Begriff wurde von Erwin Schrödinger 1935 in die physikalische Welt eingeführt [460], [461], [462]. Losgetreten wurde die Diskussion darüber durch den berühmten Artikel von Albert Einstein, Boris Podolsky und Nathan Rosen [139] sowie der Entgegnung durch Nils Bohr [69], wobei in beiden Aufsätzen der Ausdruck „verschränkter Zustand“ nicht ein einziges Mal auftaucht.

### 4.5.7.2 Das Superpositionsprinzip der Quantenmechanik

Das Superpositionsprinzip kennen wir aus der Wellenlehre. Dort besagt es, daß sich Wellen ungestört überlagern, was sich darin äußert, daß sich ihre Feldstärkevektoren bei Überlagerung einfach addieren. Genaugenommen gilt das allerdings nur für Wellen, die sich durch lineare Differentialgleichungen beschreiben lassen<sup>85</sup>. In der Quantenmechanik taucht dieses Superpositionsprinzip auf spektakuläre Weise und mit erheblich weiterreichenden Folgen wieder auf.

Ausgangspunkt ist wie üblich die Schrödingergleichung. Hierfür notieren wir (unter anderem) drei Feststellungen:

- (i) Wellenfunktionen sind Lösungen der Schrödingergleichung des betrachteten Systems.
- (ii) Die Schrödingergleichung ist eine lineare Differentialgleichung.
- (ii) Für lineare Differentialgleichungen gilt: Linearkombinationen von Lösungen sind wieder Lösungen

Das hieraus folgende Prinzip ist von fundamentaler Bedeutung und bekannt unter dem Namen

*Quantenmechanisches Superpositionsprinzip:*

*Sind  $\psi_1$  und  $\psi_2$  Lösungen der Schrödingergleichung eines physikalischen Systems, dann ist  $\psi = a\psi_1 + b\psi_2$  für beliebige  $a, b \in \mathbb{C}$  auch eine Lösung dieser Schrödingergleichung.*

Hier genügt die Beschränkung auf Koeffizienten  $a, b \in \mathbb{R}$ <sup>86</sup>. – Solche Überlagerungen von Wellenfunktionen nennt man *Superpositionen*. Entsprechend spricht man bei den zugehörigen Zuständen der dadurch beschriebenen Systeme von *Superpositionszuständen*. Es gibt auch Superpositionen aus mehr als zwei und sogar aus unendlich vielen Zuständen. Der Beweis des Superpositionsprinzips erfolgt durch einfaches Einsetzen in die Schrödingergleichung und anschließendes Ausrechnen.

Wir betrachten einige Beispiele:

a) Elektronen am Doppelspalt

Die Wellenfunktion für ein Elektron am Doppelspalt ist eine Superposition aus der Wellenfunktion für ein Elektron, das durch den linken Spalt geht, und der Wellenfunktion für ein Elektron, daß durch den rechten Spalt geht.

b) Photonen im Mach-Zehnder-Interferometer

Die Wellenfunktion eines Photons im Mach-Zehnder-Interferometer ist eine Superposition der Wellenfunktionen des Photons im einen bzw. im anderen Arm des Interferometers.

<sup>85</sup>Beispiele für lineare Differentialgleichungen in der Physik sind die Maxwell-Gleichungen in der Elektrodynamik oder allgemein die Wellengleichung. Die Feldgleichungen der Gravitation in der allgemeinen Relativitätstheorie dagegen sind nichtlinear, sodaß für Gravitationsfelder kein Superpositionsprinzip gilt.

<sup>86</sup>Siehe Anmerkung 66.

## c) Stereoisomerie

Gibt es bei einem Molekül verschiedene Versionen, bei denen zwar alle Atome jeweils die gleichen Bindungspartner haben, sich aber in ihrer räumlichen Anordnungen unterscheiden, spricht man von Stereoisomerie. Die Wellenfunktion solcher Moleküle kann eine Superposition aus den Wellenfunktionen der einzelnen Versionen des Moleküls sein.

## d) Elektronen in der Atomhülle

Die Wellenfunktion eines Elektrons in der Atomhülle eines Atoms kann als überabzählbar-unendliche Superposition aller möglichen Ortseigenzustände des Elektrons aufgefaßt werden

Die Beispiele b) und d) sind mathematisch nichttrivial. Genauer gilt folgendes:

- Es ist nicht unmittelbar klar, was Wellenfunktionen von Photonen sein sollen, da letztere relativistische Objekte sind und die Schrödingersche Quantenmechanik eine rein nichtrelativistische Theorie ist. Es gibt in der relativistischen Quantenmechanik jedoch Ersatzobjekte, die sich praktisch genauso wie Wellenfunktionen von Photonen verhalten<sup>87</sup>.
- Wellenfunktionen für Ortseigenzustände machen eine aufwendige Erweiterung des mathematischen Formalismus erforderlich, was jedoch streng durchführbar ist. Darin macht sich die Tatsache bemerkbar, daß exakt lokalisierte Teilchen natürlich Idealisierungen darstellen.

Für schulische Zwecke sind diese Sachverhalte ohne Belang.

Nun zur Interpretation des Superpositionsprinzips: Befindet sich ein System in einem Zustand, der durch eine Superposition beschrieben wird, dann hat das System sämtliche durch die einzelnen Wellenfunktionen repräsentierten Eigenschaften *gleichzeitig*. Diese sind im allgemeinen physikalisch unterscheidbar, sie können sich aus klassischer Sicht widersprechen oder sogar *einander ausschließen*. Das quantenmechanische Superpositionsprinzip geht damit in seiner Tragweite weit über dasjenige für klassische Wellen hinaus: Während letzteres nur von einer Überlagerung von *Feldstärken* spricht, handelt ersteres von einer Überlagerung von *physikalisch unterschiedlichen Zuständen* des betrachteten Systems, die sich dabei gegenseitig ausschließen können: Das System besitzt verschiedene, sich eigentlich widersprechende Eigenschaften gleichzeitig. Das bedeutet beliebig nichtklassisches Verhalten.

- Beispiel Interferometer: Hier haben wir Superpositionszustände, bei denen sich Quantenobjekte an unterschiedlichen Orten gleichzeitig befinden.
- Beispiel Stereoisomerie: Einzelne Atome von Molekülen mit Stereoisomerie befinden sich in unterschiedlichen Anordnungen zueinander gleichzeitig.

Die wichtigste Folgerung aus dem Superpositionsprinzip ist diese: Bildet man das Betragsquadrat einer quantenmechanischen Superposition, so treten im allgemeinen gemischte Terme, also *Interferenzterme* auf: Quadrieren des Betrags eines Terms wie

$$\psi = a_1 \psi_1 + a_2 \psi_2$$

<sup>87</sup>Siehe Anmerkung 28.

führt auf

$$|\psi|^2 = |a_1 \psi_1 + a_2 \psi_2|^2 = |a_1|^2 |\psi_1|^2 + |a_2|^2 |\psi_2|^2 + 2 a_1 a_2 \psi_1 \psi_2,$$

oder allgemeiner: Bildet man das Betragsquadrat von

$$\psi = \sum_n a_n \psi_n,$$

so erhält man

$$|\psi|^2 = \left| \sum_n a_n \psi_n \right|^2 = \sum_n |a_n|^2 |\psi_n|^2 + 2 \sum_n \sum_m a_n a_m \psi_n \psi_m.$$

Quantenmechanische Superpositionen sind folglich *interferenzfähig*. Das ist die wesentliche Ursache für die bereits diskutierten und alle anderen quantenmechanischen Merkwürdigkeiten. Wir halten fest:

*Der wichtigste Unterscheid zwischen klassischer Physik und Quantenmechanik ist die Existenz interferenzfähiger Superpositionen von Zuständen in der Quantenmechanik.*

Als wichtige Folge des Superpositionsprinzips gilt damit: *Die Quantenmechanik ist nicht-lokal* im Sinne einer Verletzung des Lokalitätsprinzips L3. In der Quantenmechanik treten Fernwirkungen auf, das heißt, räumlich getrennte Objekte können sich unabhängig von ihrer gegenseitigen Entfernung instantan beeinflussen. Das hat jedoch *nichts* mit irgendwelchen physikalischen Wechselwirkungen über beliebig große Entfernungen in beliebig kurzen Zeiten zu tun, und schon gar nicht wird dadurch Signalübertragung mit Überlichtgeschwindigkeit ermöglicht.

Die Verletzung der Lokalität durch das Superpositionsprinzip in der Quantenmechanik tritt auf zwei unterschiedlichen Ebenen auf, weswegen wir die Diskussion dieses Sachverhalts in zwei Abschnitte aufteilen:

- Einfache Systeme
- Zusammengesetzte Systeme

Der erste Fall wurde bereits diskutiert und kommt in diesem Abschnitt nur als kurze Wiederholung zur Sprache, der zweite Fall ist Gegenstand des nächsten Abschnitts.

Wir betrachten ein Einteilchensystem, dessen Zustand eine Superposition aus *räumlich getrennten* Zuständen ist. Ein typisches Beispiel hierfür ist ein Zwei-Wege-Interferometer (Doppelspalt, Mach-Zehnder-Interferometer oder dergleichen), bei dem dafür gesorgt wird, daß sich immer nur ein Quantenobjekt gleichzeitig in der Apparatur aufhalten kann. Wie wir bereits ausführlich gesehen haben, erhält man bei sehr häufiger Wiederholung des Versuchs *nicht* das selbe Resultat wie bei Überlagerung der Resultate der beiden Einzelspalte. Wir folgern daraus:

- Das Einzelobjekt merkt beide Spalte, das heißt, es nimmt gleichzeitig räumlich getrennte Zustände ein, die sich gegenseitig beeinflussen (Es geht auf irgendeine Weise durch beide Spalte gleichzeitig hindurch und interferiert dabei mit sich selbst).



- Es liegt hier ein Beispiel für die Nichtlokalität der Quantenmechanik im Sinne von L3 vor.

Inzwischen ist es technisch möglich, auch andere solche Superpositionen aus klassisch sich ausschließenden Zuständen gezielt experimentell herzustellen, etwa ein Atom, das sich an zwei oder mehreren Orten gleichzeitig aufhält [370] oder Ionen, die sich in Schwingungszuständen befinden, welche sich klassisch widersprechen, beispielsweise gleich- und gegenphasig gleichzeitig oder ähnliches [288]<sup>88</sup>.

### 4.5.7.3 Verschränkte Zustände

In diesem Abschnitt haben wir es wie angekündigt mit Mehrkomponentensystemen, also zusammengesetzten Systemen, zu tun. Das sprengt eigentlich den Rahmen des gymnasialen Physikunterrichts sowohl aus mathematischer als auch aus physikalischer Hinsicht bei weitem. Daher müssen wir uns hier mit einer extrem elementarisierten Darstellung begnügen<sup>89</sup>.

Zunächst ist eine Vorüberlegung zur Beschreibung zusammengesetzter Systeme erforderlich. Zusammengesetzte Systeme sind Systeme, die aus mehreren Komponenten bestehen, wie beispielsweise

- Kaskadenphotonen (Photonenpaare, die bei speziellen Kernreaktionen ausgesandt werden),
- Atome mit mehreren Elektronen in der Atomhülle,
- Moleküle

Ganz allgemein sind alle Systeme, deren Komponenten beim „Zusammenbau“ miteinander in physikalische Wechselwirkung getreten sind, im quantenmechanischen Sinn zusammengesetzte Systeme.

Im Interesse einer übersichtlichen Darstellung verwenden wir für Wellenfunktion eine stark vereinfachte Form der sogenannten *Dirac-Schreibweise*: Wir betrachten ein Quantenobjekt  $a$  mit der Eigenschaft  $E$ . Für seine Wellenfunktion  $\psi_a$  schreiben wir kurz und präzise

$$\psi_{aE} = |aE\rangle.$$

Zwei Beispiele dazu:

- (i) Von zwei Photonen sei Photon 1 waagrecht, Photon 2 senkrecht polarisiert. Für ihre Wellenfunktionen schreiben wir in der Dirac-Schreibweise

$$\psi_1 = |1 \rightarrow\rangle, \quad \psi_2 = |2 \uparrow\rangle.$$

<sup>88</sup>Die Arbeitsgruppe von R. Schnabel arbeitet derzeit am Albert-Einstein-Institut in Hannover am Aufbau eines Experiments, bei dem zwei makroskopische Spiegel in einen Zustand gebracht werden sollen, in welchem sie ebenfalls gleichzeitig gleich- und gegenphasig schwingen [379], [453].

<sup>89</sup>Die Quantenmechanische Betrachtung von Einteilchen-Systemen (wie in den bisherigen Abschnitten durchweg geschehen) stellt *immer* eine Idealisierung dar, da alle physikalischen Systeme stets in Wechselwirkung mit ihrer Umgebung stehen. Diese Umgebung ist genaugenommen das gesamte Universum. Je nach System ist diese Umgebung ganz oder teilweise zu vernachlässigen, wenn man tatsächlich etwas ausrechnen können will.

- (ii) Elektronen in der Atomhülle werden durch die Hauptquantenzahl  $n$ , die Drehimpulsquantenzahl  $l$  und die magnetische Quantenzahl  $m$  beschrieben. Entsprechend schreiben wir für ihre Wellenfunktion in der Dirac-Schreibweise

$$\psi_{nlm} = |nlm\rangle,$$

wobei für  $n, l$  und  $m$  die jeweils zutreffenden Zahlen einzusetzen sind, zum Beispiel  $|100\rangle$  für ein Elektron im  $1s$ -Zustand oder  $|321\rangle$  für ein Elektron im  $3p_y$ -Zustand.

Die Dirac-Schreibweise erlaubt eine sehr einfache Formulierung von Skalarprodukten:

$$\int_{-\infty}^{\infty} \varphi^*(x) \psi(x) dx = \langle \varphi | \psi \rangle$$

(der Stern kann für schulische Zwecke getrost ignoriert werden), und Normierungen schreiben sich entsprechend  $\langle \psi | \psi \rangle = 1$ . Dabei heißt  $\langle \psi |$  *Bra-Vektor* oder kurz *Bra* und  $|\psi\rangle$  *Ket-Vektor* oder kurz *Ket*, weswegen man auch von der *Bra-Ket-Schreibweise* spricht<sup>90</sup>.

Nun kommen wir zur (vereinfachten) formalen Beschreibung von zusammengesetzten Systemen. Hierbei gilt:

Wellenfunktionen zusammengesetzter Systeme erhält man durch Zusammenkleben der Wellenfunktionen der das zusammengesetzte System bildenden Einzelsysteme.

Der Begriff „Zusammenkleben“ ist mathematisch streng definierbar<sup>91</sup>, was für die hier verfolgten Zwecke gleichermaßen fachlich viel zu schwierig wie unnötig ist. Zur Veranschaulichung betrachten wir folgende Beispiele:

- Beispiel 1 (allgemein): Ein zusammengesetztes System bestehe aus drei Komponenten 1, 2 und 3 mit den möglichen Wellenfunktionen  $|\psi\rangle, |\varphi\rangle$  und  $|\chi\rangle$ . Eine mögliche Wellenfunktion des zusammengesetzten Systems ist dann  $|\Psi\rangle = |\psi_1\rangle \otimes |\varphi_2\rangle \otimes |\chi_3\rangle$ , oder kürzer (und etwas unpräziser)  $|\Psi\rangle = |\psi_1\rangle |\varphi_2\rangle |\chi_3\rangle$ , oder noch kürzer (und noch unpräziser)  $|\Psi\rangle = |\psi_1 \varphi_2 \chi_3\rangle$ . In diesem zusammengesetzten Zustand befindet sich System 1 im Zustand  $\psi$ , System 2 im Zustand  $\varphi$  und System 3 im Zustand  $\chi$ .
- Beispiel 2 (konkret): Ein zusammengesetztes System bestehe aus zwei Photonen, die beide jeweils waagrecht oder senkrecht polarisiert sein können und dabei durch die Wellenfunktion  $|\rightarrow\rangle$  beziehungsweise  $|\uparrow\rangle$  beschrieben werden. Mögliche Wellenfunktionen des zusammengesetzten Systems sind dann  $|\Psi\rangle = |\uparrow\uparrow\rangle$  oder  $|\Psi\rangle = |\uparrow\rightarrow\rangle$  oder ähnliches.

Natürlich gilt auch hier das *Superpositionssprinzip*, das heißt, es gibt Überlagerungen aus unterschiedlichen zusammengesetzten Zuständen. Das hat höchst merkwürdige Konsequenzen und gibt Anlaß zu einigen Bemerkungen über *Produktzustände und verschränkte Zustände*.

<sup>90</sup>Diese Bezeichnung verweist auf das englische Wort „bracket“ für „Klammer“.

<sup>91</sup>Der Fachmann erkennt dahinter das Tensorprodukt.

Zusammengesetzte Zustände wie die oben beschriebenen können als so etwas wie Produkte aus den Zuständen der einzelnen Komponenten aufgefaßt werden und heißen daher *Produktzustände*. Aus dem Superpositionsprinzip folgt jedoch, daß es Zustände gibt, die sich nicht in dieser Form schreiben lassen. Beispiele hierfür sind

$$|\Psi\rangle = \frac{1}{\sqrt{3}} (|\psi_1 \varphi_2 \chi_3\rangle + |\psi_2 \varphi_3 \chi_1\rangle + |\psi_3 \varphi_1 \chi_2\rangle)$$

oder

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|\uparrow \rightarrow\rangle + |\rightarrow \uparrow\rangle).$$

Solche Zustände heißen *verschränkte Zustände*. Die Vorfaktoren sorgen hierbei dafür, daß die zusammengesetzten Zustände normiert sind. Dabei wird vorausgesetzt, daß die Einzelzustände ebenfalls normiert und außerdem zueinander orthogonal sind, das heißt, das Skalarprodukt zweier unterschiedlicher Zustände soll jeweils verschwinden.

Die Interpretation ist formal naheliegend, aber wieder typisch quantenmechanisch seltsam: Verschränkte Zustände sind Zustände, in denen sich das zusammengesetzte System in einer Überlagerung aus unterschiedlichen Kombinationen der einzelnen Komponenten befindet. Im zweiten Beispiel bedeutet dies: Ein Zwei-Photonen-System mit der Wellenfunktion

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|\uparrow \rightarrow\rangle + |\rightarrow \uparrow\rangle)$$

befindet sich in einer Überlagerung aus

„Photon 1 senkrecht polarisiert, Photon 2 waagrecht polarisiert“

und

„Photon 1 waagrecht polarisiert, Photon 2 senkrecht polarisiert“.

Das hat bedeutende Konsequenzen:

- In diesem Zustand besitzt keines der beiden Photonen die Eigenschaft, über eine definierte Polarisation zu verfügen. Sie haben gewissermaßen jeweils beide Polarisationen gleichzeitig.
- Die beiden Photonen können in diesem Zustand nicht als getrennte Einzelsysteme betrachtet werden. Sie bilden ein gemeinsames System, das nicht getrennt werden kann, ohne den Zustand zu zerstören.

Zustände mit Wellenfunktionen der Form

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|\psi_1 \varphi_2\rangle + |\psi_2 \varphi_1\rangle)$$

heißen *Bell-Zustände*. Sie stellen nur ein mögliches Beispiel für verschränkte Zustände dar. Ein weiteres Beispiel sind *GHZ-Zustände* (Greenberger-Horne-Zeilinger-Zustände) [198], [199], [200]. Dies sind Drei-Teilchen-Zustände mit Wellenfunktionen der Form

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|\psi_1 \psi_2 \psi_3\rangle + |\varphi_1 \varphi_2 \varphi_3\rangle);$$

oder Vier-Teilchen-Zustände

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|\psi_1 \psi_2 \psi_3 \psi_4\rangle + |\varphi_1 \varphi_2 \varphi_3 \varphi_4\rangle)$$

oder allgemein  $n$ -Teilchen-Zustände

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|\psi_1 \psi_2 \dots \psi_n\rangle + |\varphi_1 \varphi_2 \dots \varphi_n\rangle).$$

Die quantitative Untersuchung verschränkter Zustände aus experimenteller wie auch theoretischer Sicht, insbesondere in Hinblick auf das Maß ihrer Verschränkung im Vergleich zu Produktzuständen, hat sich inzwischen zu eigenständigen Teilgebieten der experimentellen, theoretischen und mathematischen Quantenmechanik entwickelt.

#### 4.5.7.4 Verborgene Variable und Bellsche Ungleichungen

Das Superpositionsprinzip wurde bereits im Zusammenhang mit Interferenzeffekten als wesentliche, den Unterschied zur klassischen Physik ausmachende Eigenschaft der Quantenmechanik erwähnt. Das bestätigt sich in drastischem Ausmaß bei verschränkten Zuständen: Wie sich zeigt, läßt sich hierüber beweisen, daß sich die Quantenmechanik prinzipiell durch keine noch so raffinierte Theorie klassischer Bauart ersetzen läßt. Wir erkennen das besonders deutlich, wenn wir Messungen an verschränkten Zuständen betrachten.

Selbst wenn die Komponenten eines verschränkten Systems *weit voneinander entfernt* sind, bleiben sie *korreliert*. Das bedeutet: Wird an einer der Komponenten des verschränkten Systems etwas gemessen, merken die anderen Komponenten dies sofort, egal wie weit sie entfernt sind (im Extremfall quer durch das Universum). Nach einer Messung ist die Verschränkung im allgemeinen aufgehoben.

Wir veranschaulichen das anhand von Bell-Zuständen. Dazu betrachten wir einen Versuch, bei dem Polarisationsmessungen an verschränkten Zuständen aus zwei Photonen vorgenommen werden, etwa mit der Wellenfunktion

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|\uparrow\uparrow\rangle + |\rightarrow\rightarrow\rangle);$$

das beschreibt einen Superpositionszustand aus „beide Photonen senkrecht polarisiert“ und „beide Photonen waagrecht polarisiert“. Mit schulischen Mitteln ist das natürlich nicht real durchführbar, sodaß es hier bei einem Gedankenexperiment bleiben muß, das aufgrund seiner einfachen Struktur jedoch gut nachvollziehbar ist. Dabei sind zwei Resultate wesentlich:

- a) Wird an Photon 1 senkrechte Polarisation gemessen, ist *ab sofort* die Wahrscheinlichkeit, an Photon 2 ebenfalls senkrechte Polarisation zu messen, 100 %. Umgekehrt gilt entsprechendes. Wichtig: Vor der Messung haben *beide* Photonen *keine* definierte Polarisation.
- b) Wird Photon 1 durch einen Polarisationsfilter mit *beliebiger* Polarisation geschickt, besteht eine gewisse, von der Orientierung dieses Polarisationsfilters abhängige Wahrscheinlichkeit, das es durchgeht. Geht es durch, ist *ab sofort* die Wahrscheinlichkeit, an Photon 2 genau die gleiche Polarisation zu messen, 100 %. Wieder gilt umgekehrt entsprechendes. Dabei können die beiden Photonen räumlich und die beiden Messungen zeitlich beliebig weit voneinander getrennt sein.

Folgendes ist hierzu anzumerken:

- Sachverhalt a) ist zunächst nichts besonderes und an jedem klassischen System, das in Teilsysteme zerlegt und getrennt wird, ebenfalls feststellbar (z.B. wenn man feststellt, daß man nur einen rechten Handschuh dabei hat und dann sofort weiß, daß zu Hause ein linker Handschuh sein muß).
- Sachverhalt b) ist klassisch nicht erklärbar. Man spricht hierbei häufig von *Fernkorrelationen* oder *Action at a Distance* in der Quantenmechanik.

Hier ist jedoch größte Vorsicht vor Fehlinterpretationen angebracht: Korrelationen aufgrund von Verschränkung haben *nichts* mit irgendwelchen physikalischen Wechselwirkungen über beliebige Distanzen in Null Komma Nix zu tun und können daher weder zu gezielter Beeinflussung weit entfernter Gegenstände noch zur Nachrichtenübermittlung mit Überlichtgeschwindigkeit („Bellsches Telephon“) oder ähnlichem verwendet werden<sup>92</sup>. Solche Fernkorrelationen sind nicht direkt wahrnehmbar, sondern müssen durch Meßreihen *statistisch* nachgewiesen werden.

Das macht die Sache jedoch nicht im geringsten weniger merkwürdig. Fernkorrelationen bedeuten instantane Beeinflussung eines physikalischen Objekts durch ein anderes über beliebige Entfernungen und widersprechen damit klassischen Vorstellungen in beliebigem Ausmaß. Wer nicht ohne weiteres bereit ist, so etwas zu akzeptieren, stellt sich nun folgende Frage: Könnte es sein, daß

- Die Photonen in Wirklichkeit gar nicht korreliert sind, sondern
- Photonen verborgene, das heißt nicht meßbare zusätzliche Eigenschaften (verborgene Variablen) besitzen und damit
- bei jedem vermeintlich verschränkten Photonenpaar individuell von vorneherein feststeht, was gemessen werden wird?

Solche Theorien nennt man lokale Theorien mit verborgenen Variablen. Sie versuchen, die quantenmechanischen Nichtlokalitäten zu vermeiden. Würden sie richtig sein, dann wäre die Natur nur scheinbar von quantenmechanisch-statistischem Charakter. Wir sehen uns offensichtlich mit zwei ganz grundsätzlich unterschiedlichen Vorstellungen der Natur konfrontiert:

- **Verborgene Variable:**

Meßwerte können nicht vorausgesagt werden und scheinen statistisch verteilt zu sein, weil sie von Größen abhängen, nämlich von verborgenen Variablen, die wir nicht oder noch nicht messen können. Das Prinzip der Kausalität bleibt global erhalten.

- **Quantenmechanik:**

Meßwerte können prinzipiell nicht vorausgesagt werden; es können nur statistische Aussagen über sie gemacht werden, weil das eine grundlegende Eigenschaft der Natur selbst ist. Die Natur verletzt in diesem Sinn das Prinzip der Kausalität, auch wenn die dynamischen Gesetze der Physik (z.B. die Schrödingergleichung) natürlich streng kausal sind.

---

<sup>92</sup>Siehe Aufgabe 6.

Besondere Bedeutung kommt den verschränkten Systemen in diesem Zusammenhang zu, weil sie die Durchführung von *Entscheidungsexperimenten* ermöglichen, für die sämtliche mögliche lokale Theorien mit verborgenen Variablen einerseits und die Quantenmechanik andererseits *unterschiedliche Ergebnisse voraussagen*. Im Zusammenhang mit derartigen Experimenten ist es möglich, sogenannte *Bellsche Ungleichungen* herzuleiten, die von lokalen Theorien mit verborgenen Variablen stets erfüllt, von der Quantenmechanik dagegen immer verletzt werden. Wir betrachten nun die Grundidee eines Beispiels für solche Experimente in sehr allgemeiner Form. Dazu brauchen wir verschränkte Zweiteilchen-Zustände der Form

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|\psi_1 \varphi_2\rangle + |\psi_2 \varphi_1\rangle).$$

Gemessen werden die Größen  $A$  und  $A'$  an Teilchen 1 sowie  $B$  und  $B'$  an Teilchen 2. Die zugehörigen Meßwerte seien  $a$  und  $a'$  beziehungsweise  $b$  und  $b'$ . Dabei wird vorausgesetzt, daß es sich jeweils um zweiwertige Größen handelt, das heißt, die Meßresultate können nur zwei Werte annehmen, zum Beispiel  $a, b, a', b' \in \{-1, +1\}$ . Nun werden folgende *Korrelationsexperimente* durchgeführt:

- Meßreihe I:
  - Beobachter 1 mißt an Teilchen 1 die Größe  $A$ ,
  - Beobachter 2 mißt an Teilchen 2 die Größe  $B$ .
- Meßreihe II:
  - Beobachter 1 mißt an Teilchen 1 die Größe  $A$ ,
  - Beobachter 2 mißt an Teilchen 2 die Größe  $B'$ .
- Meßreihe III:
  - Beobachter 1 mißt an Teilchen 1 die Größe  $A'$ ,
  - Beobachter 2 mißt an Teilchen 2 die Größe  $B$ .
- Meßreihe IV:
  - Beobachter 1 mißt an Teilchen 1 die Größe  $A'$ ,
  - Beobachter 2 mißt an Teilchen 2 die Größe  $B'$ .

Zunächst sei vorausgesetzt, daß das System durch lokale Theorie mit verborgenen Variablen beschrieben werden kann. Das bedeutet:

- Jedes Meßresultat hängt ausschließlich vom gerade ankommenden Teilchen ab und von nichts sonst, insbesondere nicht vom anderen Teilchen, das mit dem ersten verschränkt ist.
- Das Lokalitätsprinzip L3 aus Abschnitt 4.5.7.1 ist voll erfüllt.
- Es gibt eine wohldefinierte Wahrscheinlichkeitsverteilung für die möglichen Meßwerte bei Messung der Größen  $A, B, A', B'$ .

Nun betrachten wir die *bedingten Wahrscheinlichkeiten*

$$\begin{aligned} \mathfrak{P}_{AB}(a, b) &= \text{Wahrscheinlichkeit, bei Messung von } A \text{ an Teilchen 1 den Wert } a \\ &\quad \text{und Messung von } B \text{ an Teilchen 2 den Wert } b \text{ zu erhalten,} \\ \mathfrak{P}_{A'B}(a', b) &= \text{Wahrscheinlichkeit, bei Messung von } A' \text{ an Teilchen 1 den Wert } a' \\ &\quad \text{und Messung von } B \text{ an Teilchen 2 den Wert } b \text{ zu erhalten,} \\ \mathfrak{P}_{AB'}(a, b') &= \text{Wahrscheinlichkeit, bei Messung von } A \text{ an Teilchen 1 den Wert } a \\ &\quad \text{und Messung von } B' \text{ an Teilchen 2 den Wert } b' \text{ zu erhalten,} \\ \mathfrak{P}_{A'B'}(a', b') &= \text{Wahrscheinlichkeit, bei Messung von } A' \text{ an Teilchen 1 den Wert } a' \\ &\quad \text{und Messung von } B' \text{ an Teilchen 2 den Wert } b' \text{ zu erhalten,} \end{aligned}$$

sowie die *Korrelationsfunktionen*

$$\begin{aligned} \mathcal{C}(A, B) &= \sum_{a,b} ab \mathfrak{P}_{AB}(a, b) \\ &= \mathfrak{P}_{AB}(+1, +1) - \mathfrak{P}_{AB}(+1, -1) - \mathfrak{P}_{AB}(-1, +1) + \mathfrak{P}_{AB}(-1, -1), \\ \mathcal{C}(A, B') &= \sum_{a,b'} ab' \mathfrak{P}_{AB'}(a, b') \\ &= \mathfrak{P}_{AB'}(+1, +1) - \mathfrak{P}_{AB'}(+1, -1) - \mathfrak{P}_{AB'}(-1, +1) + \mathfrak{P}_{AB'}(-1, -1), \\ \mathcal{C}(A', B) &= \sum_{a',b} a'b \mathfrak{P}_{A'B}(a', b) \\ &= \mathfrak{P}_{A'B}(+1, +1) - \mathfrak{P}_{A'B}(+1, -1) - \mathfrak{P}_{A'B}(-1, +1) + \mathfrak{P}_{A'B}(-1, -1), \\ \mathcal{C}(A', B') &= \sum_{a',b'} a'b' \mathfrak{P}_{A'B'}(a', b') \\ &= \mathfrak{P}_{A'B'}(+1, +1) - \mathfrak{P}_{A'B'}(+1, -1) - \mathfrak{P}_{A'B'}(-1, +1) + \mathfrak{P}_{A'B'}(-1, -1). \end{aligned}$$

Ist  $N$  die Anzahl der durchgeführten Versuche und  $n(a, b, a', b')$  die Anzahl der Messungen, bei denen die Kombination  $a, b, a', b'$  herauskommt, so folgt

$$\sum n(a, b, a', b') = N.$$

Wahrscheinlichkeiten sind relative Häufigkeiten, daher gilt

$$\begin{aligned} \mathcal{C}(A, B) &= \frac{1}{N} \sum n(a, b, a', b') ab, & \mathcal{C}(A, B') &= \frac{1}{N} \sum n(a, b, a', b') ab', \\ \mathcal{C}(A', B) &= \frac{1}{N} \sum n(a, b, a', b') a'b, & \mathcal{C}(A', B') &= \frac{1}{N} \sum n(a, b, a', b') a'b', \end{aligned}$$

und damit weiter

$$\mathcal{C}(A, B) - \mathcal{C}(A, B') = \frac{1}{N} \sum n(a, b, a', b') ab - \frac{1}{N} \sum n(a, b, a', b') ab'$$



$$\begin{aligned}
&= \frac{1}{N} \sum n(a, b, a', b') (ab - ab') \\
&= \frac{1}{N} \sum n(a, b, a', b') [ab(1 \pm ab')] \\
&\quad - \frac{1}{N} \sum n(a, b, a', b') [ab'(1 \pm a'b)].
\end{aligned}$$

Außerdem finden wir die Abschätzung

$$\begin{aligned}
|\mathcal{C}(A, B) + \mathcal{C}(A, B')| &= \left| \frac{1}{N} \sum n(a, b, a', b') [ab(1 \pm ab')] \right. \\
&\quad \left. - \frac{1}{N} \sum n(a, b, a', b') [ab'(1 \pm a'b)] \right| \\
&\leq \frac{1}{N} \sum n(a, b, a', b') |ab(1 \pm ab')| \\
&\quad + \frac{1}{N} \sum n(a, b, a', b') |ab'(1 \pm a'b)| \\
&\leq \frac{1}{N} \sum n(a, b, a', b') |ab| |1 \pm ab'| \\
&\quad + \frac{1}{N} \sum n(a, b, a', b') |ab'| |1 \pm a'b|.
\end{aligned}$$

Da für die Meßwerte  $|a| = |b| = |a'| = |b'| = 1$  gilt, erhalten wir weiter

$$\begin{aligned}
|\mathcal{C}(A, B) + \mathcal{C}(A, B')| &\leq \frac{2}{N} \sum n(a, b, a', b') \pm \frac{1}{N} \sum n(a, b, a', b') |a'b + a'b'| \\
&\leq 2 \pm [\mathcal{C}(A', B) + \mathcal{C}(A', B')]
\end{aligned}$$

und damit

$$|\mathcal{C}(A, B) - \mathcal{C}(A, B')| + |\mathcal{C}(A', B) + \mathcal{C}(A', B')| \leq 2$$

Das ist eine spezielle Form der *Bellschen Ungleichung*. – Obige Herleitung gilt in völliger Allgemeinheit. Man erhält damit folgende sehr weitreichende Aussage:

*Bellsches Theorem:*

*Jede lokale Theorie mit verborgenen Parametern muß die Bellsche Ungleichung erfüllen.*

Einige Anmerkungen hierzu:

- Die hier abgeleitete Ungleichung ist nicht die Originalversion von Bell<sup>93</sup>. Sie stammt von Clauser, Horne, Shimony und Holt und wird auch CHSH-Ungleichung genannt [102]. Clauser, Horne, Shimony und Holt betrachteten ursprünglich die Größe

$$\mathcal{S} = \mathcal{C}(A, B) - \mathcal{C}(A, B') + \mathcal{C}(A', B) + \mathcal{C}(A', B')$$

und leiteten dafür die Ungleichung

$$-2 \leq \mathcal{S} \leq 2 \quad (4.45)$$

her, die ebenfalls als CHSH-Ungleichung bezeichnet wird und zur obigen Version äquivalent ist. Es gibt weitere Beispiele für Bellsche Ungleichungen.

- Bellsche Ungleichungen haben *nichts* mit Quantenmechanik zu tun. Sie sind *rein klassische* Aussagen über das Verhalten klassischer Theorien.
- Voraussetzung für ihre Gültigkeit ist das Lokalitätspostulat L3.

Jetzt beschreiben wir die Sache aus quantenmechanischer Sicht. Wir betrachten dazu das folgende *Polarisationsexperiment mit verschränkten Photonen*<sup>94</sup>: Gegeben sei ein System, das jeweils diametral auseinanderlaufende verschränkte Photonenpaare erzeugt. Links und rechts von der Photonenquelle Q befinden sich zwei Polarisationsfilter P1 und P2 (siehe Abbildung 4.23). Die Polarisationsrichtungen von P1 und P2 lassen sich unabhängig voneinander in beliebige Richtungen einstellen;  $\alpha_1$  sei der Winkel von P1 zur Vertikalen,  $\alpha_2$  derjenige von P2. Außerdem ist ihre Entfernung zur Quelle jeweils einzeln verstellbar. Zunächst sei P2 weiter von Q entfernt als P1<sup>95</sup>. Ist  $\theta$  der Winkel zwischen Polarisierung eines Photons und Orientierung eines Polarisationsfilters, dann tritt das Photon mit der Wahrscheinlichkeit

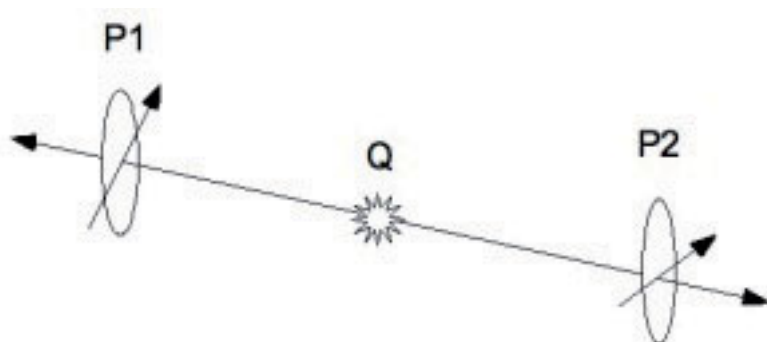


Abbildung 4.23: Schematischer Aufbau des Polarisationsexperiments

<sup>93</sup>Die zugehörigen Originalveröffentlichungen von Bell sind [47] und [48].

<sup>94</sup>Die ersten, die Experimente dieser Art mit Erfolg durchführten, waren Aspect, Grangier und Roger [23], [24]; vergleiche auch [21], [22], [25] und [26]. Inzwischen wurden solche Experimente in vielfältiger Art und großer Präzision durchgeführt. Einen Überblick liefert [54].

<sup>95</sup>Diese Entfernungseinstellung dient nur der besseren Verständlichkeit und hat keinen physikalischen Grund.

$$\mathfrak{P}(\theta) = \cos^2 \theta$$

durch den Polarisator. Der verschränkte Zustand sei wie oben von der Form

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|\uparrow\uparrow\rangle + |\rightarrow\rightarrow\rangle).$$

Als Ausgangsposition seien P1 und P2 senkrecht zueinander orientiert. Wir vereinbaren für die Messungen folgendes: „Photon wird durchgelassen“ führt zum Meßwert  $a = +1$ , „Photon wird nicht durchgelassen“ führt zum Meßwert  $a = -1$ .

Zunächst überlegen wir uns, was an den Polarisationsfiltern jeweils für sich genommen geschieht:

- Ist P1 senkrecht orientiert, so gilt folgendes:
  - Photon 1 wird mit 100 %iger Wahrscheinlichkeit durchgelassen.
  - Die Wellenfunktion des Systems nach der Messung ist  $|\Psi_{\uparrow}\rangle = |\uparrow\uparrow\rangle$ , das heißt, beide Photonen sind nun senkrecht polarisiert.
  - Photon 2 wird ebenfalls mit 100 %iger Wahrscheinlichkeit durchgelassen.
  - Ist P2 stattdessen um  $\theta$  gedreht, wird Photon 2 mit der Wahrscheinlichkeit

$$\mathfrak{P}_{\theta}(+1) = \cos^2 \theta$$

durchgelassen<sup>96</sup>.

- Die Wahrscheinlichkeit, nicht durchgelassen zu werden, ist für Photon 2

$$\mathfrak{P}_{\theta}(-1) = 1 - \cos^2 \theta = \sin^2 \theta.$$

- Wird P1 um einen beliebigen Winkel  $\alpha$  gedreht, so folgt:
  - Photon 1 wird mit der Wahrscheinlichkeit

$$\mathfrak{P}(\alpha) = \cos^2 \alpha$$

durchgelassen.

- Falls Photon 1 durchgelassen wird, ist die Wellenfunktion des Systems nach der Messung  $|\Psi_{\nearrow}\rangle = |\nearrow\nearrow\rangle$ , das heißt, in diesem Fall haben beide Photonen eine um  $\alpha$  gedrehte Polarisation.
- Wird P2 um einen beliebigen Winkel  $\beta$  gedreht, wird Photon 2 mit der Wahrscheinlichkeit

$$\mathfrak{P}_{\alpha\beta}(+1, +1) = \cos^2(\alpha - \beta)$$

durchgelassen.

- Die Wahrscheinlichkeit, daß Photon 2 nicht durchgelassen wird, ist dann

$$\mathfrak{P}_{\alpha\beta}(+1, -1) = \sin^2(\alpha - \beta).$$

<sup>96</sup>Das ist das aus der klassischen Optik bekannte *Gesetz von Malus*.

- Aus Symmetriegründen gilt entsprechend

$$\mathfrak{P}_{\alpha\beta}(-1, +1) = \sin^2(\alpha - \beta).$$

- Die Wahrscheinlichkeit dafür, daß beide Photonen nicht durchgelassen werden, beträgt

$$\mathfrak{P}_{\alpha\beta}(-1, -1) = 1 - \sin^2(\alpha - \beta) = \cos^2(\alpha - \beta).$$

Nun zum eigentlichen Versuch: Es werden sehr viele verschränkte Photonenpaare nacheinander durch die Apparatur geschickt und jeweils links und rechts notiert, ob ein Photon durchgelassen wird oder nicht. Die durchgelassenen und nicht durchgelassenen Photonen werden jeweils gezählt. Wieder betrachten wir Korrelationsexperimente:

- Meßreihe I:
  - Beobachter 1 an P1 mit Orientierung  $\alpha$ ,
  - Beobachter 2 an P2 mit Orientierung  $\beta$ .
- Meßreihe II:
  - Beobachter 1 an P1 mit Orientierung  $\alpha$ ,
  - Beobachter 2 an P2 mit Orientierung  $\beta'$ .
- Meßreihe III:
  - Beobachter 1 an P1 mit Orientierung  $\alpha'$ ,
  - Beobachter 2 an P2 mit Orientierung  $\beta$ .
- Meßreihe IV:
  - Beobachter 1 an P1 mit Orientierung  $\alpha'$ ,
  - Beobachter 2 an P2 mit Orientierung  $\beta'$ .

Damit berechnen wir die Korrelationsfunktionen  $\mathcal{C}(\alpha, \beta)$ :

$$\begin{aligned} \mathcal{C}(\alpha, \beta) &= \mathfrak{P}_{\alpha\beta}(+1, +1) - \mathfrak{P}_{\alpha\beta}(+1, -1) - \mathfrak{P}_{\alpha\beta}(-1, +1) + \mathfrak{P}_{\alpha\beta}(-1, -1) \\ &= 2 \cos^2(\alpha - \beta) - 2 \sin^2(\alpha - \beta). \end{aligned}$$

Mit einem der Additionstheoreme für trigonometrische Funktionen, nämlich

$$\cos(\phi + \theta) = \cos \phi \cos \theta - \sin \phi \sin \theta,$$

finden wir

$$\mathcal{C}(\alpha, \beta) = \cos [2(\alpha - \beta)]$$

und analog auch

$$\mathcal{C}(\alpha, \beta') = \cos [2(\alpha - \beta')]$$

$$\mathcal{C}(\alpha', \beta) = \cos [2(\alpha' - \beta)]$$

$$\mathcal{C}(\alpha', \beta') = \cos [2(\alpha' - \beta')]$$

Das bauen wir zusammen und erhalten

$$\begin{aligned}
& |\mathcal{C}(\alpha, \beta) - \mathcal{C}(\alpha, \beta')| + |\mathcal{C}(\alpha', \beta) - \mathcal{C}(\alpha', \beta')| \\
&= |\cos [2(\alpha - \beta)] - \cos [2(\alpha - \beta')]| + |\cos [2(\alpha' - \beta)] + \cos [2(\alpha' - \beta')]|
\end{aligned}$$

Nun wählen wir speziell

$$\alpha = 0^\circ, \beta = 22,5^\circ, \alpha' = 45^\circ, \beta' = 67,5^\circ$$

oder

$$\alpha = 0^\circ, \beta = 67,5^\circ, \alpha' = 135^\circ, \beta = 202,5^\circ, \beta' = 270^\circ$$

und erhalten in beiden Fällen

$$|\mathcal{C}(\alpha, \beta) - \mathcal{C}(\alpha, \beta')| + |\mathcal{C}(\alpha', \beta) - \mathcal{C}(\alpha', \beta')| = \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} = 2\sqrt{2} > 2.$$

also eine *signifikante Verletzung der Bellschen Ungleichung*. Abbildung 4.24 zeigt die Abhängigkeit der Größe

$$\mathcal{S} = \mathcal{C}(\alpha, \beta) - \mathcal{C}(\alpha, \beta') + \mathcal{C}(\alpha', \beta) - \mathcal{C}(\alpha', \beta')$$

vom Differenzwinkel  $\theta = \alpha - \beta = \alpha - \beta' = \alpha' - \beta = \alpha' - \beta'$ ; nach (4.45) bedeutet das Über- beziehungsweise Unterschreiten der gestrichelten Linien eine Verletzung der Bellschen Ungleichung, das heißt, die schraffierten Bereiche sind die nur für die Quantenmechanik, nicht aber für lokale Theorien mit verborgenen Variablen erreichbaren Bereiche. Die hier explizit diskutierten Winkel stellen gerade das Maximum beziehungsweise das Minimum dieser Kurve dar.

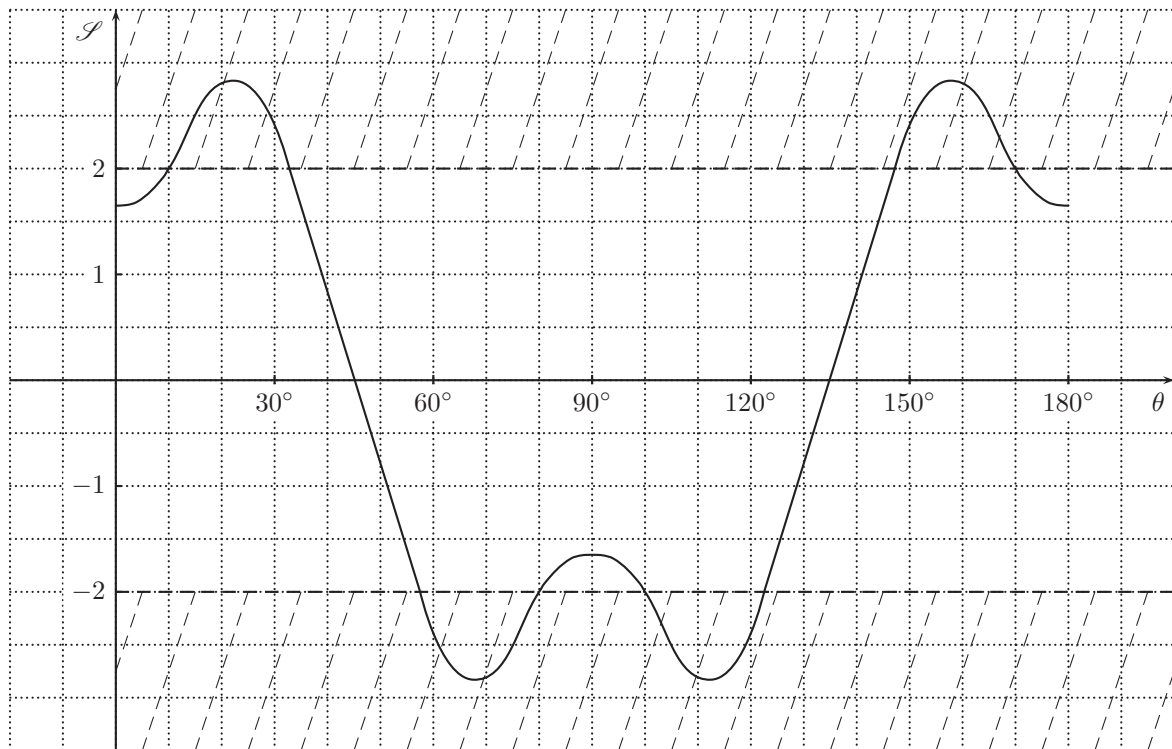


Abbildung 4.24: Verletzung der Bellschen Ungleichung durch die Quantenmechanik

Dieses Resultat läßt sich verallgemeinern:

*Bei sehr häufigen Wiederholungen von Messungen an verschränkten Systemen sagt die Quantenmechanik stärkere Korrelationen voraus als jede mögliche lokale Theorie mit verborgenen Variablen.*

Damit werden lokale Theorien mit verborgenen Variablen einerseits und die Quantenmechanik andererseits *experimentell unterscheidbar*. Die reale Durchführung des beschriebenen Polarisationsexperimentes sowie vielfältige weitere Experimente mit verschränkten Systemen sprechen ausnahmslos und eindeutig *für die Quantenmechanik*. Daraus folgt:

- Lokale Theorien mit verborgenen Variablen sind *sämtlich falsch*.
- Eine Rückkehr zur klassischen Physik mit Hilfe von lokalen Theorien mit verborgenen Variablen ist *nicht möglich*.

Oder kurz und präzise formuliert:

*Die Quantenmechanik ist die unvermeidliche und richtige Sprache zur Beschreibung der Natur.*

Die Bedeutung der Quantenmechanik geht damit über diejenige einer gewöhnlichen physikalischen Theorie im Sinne einer Sammlung von Modellen zur Beschreibung irgendwelcher Aspekte der Natur hinaus. Sie ist eine Sprache zum Formulieren von Theorien und beinhaltet ihre eigene universelle physikalische Gültigkeit: Jede grundlegende physikalische Theorie, die auf das Auftreten nichtlokaler, prinzipiell nicht nachweisbarer und damit unphysikalischer Größen verzichten will, muß in der Sprache der Quantenmechanik formuliert werden.

Einige abschließende Anmerkungen mögen gleichzeitig als Ausblick dienen:

- Die hier beschriebene Darstellung ist stark elementarisiert. Das Bellsche Theorem läßt sich auf mathematisch präzise Weise formulieren und beweisen. Die Experimente, die zeigen, daß die Quantenmechanik triumphiert, sind teilweise sehr aufwendig.
- Die historisch erste Diskussion von Korrelationen im Zusammenhang mit verschränkten Zuständen stammt von Einstein, Podolski und Rosen, weswegen die gesamte Thematik auch als EPR-Paradoxon bezeichnet wird<sup>97</sup>.
- Das bloße Vorhandensein von Korrelationen bei zusammengesetzten Systemen allein ist noch kein Argument für die Quantenmechanik und gegen jedwede klassische Alternativtheorien. Entscheidend ist die Aussage, daß solche Korrelationen bei speziellen Systemen stärker sind, als jede klassische Theorie voraussagen kann, und daß sie von der Quantenmechanik richtig vorausgesagt werden.
- Es gibt weitere Theoreme, die auf andere Art als das Bellsche Theorem eine Entscheidung zwischen Quantenmechanik einerseits und beliebigen lokalen Theorien mit verborgenen Variablen andererseits ermöglichen. Ein Beispiel hierfür ist das *Kochen-Specker-Theorem*, bei dem Widersprüche zwischen lokalen Theorien

---

<sup>97</sup>Siehe auch Anmerkung 84.

mit verborgenen Variablen und der Quantenmechanik hinsichtlich gewisser Wahrscheinlichkeitsaussagen für Spin-1-Systeme betrachtet werden [316]<sup>98</sup>. Besondere Beachtung verdienen insbesondere auch solche Resultate, die nicht nur Unterschiede bei statistischen Aussagen über Meßreihen, sondern Unterschiede bei einzelnen Messungen voraussagen. Das wichtigste Beispiel hierzu ist das *Greenberger-Horne-Zeilinger-Theorem* [198], [199], [200], das ein Experiment mit den in Abschnitt 4.5.7.3 erwähnten GHZ-Zuständen beschreibt, für welches lokale Theorien mit verborgenen Parametern und die Quantenmechanik für jede einzelne Durchführung unterschiedliche Resultate voraussagen. Generell gilt hierbei: Soweit entsprechende Experimente bereits technisch durchführbar sind, gewinnt immer die Quantenmechanik<sup>99</sup>.

- Es ist möglich, die Resultate der Quantenmechanik durch sogenannte nichtlokale Theorien mit verborgenen Variablen zu reproduzieren. Solche Theorien sind jedoch nicht nur mindestens so weit, wenn nicht weiter, von der klassischen Physik entfernt wie die Quantenmechanik selbst, sodaß sich jeder Anhänger der klassischen Physik ernsthaft fragen muß, was damit überhaupt gewonnen sein soll. Sie weisen auch formale, erkenntnistheoretische und physikalische Ungereimtheiten auf, die es als höchst zweifelhaft erscheinen lassen, daß sie ernstzunehmende Konkurrenten der Quantenmechanik sein könnten<sup>100</sup>.
- Oben wurden nur verschränkte Systeme aus räumlich separierten Komponenten betrachtet. Das läßt sich verallgemeinern, was die Verwendung des Begriffs „Nichtseparierbarkeit“ anstelle von „Nichtlokalität“ nahelegt.
- Das Prinzip der Nichtseparierbarkeit ist eine direkte Folge des quantenmechanischen Superpositionsprinzips. Die Existenz interferenzfähiger Superpositionen erweist sich dabei erneut als wichtigster Unterschied zwischen Quantenmechanik und klassischer Physik.
- Die Frage, warum wir im meso- und makroskopischen Bereich eine klassische Welt wahrnehmen und nicht ständig und überall irgendwelche verrückten Superpositionen aus sich eigentlich widersprechenden Zuständen, läßt sich womöglich mit dem Prinzip der Verschränkung beantworten. Die Verschränkung praktisch jeden Systems mit seiner Umgebung macht solche Superpositionen bei Systemen oberhalb des mikroskopischen Bereichs in den meisten Fällen unbeobachtbar. Man nennt dieses Phänomen *Dekohärenz*. Die Wahrnehmung einer klassischen Welt ist somit

<sup>98</sup>Vergleiche auch [489]. Ausführliche Informationen dazu findet man in [108] und [424]; neuere Entwicklungen werden in [315] beschrieben, siehe auch hierzu [108].

<sup>99</sup>Das erste Experiment dierser Art gelang Bouwmeester et al. [73] mit Drei-Teilchen-GHZ-Zuständen der Form

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|\uparrow\uparrow\uparrow\rangle + |\downarrow\downarrow\downarrow\rangle);$$

vergleiche auch [541]. Ein Experiment von Pan und Zeilinger [404] liefert GHZ-Korrelationen bei den Vier-Teilchen-GHZ-Zuständen

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|\uparrow\downarrow\downarrow\uparrow\rangle - |\downarrow\uparrow\uparrow\downarrow\rangle).$$

Weitere Informationen hierzu findet man in [54] und in [442].

<sup>100</sup>Siehe auch Anmerkung 120 aus Seite 172.



ein *rein quantenmechanisches Phänomen*. Das in Abschnitt 4.5.4.2 andeutungsweise erwähnte Problem des quantenmechanischen Meßprozesses kann dadurch jedoch nicht gelöst werden. Dekohärenz kann zwar für den Beobachter einen scheinbaren Kollaps der Wellenfunktion bewirken, die Wellenfunktion von System und Umgebung bleibt jedoch unreduziert. Insbesondere läßt sich mit der Theorie der Dekohärenz die Dynamik quantenmechanischer Meßprozesse insofern nicht beschreiben, als unklar bleibt, wie die Auswahl des erhaltenen Meßwertes aus dem gesamten Spektrum der möglichen Meßwerte funktioniert<sup>101</sup>.

- In der Quantenfeldtheorie findet man ebenfalls maximale Verletzung der Bellschen Ungleichungen [502]. Darüberhinaus tauchen dort noch *sehr viel stärkere* Nichtlokalitäts- und Nichtseparierbarkeitseigenschaften auf. Ein Beispiel für extreme Nichtlokalitätseffekte der Quantenfeldtheorie ist das *Reeh-Schlieder-Theorem* [430], das besagt, daß jeder Vektor im Hilbertraum aller Zustände beliebig genau durch Zustände approximiert werden kann, die man erhält, indem man geeignete, auf einer beliebigen beschränkten offenen Teilmenge der Raumzeit definierte Observable auf den Vakuumzustand anwendet<sup>102</sup>, und aus dem unter anderem folgt, daß  $N$ -Teilchen-Zustände niemals auf beschränkte Raumzeit-Gebiete lokalisiert sein können, sondern stets auf der gesamten Raumzeit nicht verschwinden<sup>103</sup>.
- Die quantenmechanische Nichtseparierbarkeit könnte ein Hinweis darauf sein, daß die Natur eine grundlegendere Struktur als diejenige der Raum-Zeit haben und letztere nur ein abgeleitetes Konzept beziehungsweise unsere Art der Wahrnehmung dieser grundlegenderen Struktur sein könnte.

## Aufgaben

1. Zeige, daß der verschränkte Zustand

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|\uparrow\uparrow\rangle + |\rightarrow\rightarrow\rangle)$$

normiert ist. Dabei darf vorausgesetzt werden, daß die Komponentenzustände jeweils normiert und paarweise orthogonal sind.

2. Überprüfe, ob die Wellenfunktionen

$$|\Psi_1\rangle = \frac{1}{2} (|\uparrow\uparrow\rangle - |\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle - |\downarrow\downarrow\rangle),$$

<sup>101</sup>Ein Versuch in diese Richtung wird in [385] beschrieben.

<sup>102</sup>Die genaue Formulierung lautet folgendermaßen:

*Ist  $\Omega$  der Vakuumzustand,  $\mathcal{G}$  ein beliebiges offenes Gebiet in der Raumzeit und  $\mathcal{R}(\mathcal{G})$  die Menge aller Observablen auf  $\mathcal{G}$ , dann ist  $\{\hat{A}\Omega \mid \hat{A} \in \mathcal{R}(\mathcal{G})\}$  dicht im Hilbertraum  $\mathcal{H}$  aller Zustände.*

Für Details siehe [16], [142], [208] oder [496].

<sup>103</sup>Genauer gesagt können die zugehörigen Projektionsoperatoren niemals Elemente einer zu irgendeinem beschränkten Gebiet gehörenden Operator-Algebra sein. Näheres hierzu steht in [425]; vergleiche auch [217].

$$|\Psi_2\rangle = \frac{1}{2} (|\uparrow\uparrow\rangle + |\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle - |\downarrow\downarrow\rangle)$$

Produktzustände oder verschränkte Zustände beschreiben.

Hinweis: Zeige, daß man diese Zustände auch in der Form

$$\begin{aligned} |\Psi_1\rangle &= \frac{1}{\sqrt{2}} (|\uparrow\rangle + |\downarrow\rangle) \otimes \frac{1}{\sqrt{2}} (|\uparrow\rangle - |\downarrow\rangle), \\ |\Psi_2\rangle &= \frac{1}{2} \left[ \frac{1}{\sqrt{2}} (|\uparrow\rangle + |\downarrow\rangle) \otimes \frac{1}{\sqrt{2}} (|\uparrow\rangle + |\downarrow\rangle) \right. \\ &\quad + \frac{1}{\sqrt{2}} (|\uparrow\rangle + |\downarrow\rangle) \otimes \frac{1}{\sqrt{2}} (|\uparrow\rangle - |\downarrow\rangle) \\ &\quad + \frac{1}{\sqrt{2}} (|\uparrow\rangle - |\downarrow\rangle) \otimes \frac{1}{\sqrt{2}} (|\uparrow\rangle - |\downarrow\rangle) \\ &\quad \left. + \frac{1}{\sqrt{2}} (|\uparrow\rangle - |\downarrow\rangle) \otimes \frac{1}{\sqrt{2}} (|\uparrow\rangle + |\downarrow\rangle) \right] \end{aligned}$$

schreiben kann.

3. Finde weitere Beispiele für verschränkte Zustände aus drei Teilchen. Was geschieht, wenn man an diesen Zuständen jeweils Größen mißt, für welche die Komponenten der verschränkten Zustände Eigenzustände darstellen?
4. Beschreibe den Vorgang der Messung der Polarisation an einem verschränkten System im Zustand

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|\uparrow\rightarrow\rangle + |\rightarrow\uparrow\rangle).$$

5. Erläutere die Aussage: „Messungen zerstören Korrelationen.“
6. Begründe mit Hilfe der ausschließlich statistischen Nachweismöglichkeiten von Korrelationseffekten bei verschränkten Systemen, daß solche Systeme keine Informationsübertragung mit Überlichtgeschwindigkeit ermöglichen.
7. Betrachte das oben beschriebene Polarisationsexperiment für drei beliebige Winkel  $\alpha, \beta, \gamma$ .

- a) Zeige, daß hierfür eine Bellsche Ungleichung der Form

$$\mathfrak{P}_{\alpha\beta}(+1, +1) \leq \mathfrak{P}_{\alpha\gamma}(+1, +1) + \mathfrak{P}_{\beta\gamma}(+1, -1)$$

gilt<sup>104</sup>.

- b) Finde drei Winkel, für die diese Bellsche Ungleichung quantenmechanisch verletzt wird. Weise dies explizit rechnerisch nach.

<sup>104</sup>Das ist eine vereinfachte Version der ursprünglichen, von Bell selbst gefundenen Ungleichung.

8. Gegeben seien drei Teilchen mit Spin  $1/2$ , die sich gerade an den drei unterschiedlichen Orten  $A, B, C$  aufhalten. Für jedes Teilchen bezeichne  $|\uparrow_1\rangle$  den Zustand mit Spin aufwärts in  $x_1$ -Richtung,  $|\uparrow_2\rangle$  denjenigen mit Spin aufwärts in  $x_2$ -Richtung und  $|\uparrow_3\rangle$  denjenigen mit Spin aufwärts in  $x_3$ -Richtung.  $|\downarrow_1\rangle, |\downarrow_2\rangle$  und  $|\downarrow_3\rangle$  sind die entsprechenden Zustände mit Spin abwärts. Die drei Teilchen seien in einem GHZ-Zustand, der gleichermaßen durch die äquivalenten Wellenfunktionen

$$\begin{aligned} |\Psi_{ABC}\rangle &= \frac{1}{\sqrt{2}} (|\uparrow_{3A}\uparrow_{3B}\uparrow_{3C}\rangle + |\downarrow_{3A}\downarrow_{3B}\downarrow_{3C}\rangle) \\ &= \frac{1}{2} (|\uparrow_{2A}\downarrow_{2B}\uparrow_{1C}\rangle + |\downarrow_{2A}\uparrow_{2B}\uparrow_{1C}\rangle + |\uparrow_{2A}\uparrow_{2B}\downarrow_{1C}\rangle + |\downarrow_{2A}\downarrow_{2B}\downarrow_{1C}\rangle) \\ &= \frac{1}{2} (|\uparrow_{1A}\uparrow_{1B}\uparrow_{1C}\rangle + |\uparrow_{1A}\downarrow_{1B}\downarrow_{1C}\rangle + |\downarrow_{1A}\uparrow_{1B}\downarrow_{1C}\rangle + |\downarrow_{1A}\downarrow_{1B}\uparrow_{1C}\rangle) \end{aligned}$$

beschrieben wird. Bei Messungen des Spins in einer beliebigen Richtung kommen stets die Meßwerte  $s_{iJ} = +1$  für Spin aufwärts und  $s_{iJ} = -1$  für Spin abwärts heraus, mit  $i = 1, 2, 3$  und  $J = A, B, C$ . Gemessen wird stets an allen drei Teilchen gleichzeitig eine beliebige Kombination der Spinrichtungen<sup>105</sup>.

- a) Zeige, daß gemäß der Quantenmechanik für alle möglichen Meßwerte

$$s_{2A} s_{2B} s_{1C} = s_{1A} s_{2B} s_{2C} = s_{2A} s_{1B} s_{2C} = -1, \quad s_{1A} s_{1B} s_{1C} = +1$$

gilt.

- b) Zeige, daß jede lokale Theorie mit verborgenen Variablen im Widerspruch zu Teilaufgabe a)

$$s_{1A} s_{1B} s_{1C} = -1$$

liefert.

- c) Experimentell bestätigt sich das Ergebnis der Quantenmechanik. Warum ist der hier beschriebene Sachverhalt ein stärkeres Kriterium zu ihren Gunsten als das in Abschnitt 4.5.7.4 beschriebene Polarisationsexperiment?

<sup>105</sup>Genauereres dazu steht in [198], [199] und [200]; vergleiche auch [27], [187] und [404].

# Anhang A

## Der thermodynamische Limes

Die thermodynamische Betrachtungsweise physikalischer Systeme, wie sie in den Abschnitten 1.2.4, 1.2.5 und 1.3.1.1 diskutiert wurde, stellt, sofern sie sachgemäß auf geeignete Konfigurationen angewendet wird, eines der wirksamsten Werkzeuge zur Beschreibung der Natur dar. Dabei wird im allgemeinen zwar berücksichtigt, daß man es genauesehen stets mit einer Grenzwertbetrachtung zu tun hat, nämlich dem Limes von Systemen mit unendlich vielen Teilchen, es wird jedoch so gut wie nie die Frage gestellt, ob dieser Limes in den einzelnen Fällen überhaupt existiert. Nun ist es in der theoretischen Physik üblich, solche Fragen generell auszublenden, insbesondere, wenn man daran interessiert ist, tatsächlich konkrete Resultate zu erzielen; dennoch erscheint es gerade aus der Sicht der mathematischen Physik besonders interessant beziehungsweise ist es geradezu eine ihrer Hauptaufgabenstellungen, Existenzfragen dieser Sorte zu beantworten. Ähnlich wie bei den Pfadintegralen wird auch hier in exemplarischer Weise deutlich, wie verbreitet angewendete und vermeintlich nahezu triviale Grenzwertbetrachtungen bei genauerer Betrachtung höchst subtile Angelegenheiten sind. Deshalb wenden wir uns hier noch ein wenig der mathematischen Thermodynamik zu; der Vergleich mit der elementaren Darstellung der oben erwähnten Abschnitte ist lehrreich und sei der Leserin und dem Leser nahegelegt<sup>1</sup>.

### A.1 Thermodynamische Systeme

Unter einem *klassischen thermodynamischen System*  $S$  verstehen wir hier ein Ensemble aus einer sehr großen Anzahl identischer Subsysteme, die sich im Konfigurationsraum  $\mathbb{R}^d$  bewegen und über zusätzliche innere Freiheitsgrade verfügen können. Die Subsysteme werden der Einfachheit halber als *Teilchen* bezeichnet. Zusätzlich sei für jedes  $N \in \mathbb{N}$  eine translationsinvariante Funktion

$$V: \prod_{i=1}^N \mathbb{R}^d \longrightarrow \mathbb{R} \cup \{\infty\}$$

---

<sup>1</sup>Erste Existenzbeweise für den thermodynamischen Limes wurden unabhängig voneinander von Van Hove [254] sowie von Yang und Lee [533] veröffentlicht. Der hier beschriebene Beweis wurde von Ruelle gefunden [437]. Wir folgen hier im wesentlichen dessen Monographie [439]. Weitere Details findet man dort; vergleiche auch [173], [440] und [505].

definiert, die *potentielle Energie*: Befinden sich  $N$  Teilchen der Masse  $m$  an den Orten  $x_1, x_2, \dots, x_N \in \mathbb{R}^d$ , dann ist deren potentielle Energie durch  $V(x_1, x_2, \dots, x_N)$  gegeben. Für jedes  $N \in \mathbb{N}$  läßt sich  $V$  in der Form

$$V(x_1, x_2, \dots, x_N) = \sum_{j \geq 2} \sum_{1 \leq i_1 < \dots < i_j \leq N} \Phi^j(x_{i_1}, x_{i_2}, \dots, x_{i_j})$$

darstellen, mit reellwertigen, Lebesgue-meßbaren, translations- und permutationsinvarianten Funktionen

$$\Phi^j: \prod_{i=1}^j \mathbb{R}^d \longrightarrow \mathbb{R} \cup \{\infty\}.$$

Die  $\Phi^j$  nennt man  $j$ -Körper-Potentiale; die äußere Summe beginnt bei 2, weil man aufgrund der beliebigen Festlegbarkeit des Nullpunkts der potentiellen Energie stets  $\Phi^0 = \Phi^1 = 0$  setzen kann. Die Wechselwirkungen der Teilchen eines thermodynamischen Systems werden durch die zugehörige Folge  $\{\Phi^j\}_{j \geq 2}$  von  $j$ -Körper-Potentialen eindeutig bestimmt.

Denken wir uns ein System  $\mathcal{S}$  mit  $n$  Teilchen in zwei Teilsysteme  $\mathcal{S}_1$  und  $\mathcal{S}_2$  mit  $N_1$  beziehungsweise  $N_2$  Teilchen zerlegt, so gilt für dessen potentielle Energie

$$\begin{aligned} V(x'_1, x'_2, \dots, x'_{N_1}, x''_1, x''_2, \dots, x''_{N_2}) &= V(x'_1, x'_2, \dots, x'_{N_1}) + V(x''_1, x''_2, \dots, x''_{N_2}) \\ &\quad + V_{N_1, N_2}(x'_1, x'_2, \dots, x'_{N_1}, x''_1, x''_2, \dots, x''_{N_2}), \end{aligned}$$

mit

$$\begin{aligned} &V_{N_1, N_2}(x'_1, x'_2, \dots, x'_{N_1}, x''_1, x''_2, \dots, x''_{N_2}) \\ &= \sum_{k \geq 1} \sum_{l \geq 1} \sum_{1 \leq i_1 < \dots < i_k \leq N_1} \sum_{1 \leq j_1 < \dots < j_l \leq N_2} \Phi^{k+l}(x'_1, x'_2, \dots, x'_{N_1}, x''_1, x''_2, \dots, x''_{N_2}). \end{aligned}$$

Damit können wir einige wesentliche Begriffe formulieren. Wir beginnen mit zwei zusätzlichen Eigenschaften für Potentiale.

**Definition:**  $\{\Phi^j\}_{j \geq 2}$  sei eine Folge von  $j$ -Körper-Potentialen wie oben.

- (i)  $\{\Phi^j\}_{j \geq 2}$  heißt *temperiert*, wenn es Konstanten  $\alpha > d, A \geq 0$  und  $R > 0$  gibt, so daß

$$V_{n_1, n_2}(x'_1, x'_2, \dots, x'_{n_1}, x''_1, x''_2, \dots, x''_{n_2}) \leq A n_1 n_2 r^{-\alpha}$$

gilt für alle  $\|x''_j - x'_i\| \geq r \geq R$  und alle  $i = 1, 2, \dots, N_1, j = 1, 2, \dots, N_2$ .

- (ii)  $\{\Phi^j\}_{j \geq 2}$  heißt *stabil*, wenn es eine Konstante  $B \geq 0$  gibt, so daß

$$V(x_1, x_2, \dots, x_n) \geq -NB$$

gilt für alle  $N \in \mathbb{N}$  und alle  $x_1, x_2, \dots, x_N \in \mathbb{R}^d$ .

Außerdem führen wir zwei grundlegende Begriffe der statistischen Mechanik ein. Dazu seien  $p_1, p_2, \dots, p_N \in (\mathbb{R}^d)^*$  die Impulse der Teilchen.

**Definition:**  $\Lambda$  sei eine beschränkte Lebesgue-meßbare Teilmenge des  $\mathbb{F}^d$  vom Lebesgue-maß  $\mathbf{m}(\Lambda)$ , außerdem sei  $E$  die Gesamtenergie des betrachteten Systems.

(i) Unter der *mikrokanonischen Zustandssumme* versteht man die Größe

$$\begin{aligned}\Omega(\Lambda, N, E) &= \frac{1}{N!} \int_{\Lambda} \int_{\Lambda} \cdots \int_{\Lambda} \chi_{[-\infty, 0]} \left( \sum_{i=1}^N \frac{p_i^2}{2m} + V(x_1, x_2, \dots, x_N) - E \right) \times \\ &\quad \times dx_1 dx_2 \cdots dx_N \\ &= \frac{1}{N!} \mathbf{m} \left\{ (x_1, x_2, \dots, x_N) \subset \Lambda^N \left| \sum_{i=1}^N \frac{p_i^2}{2m} + V(x_1, x_2, \dots, x_N) \leq E \right. \right\}.\end{aligned}$$

(ii) Die *Konfigurations-Entropie* ist definiert durch

$$S(\Lambda, N, E) = \ln \Omega(\Lambda, N, E).$$

$S$  ist eine monoton steigende Funktion von  $\Lambda$  und  $E$ , so daß sie für festes  $\Lambda$  und festes  $N$  invertierbar ist. Die resultierende Funktion  $E(\Lambda, N, S)$  ist folglich eine monoton steigende Funktion von  $S$  und damit eine monoton fallende Funktion von  $\Lambda$ . Mit  $d(\Lambda)$  bezeichnen wir wie üblich das Supremum der Abstände zweier Punkte aus  $\Lambda$ .

Wir benötigen noch eine wichtige Konsequenz aus der Temperiertheit von Potentialen, und zwar den folgenden

**Satz:** Ist die Potentialfolge  $\{\Phi^j\}_{j \geq 2}$  temperiert, so gilt für disjunkte Gebiete

$\Lambda_1, \Lambda_2, \dots, \Lambda_n \subset \mathbb{F}^d$  mit gegenseitigem Abstand  $r \geq R$

$$E \left( \bigcup_{i=1}^n \Lambda_i, \sum_{i=1}^n N_i, \sum_{i=1}^n S_i \right) \leq \sum_{i=1}^n E(\Lambda_i, N_i, S_i) + \frac{A}{2r^\alpha} \left( \sum_{i=1}^n N_i \right)^2. \quad (\text{A.1})$$

BEWEIS: Sei zunächst  $n = 2$  und  $x'_1, x'_2, \dots, x'_{N_1} \in \Lambda_1$ ,  $x''_1, x''_2, \dots, x''_{N_2} \in \Lambda_2$  so daß  $V(x'_1, x'_2, \dots, x'_{N_1}) \leq E_1$  und  $V(x''_1, x''_2, \dots, x''_{N_2}) \leq E_2$ . Dann folgt aus der Temperiertheitsbedingung

$$\begin{aligned}V(x'_1, x'_2, \dots, x'_{N_1}, x''_1, x''_2, \dots, x''_{N_2}) &= V(x'_1, x'_2, \dots, x'_{N_1}) + V(x''_1, x''_2, \dots, x''_{N_2}) \\ &\quad + V_{N_1, N_2}(x'_1, x'_2, \dots, x'_{N_1}, x''_1, x''_2, \dots, x''_{N_2}) \\ &\leq E_1 + E_2 + \frac{AN_1N_2}{r^\alpha}.\end{aligned}$$

Die Anzahl der Konfigurationsmöglichkeiten in  $\Lambda_1 \cup \Lambda_2$ , für welche die potentielle Energie kleiner ist als  $E_1 + E_2 + AN_1N_2/r^\alpha$ , ist somit größer als das Produkt der Anzahl der Konfigurationsmöglichkeiten in  $\Lambda_1$ , für welche die potentielle Energie kleiner als  $E_1$  ist, multipliziert mit der Anzahl der Konfigurationsmöglichkeiten in  $\Lambda_2$ , für welche sie kleiner als  $E_2$  ist, was sich unter Verwendung des Lebesgue-Maßes in der Form

$$\frac{\mathbf{m}(\mathcal{G}_{1;2})}{(N_1 + N_2)!} \geq \frac{\mathbf{m}(\mathcal{G}_1) \mathbf{m}(\mathcal{G}_2)}{N_1! N_2!}$$

schreiben läßt, mit

$$\begin{aligned}\mathcal{G}_1 &= \{(x_1, x_2, \dots, x_{N_1}) \in \Lambda_1 \mid V(x_1, x_2, \dots, x_{N_1}) \leq E_1\}, \\ \mathcal{G}_2 &= \{(x_1, x_2, \dots, x_{N_2}) \in \Lambda_2 \mid V(x_1, x_2, \dots, x_{N_2}) \leq E_2\}, \\ \mathcal{G}_{1;2} &= \left\{ (x_1, x_2, \dots, x_{N_1+N_2}) \in \Lambda_1 \cup \Lambda_2 \mid \right. \\ &\quad \left. V(x_1, x_2, \dots, x_{N_1+N_2}) \leq E_1 + E_2 + \frac{AN_1N_2}{r^\alpha} \right\}.\end{aligned}$$

Mit der Definition der mikrokanonischen Zustandssumme  $\Omega(\Lambda, N, E)$  ergibt sich daraus unmittelbar

$$\Omega\left(\Lambda_1 \cup \Lambda_2, N_1 + N_2, E_1 + E_2 + \frac{AN_1N_2}{r^\alpha}\right) \geq \Omega(\Lambda_1, N_1, E_1) \Omega(\Lambda_2, N_2, E_2),$$

mit derjenigen der Konfigurations-Entropie weiter

$$S\left(\Lambda_1 \cup \Lambda_2, N_1 + N_2, E_1 + E_2 + \frac{AN_1N_2}{r^\alpha}\right) \geq S(\Lambda_1, N_1, E_1) + S(\Lambda_2, N_2, E_2),$$

und da  $S$  und  $E$  monoton steigende Umkehrfunktionen voneinander sind, liefert das

$$E(\Lambda_1 \cup \Lambda_2, N_1 + N_2, E_1 + E_2) \leq E(\Lambda_1, N_1, E_1) + E(\Lambda_2, N_2, E_2) + \frac{AN_1N_2}{r^\alpha}.$$

Wendet man diese Ungleichung  $n - 1$  mal an und benutzt außerdem die Relation

$$2 \sum_{i < j} N_i N_j \leq \left( \sum_i N_i \right)^2,$$

so folgt die Behauptung. □

## A.2 Existenz des thermodynamischen Limes

Wir kommen jetzt zum Beweis der Existenz des thermodynamischen Limes. Dies wird zunächst für die mikrokanonische Gesamtheit gezeigt und sodann daraus auch für die kanonische und die großkanonische Gesamtheit gefolgert. Der Beweis für die mikrokanonische Gesamtheit läßt sich in zwei Aspekte unterteilen, einen eher technischen und einen, welcher die physikalischen Konsequenzen zum Ausdruck bringt. Dazu sei  $\{\Phi^j\}_{j \geq 2}$  eine temperierte, stabile Folge von Potentialen. Inhalt des ersten Aspektes ist der folgende

**Satz:** *Es seien  $\theta$  eine Konstante mit  $1 < 2^{d/\alpha} < \theta < 2$ , außerdem  $\mathcal{G} \subset \mathbb{R}^2$  ein konvexes Gebiet,  $\mathcal{E} : \mathcal{G} \rightarrow \mathbb{R}$  eine stetige, konvexe, im ersten Argument monoton steigende Funktion mit  $\mathcal{E}(\xi, \eta) = +\infty$  für  $(\xi, \eta) \notin \overline{\mathcal{G}}$  und  $\{\Lambda\} \subset \mathcal{P}(\mathbb{R}^d)$  eine unendliche Familie Lebesgue-meßbarer Mengen mit folgenden Eigenschaften:*



(i)  $\lim \mathfrak{m}(\Lambda) = \infty$

(ii) Es gibt eine reelle Funktion  $g$  mit  $\lim_{\xi \rightarrow 0} g(\xi) = 0$ , so daß für genügend kleine  $\varepsilon$  und alle  $\Lambda$  die Relation

$$\frac{\mathfrak{m}\{x \in \Lambda \mid \forall y \in \partial\Lambda \ \|x - y\| \leq \varepsilon d(\Lambda)\}}{\mathfrak{m}(\Lambda)} \leq g(\xi)$$

erfüllt ist.

(iii)  $\lim_{\Lambda \rightarrow \infty} N/\mathfrak{m}(\Lambda) = \xi/L$  und  $\lim_{\Lambda \rightarrow \infty} S/\mathfrak{m}(\Lambda) = \eta/L$  für  $(\xi, \eta) \in \mathcal{G}$  und  $L > R/(2 - \theta)$ .

Dann gilt

a) für  $(\xi, \eta) \in \mathcal{G}$ :

$$\lim_{\Lambda \rightarrow \infty} \frac{E(\Lambda, N, S)}{\mathfrak{m}(\Lambda)} = \frac{\mathcal{E}(\xi, \eta)}{L}.$$

b) für  $(\xi, \eta) \notin \overline{\mathcal{G}}$ :

$$\lim_{\Lambda \rightarrow \infty} \frac{E(\Lambda, N, S)}{\mathfrak{m}(\Lambda)} = +\infty,$$

c) und für  $(\xi, \eta) \in \partial\mathcal{G}$ :

$$\liminf_{\Lambda \rightarrow \infty} \frac{E(\Lambda, N, S)}{\mathfrak{m}(\Lambda)} \geq \liminf_{(\xi', \eta') \rightarrow (\xi, \eta)} \mathcal{E}(\xi', \eta').$$

BEWEIS: a) Der Beweis erfolgt in zwei Schritten. Zunächst beschränken wir uns auf eine spezielle Folge würfelförmiger Gebiete im  $\mathbb{F}^d$ . Dazu sei  $\Lambda_n$  ein Würfel der Kantenlänge  $L_n = 2^n L - \theta^n R/(2 - \theta)$  im  $\mathbb{F}^d$ . Evidenterweise lassen sich  $2^d$  identische Kopien  $\Lambda_{n,i}$  dieses Würfels parallel zueinander im nächstgrößeren Würfel  $\Lambda_{n+1}$  unterbringen, wenn deren gegenseitiger Abstand jeweils höchstens den Wert

$$R_n = L_{n+1} - 2L_n = \frac{(-\theta^{n+1} + 2\theta^n)R}{2 - \theta} = \theta^n R$$

hat; für diesen gilt  $R_n > R$ . Damit haben wir gerade die Situation der Ungleichung (A.1) und können  $E$  abschätzen; wir erhalten

$$E\left(\bigcup_{i=1}^{2^d} \Lambda_{n,i}, \sum_{i=1}^{2^d} N_i, \sum_{i=1}^{2^d} S_i\right) \leq \sum_{i=1}^{2^d} E(\Lambda_n, N_i, S_i) + \frac{A}{2R_n^\alpha} \left(\sum_{i=1}^{2^d} N_i\right)^2.$$

Mit Hilfe der Monotonie von  $E$  als Funktion von  $\Lambda$  folgt daraus

$$E\left(\Lambda_{n+1}, \sum_{i=1}^{2^d} N_i, \sum_{i=1}^{2^d} S_i\right) \leq \sum_{i=1}^{2^d} E(\Lambda_n, N_i, S_i) + \frac{A}{2R_n^\alpha} \left(\sum_{i=1}^{2^d} N_i\right)^2 \tag{A.2}$$

und damit weiter

$$E(\Lambda_{n+1}, 2^d N, 2^d S) \leq 2^d E(\Lambda_n, N, S) + \frac{A}{2R_n^\alpha}. \tag{A.3}$$

Diese Abschätzung verwenden wir nun, um die Existenz eines Grenzwerts von  $E/2^{nd}$  für  $n \rightarrow \infty$  nachzuweisen.

Dazu betrachten wir für  $(\xi, \eta)$  mit  $2^{nd}\xi \in \mathfrak{N}$  und  $\eta \in \mathfrak{E}$  die Funktion

$$a_n(\xi, \eta) = \frac{1}{2^{nd}} E(\Lambda_n, 2^{nd}\xi, 2^{nd}\eta) - \frac{A\xi^2}{2} \sum_{m=0}^{n-1} \frac{2^{(m+1)d}}{R_m^\alpha}.$$

Betrachtet man  $n$  als Laufvariable, so wird daraus eine Folge  $(a_n)_{n \in \mathfrak{N}}$  mit positiven Folgeelementen; da sie gemäß (A.3) auch monoton fällt, ist sie konvergent. Ihr Grenzwert sei  $a(\xi, \eta)$ . Damit gilt es, das Konvergenzverhalten des zweiten Summanden von  $a_n$  zu überprüfen.

Für die Summe findet man

$$\sum_{m=0}^{\infty} \frac{2^{(m+1)d}}{R_m^\alpha} = \frac{2^d}{(2-\theta)^\alpha R^\alpha} \sum_{m=0}^{\infty} \left(\frac{2^d}{\theta^\alpha}\right)^m$$

und damit eine geometrische Reihe; diese konvergiert für  $2^d/\theta^\alpha < 1$ , was nach Voraussetzung der Fall ist. Zusammengenommen folgt daraus, daß für  $2^{nd}\xi \in \mathfrak{N}$  und  $\eta \in \mathfrak{E}$  der Grenzwert

$$\mathcal{E}(\xi, \eta) := \lim_{n \rightarrow \infty} \frac{1}{2^{nd}} E(\Lambda_n, 2^{nd}\xi, 2^{nd}\eta)$$

existiert und die Form

$$\mathcal{E}(\xi, \eta) = a(\xi, \eta) + \frac{A\beta^2}{2} \sum_{m=0}^{\infty} \frac{2^{(m+1)d}}{R_m^\alpha}$$

hat. Dabei ist auch der Fall  $\mathcal{E}(\xi, \eta) = +\infty$  mit eingeschlossen, der genau dann eintritt, wenn auch  $E(\Lambda_n, 2^{nd}\xi, 2^{nd}\eta) = +\infty$  gilt. Der Funktion  $\mathcal{E}$  muß nun die weitere Aufmerksamkeit zugewendet werden; insbesondere muß sie auf beliebige Argumente  $(\xi, \eta) \in \mathcal{G}$  fortgesetzt werden.

Zunächst schätzen wir  $\mathcal{E}$  nach oben ab. Das geschieht unter Verwendung der Ungleichung (A.2). Aus dieser folgt

$$2^{-(n+1)d} E\left(\Lambda_{n+1}, \sum_{i=1}^{2^d} N_i, \sum_{i=1}^{2^d} S_i\right) \leq 2^{-(n+1)d} \left[ \sum_{i=1}^{2^d} E(\Lambda_n, N_i, S_i) + \frac{A}{2R_n^\alpha} \left(\sum_{i=1}^{2^d} N_i\right)^2 \right]$$

und für  $n \rightarrow \infty$  weiter

$$\mathcal{E}\left(\frac{\xi_1 + \xi_2}{2}, \frac{\eta_1 + \eta_2}{2}\right) \leq \frac{\mathcal{E}(\xi_1, \eta_1) + \mathcal{E}(\xi_2, \eta_2)}{2}. \quad (\text{A.4})$$

Wählt man nun ein  $\chi \in \mathfrak{I}$  mit  $\chi = p/2^q$  und  $0 \leq \chi \leq 1$ , so liefert  $q$ -fache Anwendung von (A.4)

$$\mathcal{E}(\chi\xi_1 + (1-\chi)\xi_2, \chi\eta_1 + (1-\chi)\eta_2) \leq \chi\mathcal{E}(\xi_1, \eta_1) + (1-\chi)\mathcal{E}(\xi_2, \eta_2). \quad (\text{A.5})$$

Hieraus erhält man für  $l \leq 2^{nd}$  durch wiederholte Anwendung von (A.2)

$$\mathcal{E}(N\xi_1 + (2^{nd} - N)\xi_2, N\eta_1 + (2^{nd} - N)\eta_2) \leq N\mathcal{E}(\xi_1, \eta_1) + (2^{nd} - N)\mathcal{E}(\xi_2, \eta_2)$$

oder

$$E(\Lambda_n, N, NS_1 + (2^{nd} - N) S_0) \leq NE(\Lambda_0, 1, S_1) + (2^{nd} - N) E(\Lambda_0, 0, S_0) + \frac{NA}{2} \sum_{m=0}^{n-1} \frac{2^{(m+1)d}}{R_m^\alpha}. \tag{A.6}$$

Offensichtlich gilt  $E(\Lambda_0, 0, S_0) = E(\Lambda_0, 1, S_1) = 0$ , falls  $S_0 \leq 0$  und  $S_1 \leq \ln \mathbf{m}(\Lambda_0)$ , wodurch sich (A.6) für  $N \leq 2^{nd}$  und  $S/N \leq \ln \mathbf{m}(\Lambda_0)$  zu

$$E(\Lambda_n, N, S) \leq \frac{NA}{2} \sum_{m=0}^{n-1} \frac{2^{(m+1)d}}{R_m^\alpha} \tag{A.7}$$

vereinfacht. Daraus folgt mit

$$\mathcal{E}(\xi, \eta) \leq \frac{\xi A}{2} \sum_{m=0}^{n-1} \frac{2^{(m+1)d}}{R_m^\alpha} \tag{A.8}$$

für  $0 \leq \xi \leq 1$  und  $\eta \leq \xi \ln \mathbf{m}(\Lambda_0)$  die gewünschte Abschätzung.

Als nächstes müssen wir  $\mathcal{E}(\xi, \eta)$  geeignet erweitern. Genauergesagt zeigen wir, daß sich  $\mathcal{E}(\xi, \eta)$  auf einer geeigneten konvexen Menge  $\mathcal{G} \subset \mathbb{F}^2$  stetig zu einer konvexen, im zweiten Argument monoton steigenden Funktion fortsetzen läßt, für die  $\mathcal{E}(\xi, \eta) = +\infty$  gilt, falls  $(\xi, \eta) \notin \bar{\mathcal{G}}$ . Dazu betrachten wir die Menge

$$\Gamma = \{(\xi, \eta) \in \mathbb{F}^d \mid \mathcal{E}(\xi, \eta) < \infty\}$$

sowie deren Abschluß  $\bar{\Gamma}$  und setzen  $\mathcal{G} = \bar{\Gamma} \setminus \partial\bar{\Gamma}$ . (A.5) lehrt, daß  $\bar{\Gamma}$  und damit auch  $\mathcal{G}$  konvex ist. (A.8) zeigt, daß  $\mathcal{E}$  auf der Menge

$$\mathcal{A} = \{(\xi, \eta) \in \mathbb{F}^d \mid 0 \leq \xi \leq 1 \wedge \eta \leq \xi \ln \mathbf{m}(\Lambda_0)\}$$

nach oben beschränkt ist. Folglich gilt  $\mathcal{A} \subset \mathcal{G}$  und somit  $\mathcal{G} \neq \emptyset$ . Als nächstes zeigen wir, daß  $\mathcal{E}$  in der Umgebung eines jeden Punktes von  $\mathcal{G}$  beschränkt ist. Für alle  $(\xi, \eta) \in \mathcal{G}$  gibt es ein  $(\tilde{\xi}, \tilde{\eta}) \in \Gamma$ , ein  $\chi = p/2^q$  mit  $0 \leq \chi \leq 1$  und ein  $r > 0$ , so daß

$$\{\chi(\tilde{\xi}, \tilde{\eta}) + (1 - \chi)(\hat{\xi}, \hat{\eta}) \mid 0 \leq \hat{\xi} \leq 1 \wedge \hat{\eta} \leq \xi \ln \mathbf{m}(\Lambda_0)\} \supset \{(\xi, \eta) \mid \|(\xi, \eta) - (\xi, \eta)\| < r\}$$

gilt. Gemäß (A.5) ist  $\mathcal{E}$  damit in einer Umgebung eines jeden Punktes von  $\mathcal{G}$  nach oben beschränkt. Die Stabilitätsbedingung der vorgegebenen Potentialfolge liefert gleichzeitig

$$\mathcal{E}(\xi, \eta) \leq -\xi B$$

und damit Beschränktheit von  $\mathcal{E}$  auch nach unten.

Damit kommen wir zur stetigen Fortsetzung der Funktion  $\mathcal{E}$  auf  $\mathcal{G}$ . Dazu genügt es zu zeigen, daß eine Erweiterung von  $\mathcal{E}$  auf eine geeignete Obermenge von  $\mathcal{G}$  dort gleichmäßig

stetig ist. Eine solche findet man folgendermaßen:  $\mathcal{B}_1$  sei eine Kreisscheibe in  $\mathcal{G}$ , so daß es ein  $r > 0$  gibt mit  $\|\mathcal{E}(\xi, \eta)\| < r$  für alle  $(\xi, \eta) \in \mathcal{B}_1$ , dazu sei  $\mathcal{B}_2$  eine Kreisscheibe mit demselben Radius wie  $\mathcal{B}_1$ , aber dem halben Radius. Für jedes  $r$  erhält man ein Paar von Kreisscheiben  $(\mathcal{B}_1, \mathcal{B}_2)$ , und  $\mathcal{G}$  läßt sich mit Kreisscheiben der Form  $\mathcal{B}_2$  überdecken. Weist man nun die gleichmäßige Steigkeit von  $\mathcal{E}$  auf einer Kreisscheibe  $\mathcal{B}_2$  nach, so ist gleichzeitig die Existenz einer stetigen Erweiterung auf ganz  $\mathcal{G}$  gezeigt.

Startpunkt ist wieder die Ungleichung (A.5); angewandt auf

$$\begin{aligned}\xi_1 &= \xi'' + 2^q (\xi' - \xi''), & \eta_1 &= \eta'' + 2^q (\eta' - \eta'') \\ \xi_2 &= \xi'', & \eta_2 &= \eta''\end{aligned}$$

mit  $\delta = 1/2^q$  liefert diese

$$\mathcal{E}(\xi', \eta') \leq \frac{1}{2^q} \mathcal{E}(\xi'' + 2^q (\xi' - \xi''), \eta'' + 2^q (\eta' - \eta'')) + \left(1 - \frac{1}{2^q}\right) \mathcal{E}(\xi'', \eta'')$$

und damit auch

$$\mathcal{E}(\xi', \eta') - \mathcal{E}(\xi'', \eta'') \leq \frac{1}{2^q} [\mathcal{E}(\xi'' + 2^q (\xi' - \xi''), \eta'' + 2^q (\eta' - \eta'')) - \mathcal{E}(\xi'', \eta'')].$$

Vertauscht man die einfach und die doppelt gestrichenen Variablen, erhält man analog

$$\mathcal{E}(\xi'', \eta'') - \mathcal{E}(\xi', \eta') \leq \frac{1}{2^q} [\mathcal{E}(\xi' + 2^q (\xi'' - \xi'), \eta' + 2^q (\eta'' - \eta')) - \mathcal{E}(\xi', \eta')].$$

Nun betrachten wir für  $(\xi', \eta'), (\xi'', \eta'') \in \mathcal{B}_2$  die Grenzübergänge  $\xi' \rightarrow \xi'', \eta' \rightarrow \eta''$  und  $q \rightarrow \infty$ . Für diese gilt  $(\xi'' + 2^q (\xi' - \xi''), \eta'' + 2^q (\eta' - \eta'')) \in \mathcal{B}_1$  und  $(\xi' + 2^q (\xi'' - \xi'), \eta' + 2^q (\eta'' - \eta')) \in \mathcal{B}_1$  und damit auch

$$|\mathcal{E}(\xi', \eta') - \mathcal{E}(\xi'', \eta'')| \leq 2^{1-q} r.$$

Wegen  $q \rightarrow \infty$  ist das gleichbedeutend mit der gleichmäßigen Stetigkeit von  $\mathcal{E}$ . Weil punktweise Konvergenz einer monotonen Funktion gegen eine stetige Funktion auf einem abgeschlossenen Intervall die gleichmäßige Konvergenz nach sich zieht, ist damit auch die gleichmäßige Konvergenz von  $\mathcal{E}(\xi, \eta)$  auf Intervallen der Form  $[\eta_1, \eta_2]$  nachgewiesen.

Jetzt verallgemeinern wir das ganze auf beliebige Gebiete  $\Lambda \subset \mathbb{R}^d$ , indem wir anschaulich gesprochen solche Gebiete in geeigneter Weise mit Würfeln der Form  $\Lambda_n$  anfüllen.

Um die Existenz des Grenzwerts  $\lim_{\Lambda \rightarrow \infty} \frac{E(\Lambda, N, S)}{\mathfrak{m}(\Lambda)}$  nachzuweisen, genügt es, zu zeigen,

daß ein und derselbe Wert gleichzeitig eine obere Schranke des Supremums und eine

untere Schranke des Infimums von  $\frac{E(\Lambda, N, S)}{\mathfrak{m}(\Lambda)}$  ist.

Zur Abschätzung des Supremums wählen wir ein  $N_0 \in \mathbb{N}$  und ein  $\delta > 0$  so, daß  $N_0/2^{nd} = \xi + \delta$  gilt. Ist  $\Lambda$  genügend groß, enthält es  $M + q + 1$  disjunkte Würfel der Kantenlänge  $rL_n$  mit

$$1 < r < \left(\frac{N_0}{2^{nd} \xi}\right)^{1/d},$$

$$\begin{aligned} N &= MN_0 + Q_0, & 0 &\leq Q_0 < N_0, \\ Q_0 &= q 2^{nd} + Q_1, & 0 &\leq Q_1 < 2^{nd}, \\ & & \frac{L_n^d}{N} &\longrightarrow 0, \quad \frac{L_n^\alpha}{N} \longrightarrow \infty. \end{aligned}$$

Damit lassen sich  $M + Q + 1$  Kopien  $\Lambda_{n,i}$  des Würfels  $\Lambda_n$  parallel zueinander in  $\Lambda$  unterbringen, wobei deren gegenseitiger Abstand mindestens  $(r - 1) L_n$  ist. Die Eigenschaft von  $E(\Lambda, N, S)$ , als Funktion von  $\Lambda$  monoton fallend zu sein, sowie (A.1) und (A.7) liefern nun

$$\begin{aligned} E(\Lambda, N, S) &\leq E\left(\bigcup_{i=1}^{M+Q+1} \Lambda_{n,i}, MN_0 + q 2^{nd} + Q_1, S\right) \\ &\leq M E\left(\Lambda_n, N_0, \frac{1}{m} [S - Q_0 \ln \mathfrak{m}(\Lambda_0)]\right) + \frac{Q_0 A}{2} \sum_{l=0}^{n-1} \frac{2^{l+1} d}{R_i^d} + \frac{AN^2}{2(r-1)^d L_n^d}, \end{aligned}$$

und man erhält eine Abschätzung der Energiedichte in Form von

$$\begin{aligned} \frac{E(\Lambda, N, S)}{\mathfrak{m}(\Lambda)} &\leq \frac{N}{\mathfrak{m}(\Lambda)} \left[ \frac{MN_0}{2^{nd} N (\xi + \delta)} E\left(\Lambda_n, N_0, \frac{1}{m} [S - Q_0 \ln \mathfrak{m}(\Lambda_0)]\right) \right. \\ &\quad \left. + \frac{Q_0 A}{2N} \sum_{l=0}^{n-1} \frac{2^{l+1} d}{R_i^d} + \frac{AN}{2(r-1)^d L_n^d} \right]. \end{aligned}$$

Für  $\Lambda \rightarrow \infty$  verschwinden der erste und der dritte Summand, während der Summenterm stehenbleibt, und man findet aufgrund der gleichmäßigen Konvergenz von  $\mathcal{E}$

$$\limsup_{\Lambda \rightarrow \infty} \frac{E(\Lambda, N, S)}{\mathfrak{m}(\Lambda)} \leq \frac{\xi}{L^d (\xi + \delta)} \mathcal{E}\left(\xi + \delta, \frac{\xi + \delta}{\xi} \eta\right).$$

Für  $\delta \rightarrow 0$  schließlich ergibt sich

$$\limsup_{\Lambda \rightarrow \infty} \frac{E(\Lambda, N, S)}{\mathfrak{m}(\Lambda)} \leq \frac{\mathcal{E}(\xi, \eta)}{L^d}.$$

Um das Infimum abzuschätzen, wählen wir  $C > 0$  und  $n \in \mathbb{N}$  so, daß der Würfel  $\Lambda_n$  eine parallelverschobene Kopie  $\Lambda'$  von  $\Lambda$  enthält und für deren Inhalte

$$C \leq \frac{\mathfrak{m}(\Lambda)}{\mathfrak{m}(\Lambda_n)} \leq \frac{1}{2}$$

gilt. Zusätzlich wählen wir ein  $\delta < 0$  so, daß  $\xi + \delta = p/2^q$  gilt mit geeigneten  $p, q \in \mathbb{N}$ . Weiter sei

$$\Lambda'' = \left\{ x \in \Lambda_n \mid \min\{|x - y| \mid y \in \Lambda'\} > \frac{\theta^n R}{2 - \theta} \right\}.$$

Gilt im Sinne der Voraussetzungen (i) – (iii) des Satzes  $\lim \Lambda = \infty$ , gilt auch  $\lim \Lambda'' = \infty$  sowie

$$\lim_{\Lambda \rightarrow \infty} \frac{\mathfrak{m}(\Lambda) + \mathfrak{m}(\Lambda'')}{\mathfrak{m}(\Lambda_n)} = 1.$$

Wiederum liefern die Monotonie von  $\mathcal{E}(\Lambda, N, S)$  und (A.1) eine Abschätzung für  $E$ , in diesem Fall

$$E(\Lambda_n, 2^{nd}(\xi + \delta), 2^{nd}\eta) \leq E(\Lambda, N, S) + E(\Lambda'', 2^{nd}(\xi + \delta) - N, 2^{nd}\eta - S) \\ + \frac{2^{2nd}(\xi + \delta)^2(2 - \theta)^\alpha A}{(\theta^N R)^\alpha}.$$

Im Grenzfall  $\Lambda \rightarrow \infty$  und damit  $n \rightarrow \infty$  folgt daraus für die Energiedichte

$$\liminf_{\Lambda \rightarrow \infty} \frac{E(\Lambda, N, S)}{\mathfrak{m}(\Lambda)} \geq \liminf_{\Lambda \rightarrow \infty} \left[ \frac{\mathfrak{m}(\Lambda_n)}{\mathfrak{m}(\Lambda)} \frac{E(\Lambda_n, (\xi + \delta)/2^{nd}, 2^{nd}\eta)}{L^d 2^{nd}} \right. \\ \left. - \frac{\mathfrak{m}(\Lambda_n) - \mathfrak{m}(\Lambda)}{\mathfrak{m}(\Lambda)} \frac{E(\Lambda'', 2^{nd}(\xi + \delta) - N, 2^{nd}\eta - S)}{\mathfrak{m}(\Lambda'')} \right],$$

und für  $\delta \rightarrow 0$  erhält man

$$\liminf_{\Lambda \rightarrow \infty} \frac{E(\Lambda, N, S)}{\mathfrak{m}(\Lambda)} \geq \frac{\mathcal{E}(\xi, \eta)}{L^d}.$$

Der Limes der Energiedichte existiert folglich für beliebige konvexe Gebiete und hat den Wert

$$\lim_{\Lambda \rightarrow \infty} \frac{E(\Lambda, N, S)}{\mathfrak{m}(\Lambda)} = \frac{\mathcal{E}(\xi, \eta)}{L^d}.$$

b) Angenommen, für  $(\xi, \eta) \notin \bar{\Lambda}$  gelte

$$\lim_{\Lambda \rightarrow \infty} \frac{E(\Lambda, N, S)}{\mathfrak{m}(\Lambda)} = \chi < \infty.$$

Es seien  $\Lambda_n, \Lambda'$  und  $\Lambda''$  definiert wie oben. Gemäß dem Satz von Bolzano-Weierstraß gibt es dann eine konvergente Teilfolge  $\{\Lambda_{n'}\}_{n' \in \mathfrak{N}}$  von  $\{\Lambda_n\}_{n \in \mathfrak{N}}$ ; deren Grenzwert werde mit  $b$  bezeichnet. Außerdem gelte

$$\left( \lim_{\Lambda'' \rightarrow \infty} \frac{N''}{\mathfrak{m}(\Lambda'')}, \lim_{\Lambda'' \rightarrow \infty} \frac{S''}{\mathfrak{m}(\Lambda'')} \right) = (\xi'', \eta'')$$

mit

$$(\xi'', \eta'') \in \mathcal{G}, \quad N + N'' = 2^{nd} \tilde{\xi}, \quad S + S'' = 2^{nd} \tilde{\eta}.$$

Analog zu oben liefert Anwendung von (A.1)

$$\chi \liminf_{\Lambda \rightarrow \infty} \frac{E(\Lambda, N, S)}{\mathfrak{m}(\Lambda)} + (1 - \chi) \frac{\mathcal{E}(\xi'', \eta'')}{L^d} \geq \frac{1}{L^d} \lim_{n \rightarrow \infty} \frac{E(\Lambda_n, 2^{nd} \tilde{\xi}, 2^{nd} \tilde{\eta})}{2^{nd}}. \quad (\text{A.9})$$

Wählt man nun  $(\tilde{\xi}, \tilde{\eta}) \notin \bar{\mathcal{G}}$ , so gilt

$$E(\Lambda_n, 2^{nd} \tilde{\xi}, 2^{nd} \tilde{\eta}) = +\infty,$$

ein Widerspruch.

c) Angenommen, für  $(\xi, \eta) \in \partial \mathcal{G}$  gelte

$$\liminf_{\Lambda \rightarrow \infty} \frac{E(\Lambda, N, S)}{\mathfrak{m}(\Lambda)} < \liminf_{(\xi, \eta) \rightarrow (\xi, \eta)} \mathcal{E}(\xi, \eta).$$

Die Bezeichnungen seien wie eben gewählt, zusätzlich seien  $\tilde{\xi} = \xi + t\xi$  und  $\tilde{\eta} = \eta + t\eta$  mit  $(\xi, \eta) \in \mathbb{R}^d$  und  $t \in \mathbb{R}$ . Für  $t \rightarrow 0$  folgt aus (A.9)

$$\liminf_{\Lambda \rightarrow \infty} \frac{E(\Lambda, N, S)}{\mathfrak{m}(\Lambda)} \geq \mathcal{E}(\xi, \eta),$$

ein Widerspruch. □

Anstelle  $E$  als Funktion von  $\Lambda, N$  und  $S$  ist es oft zweckmäßig, andere thermodynamische Variablen zu betrachten, etwa  $S$  als Funktion von  $\Lambda, N$  und  $E$ . Es ist in analoger Weise möglich, auch hierfür die Existenz thermodynamischer Limes zu beweisen. Details dazu findet man unter anderem in [439], [440] und [505].

### A.3 Die mikrokanonische Gesamtheit

Wir kommen nun zum zweiten Aspekt, den physikalischen Konsequenzen. Dabei wird sich gleichzeitig zeigen, daß die Existenz des thermodynamischen Limes für die mikrokanonische Gesamtheit dasselbe für andere Gesamtheiten mit beinhaltet.

Zunächst bleiben wir noch bei der mikrokanonischen Gesamtheit. Vermöge einer Umskalierung führen wir die physikalischen Variablen  $\rho = \xi/L^{nd}$ ,  $s = \eta/L^{nd}$ ,  $\varepsilon = \mathcal{E}/L^{nd}$  ein; das sind die thermodynamischen Grenzwerte der *Teilchendichte*, der *Entropiedichte* und der *Energiedichte*. Das Supremum von  $\rho$  sei  $\hat{\rho}$ , für dieses gilt  $\hat{\rho} \in [0, \infty]$ <sup>2</sup>. Das Infimum von  $\varepsilon$  für  $0 \leq \rho < \hat{\rho}$  sei  $\varepsilon_0(\rho)$ , eine konvexe, stetige Funktion auf  $[0, \hat{\rho})$  mit  $\varepsilon_0(0) = 0$  und  $\varepsilon_0(\rho) \geq -\rho B$ . Damit ist  $s$  eine konkave, stetige Funktion auf  $\Theta = \{(\rho, \varepsilon) \in \mathbb{R}^2 \mid 0 \leq \rho < \hat{\rho}, \varepsilon > \varepsilon_0(\rho)\}$ , die als Funktion von  $\varepsilon$  monoton steigt mit  $s(0, \varepsilon) = 0$  für  $\varepsilon > 0$ . Weiter sei

$$\begin{aligned} \Sigma &= \{(\rho, s, \varepsilon) \in \mathbb{R}^3 \mid (L^{nd}\rho, L^{nd}s) \in \mathcal{G}, \varepsilon = L^{nd}\mathcal{E}(\xi, \eta)\} \\ &\cup \{(\rho, s, \varepsilon) \in \mathbb{R}^3 \mid \rho = 0, s < 0, \varepsilon = 0\} \\ &\cup \bigcup \left\{ \{(\rho, s, \varepsilon) \in \mathbb{R}^3 \mid \rho = \tilde{\beta}/L^{nd}, s = \tilde{\eta}/L^{nd}, \varepsilon \geq \liminf_{(\xi, \eta) \rightarrow (\xi, \eta)} \mathcal{E}(\xi, \eta)\} \mid \right. \\ &\quad \left. (\tilde{\beta}, \tilde{\eta}) \in \partial \mathcal{G}, \xi \neq 0, \eta \geq 0 \right\}. \end{aligned}$$

Für

$$\lim_{\Lambda \rightarrow \infty} \frac{N}{\mathfrak{m}(\Lambda)} = \rho, \quad \lim_{\Lambda \rightarrow \infty} \frac{S(N, E)}{\mathfrak{m}(\Lambda)} = s, \quad \lim_{\Lambda \rightarrow \infty} \frac{E}{\mathfrak{m}(\Lambda)} = \varepsilon$$

folgt dann unmittelbar  $(\rho, s, \varepsilon) \in \Sigma$ , außerdem gilt

<sup>2</sup>Wenn die Teilchen des Systems punktförmig sind, gilt  $\hat{\rho} = \infty$ . Stellen sie undurchdringliche, inkompressible Systeme dar, so ist  $\hat{\rho}$  die Dichte der dichtesten Packung dieser Teilchen.



a) für  $(\rho, \varepsilon) \in \Theta$ :

$$\lim_{\Lambda \rightarrow \infty} \frac{S(\Lambda, N, E)}{\mathfrak{m}(\Lambda)} = s(\rho, \varepsilon),$$

b) für  $(\rho, \varepsilon) \in \partial \Theta$ :

$$\liminf_{\Lambda \rightarrow \infty} \frac{S(\Lambda, N, E)}{\mathfrak{m}(\Lambda)} \leq \limsup_{(\xi, \eta) \rightarrow (\rho, \varepsilon)} s(\xi, \eta), \quad (\text{A.10})$$

c) und für  $(\rho, \varepsilon) \in \mathbb{F}^2 \setminus \Theta$ :

$$\lim_{\Lambda \rightarrow \infty} \frac{S(\Lambda, N, E)}{\mathfrak{m}(\Lambda)} = -\infty.$$

Wir können die bisher beschriebenen Sachverhalte nun auf die mikrokanonische Zustandssumme anwenden. Diese schreiben wir in der Form

$$\begin{aligned} \Omega(\Lambda, N, E) &= \int_{\mathbb{Z}^d} \int_{\mathbb{Z}^d} \cdots \int_{\mathbb{Z}^d} \delta \left( \sum_{i=1}^N \frac{p_i^2}{2m} - a \right) dp_1 dp_2 \cdots dp_N da \times \\ &\quad \times \frac{1}{N!} \int_{\Lambda} \int_{\Lambda} \cdots \int_{\Lambda} \chi_{[-\infty, 0]}(V(x_1, x_2, \dots, x_N) + a - E) dx_1 dx_2 \cdots dx_N \\ &= \int \Omega_{\text{kin}}(N, a) \Omega_{\text{conf}}(\Lambda, N, E - a) da \end{aligned}$$

mit der *Konfigurations-Zustandssumme*

$$\Omega_{\text{conf}}(\Lambda, N, E) = \frac{1}{N!} \int_{\Lambda} \int_{\Lambda} \cdots \int_{\Lambda} \chi_{[-\infty, 0]}(V(x_1, x_2, \dots, x_N) - E) dx_1 dx_2 \cdots dx_N$$

und der *kinetischen Zustandssumme*

$$\Omega_{\text{kin}}(N, E) = \int_{\mathbb{Z}^d} \int_{\mathbb{Z}^d} \cdots \int_{\mathbb{Z}^d} \delta \left( \sum_{i=1}^N \frac{p_i^2}{2m} - E \right) dp_1 dp_2 \cdots dp_N.$$

Der Wert dieses Ausdrucks entspricht dem Inhalt der Oberfläche  $O_N$  einer  $N$ -dimensionalen Kugel mit Radius  $\sqrt{2m/E}$ . Um diese zu berechnen, betrachten wir das Integral

$$\begin{aligned} \mathcal{I}_N &= \int_{\mathbb{Z}^N} e^{-\|x\|^2} dx = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp \left( - \sum_{i=1}^N x_i^2 \right) dx_1 dx_2 \cdots dx_N \\ &= \left( \int_{-\infty}^{\infty} e^{-\xi^2} d\xi \right)^N = \pi^{N/2}, \end{aligned}$$

das wir auf die Oberfläche  $\omega_N$  der  $N$ -dimensionalen Einheitskugel zurückführen. Einführung von Polarkoordinaten sowie die Koordinatentransformation  $\chi = \xi^2$  liefern

$$\mathcal{I}_N = \int_0^{\infty} \int_{\omega_N} e^{-\xi^2} \xi^{N-1} d\omega_N d\xi = \omega_N \int_0^{\infty} e^{-\xi^2} \xi^{N-1} d\xi$$

$$= \frac{\omega_N}{2} \int_0^\infty e^{-\chi} \chi^{N/2-1} d\chi = \frac{\omega_N}{2} \Gamma(N/2),$$

mit der Gamma-Funktion<sup>3</sup>

$$\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt;$$

daraus folgt für die Oberfläche der  $N$ -dimensionale Einheitskugel

$$\omega_N = \frac{2\pi^{N/2}}{\Gamma(N/2)}$$

und für die Oberfläche einer  $N$ -dimensionalen Kugel mit Radius  $R$

$$O_N = \frac{2\pi^{N/2} R^N}{\Gamma(N/2)}.$$

Für die kinetische Zustandssumme ergibt sich daraus

$$\Omega_{\text{kin}}(N, E) = \frac{(2\pi m E)^{Nd/2}}{E \Gamma(Nd/2)}.$$

Hiermit definieren wir die kinetische Entropiedichte

$$s_{\text{kin}}(\rho, \varepsilon) = \lim_{\Lambda \rightarrow \infty} \frac{\ln \Omega_{\text{kin}}(N, E)}{\mathfrak{m}(\Lambda)} = \frac{d\rho}{2} \left( \ln \frac{4\pi m \varepsilon}{d\rho} + 1 \right).$$

und damit wiederum die gesamte Entropiedichte

$$\mathfrak{s}(\rho, \varepsilon) = \sup_{\varepsilon' \in (0, \varepsilon - \varepsilon_0(\rho))} [s_{\text{kin}}(\rho, \varepsilon') + s(\rho, \varepsilon - \varepsilon')].$$

Wie alle thermodynamisch relevanten Größen läßt sich die Entropie aus der Zustandssumme berechnen. Es gilt der folgende

**Satz:** Für  $(\rho, \varepsilon) \in \Theta$  sei

$$\lim_{\Lambda \rightarrow \infty} \frac{N}{\mathfrak{m}(\Lambda)} = \rho, \quad \lim_{\Lambda \rightarrow \infty} \frac{E}{\mathfrak{m}(\Lambda)} = \varepsilon.$$

Dann gilt für die mikrokanonische Gesamtheit

$$\lim_{\Lambda \rightarrow \infty} \frac{\ln \Omega(\Lambda, N, E)}{\mathfrak{m}(\Lambda)} = \mathfrak{s}(\rho, \varepsilon).$$

<sup>3</sup>Kennt man sich mit der Gamma-Funktion aus, erhält man das in Abschnitt 2.2.2 gewonnene und hier ebenfalls verwendete Resultat

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$$

unmittelbar über

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \Gamma(-1/2) = \sqrt{\pi}.$$

BEWEIS: Einerseits folgt aus (A.10) für genügend große  $\Lambda$

$$\Omega_{\text{kin}}(N, E) > \exp[\mathbf{m}(\Lambda) s_{\text{kin}}(\rho, \varepsilon)]$$

und damit in einem geeigneten Intervall  $(a_1, a_2)$  für alle  $a \in (a_1, a_2)$

$$\Omega_{\text{kin}}(N, E) \Omega_{\text{conf}}(\Lambda, N, E) > \exp\{\mathbf{m}(\Lambda) [\mathfrak{s}_{\text{kin}}(\rho, \varepsilon) - \xi]\};$$

das wiederum zieht

$$\lim_{\Lambda \rightarrow \infty} \frac{\ln \Omega(\Lambda, N, E)}{\mathbf{m}(\Lambda)} \geq \mathfrak{s}(\rho, \varepsilon) \quad (\text{A.11})$$

nach sich. Andererseits verschwindet  $\Omega_{\text{kin}}(N, a) \Omega_{\text{conf}}(\Lambda, N, E - a)$  außerhalb des abgeschlossenen Intervalls  $[0, a]$ , wenn  $a$  genügend groß ist; das folgt aus der Definition dieses Ausdrucks sowie aus der vorausgesetzten Stabilität der zugehörigen Potentialfolge. Zusätzlich gilt auf  $[0, a]$

$$\Omega_{\text{kin}}(N, E) \Omega_{\text{conf}}(\Lambda, N, E) < \exp\{\mathbf{m}(\Lambda) [\mathfrak{s}_{\text{kin}}(\rho, \varepsilon) + \xi]\}.$$

denn andernfalls gäbe es eine Folge  $\{(\Lambda_n, N_n, E_n)\}_{n \in \mathfrak{N}}$  mit

$$\lim_{\Lambda_n \rightarrow \infty} \frac{N_n}{\mathbf{m}(\Lambda_n)} = \rho, \quad \lim_{\Lambda_n \rightarrow \infty} \frac{E_n}{\mathbf{m}(\Lambda_n)} = \varepsilon$$

und ein  $C > 0$ , so daß

$$\left| \frac{S(\Lambda, N, E)}{\mathbf{m}(\Lambda)} - s(\rho, \varepsilon) \right| \geq C,$$

ein Widerspruch zu Aussage a) oben. Daraus folgt

$$\lim_{\Lambda \rightarrow \infty} \frac{\ln \Omega(\Lambda, N, E)}{\mathbf{m}(\Lambda)} \leq \mathfrak{s}(\rho, \varepsilon), \quad (\text{A.12})$$

und (A.11) und (A.12) liefern gemeinsam die Behauptung.  $\square$

Analog zeigt man für  $(\rho, \varepsilon) \in \partial \Theta$

$$\liminf_{\Lambda \rightarrow \infty} \frac{\ln \Omega(\Lambda, N, E)}{\mathbf{m}(\Lambda)} \leq \limsup_{(\xi, \eta) \rightarrow (\rho, \varepsilon)} \mathfrak{s}(\xi, \eta)$$

und für  $(\rho, \varepsilon) \in \mathfrak{R}^2 \setminus \Theta$

$$\lim_{\Lambda \rightarrow \infty} \frac{\ln \Omega(\Lambda, N, E)}{\mathbf{m}(\Lambda)} = -\infty.$$

## A.4 Die kanonische Gesamtheit

Der Übergang von der mikrokanonischen zur kanonischen Gesamtheit geschieht formal, indem man die Energie  $E$  durch einen Parameter  $\beta > 0$  ersetzt, der proportional zum Kehrwert der absoluten Temperatur  $T$  ist. Wir definieren die *kanonische Zustandssumme*

$$Z = \left( \frac{2\pi m}{\beta} \right)^{Nd/2} \frac{1}{N!} \int_{\Lambda} \int_{\Lambda} \cdots \int_{\Lambda} \exp[\beta V(x_1, x_2, \dots, x_N)] dx_1 dx_2 \cdots dx_N$$

sowie für  $0 \leq \rho < \hat{\rho}$  und  $\beta > 0$  die *spezifische freie Energie*

$$f(\rho, \beta) = \inf_{\varepsilon > \varepsilon_0(\rho)} \left\{ \varepsilon - \frac{s(\rho, \varepsilon)}{\beta} \right\}.$$

$f$  ist eine konvexe Funktion von  $\rho$  und eine konkave Funktion von  $1/\beta$ . Die folgende zentrale Aussage liefert gleichzeitig den Zusammenhang zwischen den Variablen der mikrokanonischen und der kanonischen Gesamtheit.

**Satz:** Für  $0 \leq \rho < \hat{\rho}$  sei

$$\lim_{\Lambda \rightarrow \infty} \frac{N}{\mathbf{m}(\Lambda)} = \rho.$$

Dann gilt für die kanonische Gesamtheit

$$\lim_{\Lambda \rightarrow \infty} \frac{\ln Z(\Lambda, N, \beta)}{\mathbf{m}(\Lambda)} = -\frac{(2\pi m)^{Nd/2}}{\beta^{Nd/2-1}} f(\rho, \beta).$$

BEWEIS: Es sei  $\delta > 0$ . Dann gibt es ein  $\varepsilon'$  mit

$$\varepsilon' < f(\rho, \beta) + \frac{s(\rho, \varepsilon') + \delta}{\beta}.$$

Für genügend großes  $\Lambda$  können wir die Zustandssumme abschätzen, erhalten

$$\begin{aligned} Z(\Lambda, N, \beta) &\geq \left( \frac{2\pi m}{\beta} \right)^{Nd/2} \frac{1}{N!} \int_{\Lambda} \int_{\Lambda} \cdots \int_{\Lambda} \chi_{[-\infty, 0]}(V(x_1, x_2, \dots, x_N) - \varepsilon' \mathbf{m}(\Lambda)) \times \\ &\quad \times \exp[\beta V(x_1, x_2, \dots, x_N)] dx_1 dx_2 \cdots dx_N \\ &\geq \left( \frac{2\pi m}{\beta} \right)^{Nd/2} \exp[-\beta \varepsilon' \mathbf{m}(\Lambda)] \Omega(\Lambda, N, \varepsilon' \mathbf{m}(\Lambda)) \\ &\geq \left( \frac{2\pi m}{\beta} \right)^{Nd/2} \exp\{\mathbf{m}(\Lambda) [-\beta \varepsilon' + s(\beta, \varepsilon') - \delta]\} \\ &\geq \left( \frac{2\pi m}{\beta} \right)^{Nd/2} \exp\{\mathbf{m}(\Lambda) [-\beta f(\rho, \beta) - 2\delta]\} \end{aligned}$$

und daraus sofort

$$\lim_{\Lambda \rightarrow \infty} \frac{\ln Z(\Lambda, N, \beta)}{\mathbf{m}(\Lambda)} \geq -\frac{(2\pi m)^{Nd/2}}{\beta^{Nd/2-1}} f(\rho, \beta). \quad (\text{A.13})$$

Nun wählen wir  $-\rho B < \varepsilon_1 < \varepsilon_2 < \cdots < \varepsilon_{N+1}$  so, daß

$$\varepsilon_i < f(\rho, \beta) + \frac{s(\rho, \varepsilon_{i+1}) + \delta}{\beta}$$

gilt für  $i = 1, 2, \dots, N$ , setzen

$$Z_i = \left( \frac{2\pi m}{\beta} \right)^{Nd/2} \frac{1}{N!} \int_{\Lambda} \int_{\Lambda} \cdots \int_{\Lambda} \chi_{[\mathbf{m}(\Lambda) \varepsilon_i, \mathbf{m}(\Lambda) \varepsilon_{i+1}]}(V(x_1, x_2, \dots, x_N) - \varepsilon_i \mathbf{m}(\Lambda)) \times$$

$$\times \exp [\beta V(x_1, x_2, \dots, x_N)] dx_1 dx_2 \cdots dx_N$$

und schätzen damit die kanonische Zustandssumme ab gemäß

$$Z \geq \sum_{i=1}^N Z_i.$$

Die Abweichung können wir ebenfalls abschätzen und erhalten

$$Z - \sum_{i=1}^N Z_i \leq \frac{[\mathbf{m}(\Lambda)]^N}{N!} \exp(-\beta \mathbf{m}(\Lambda) \varepsilon_{N+1});$$

dieser Term ist für genügend große  $\varepsilon_{N+1}$  offensichtlich vernachlässigbar. Daraus folgt

$$\lim_{\Lambda \rightarrow \infty} \frac{\ln Z(\Lambda, N, \beta)}{\mathbf{m}(\Lambda)} \leq -\frac{(2\pi m)^{Nd/2}}{\beta^{Nd/2-1}} f(\rho, \beta) \quad (\text{A.14})$$

und mit (A.13) und (A.14) zusammen die Behauptung.  $\square$

Wieder zeigt man analog für  $\rho = \hat{\rho}$

$$\lim_{\Lambda \rightarrow \infty} \frac{\ln Z(\Lambda, N, \beta)}{\mathbf{m}(\Lambda)} \leq -\frac{(2\pi m)^{Nd/2}}{\beta^{Nd/2-1}} \lim_{\substack{\rho \rightarrow \hat{\rho} \\ \rho < \hat{\rho}}} f(\rho, \beta)$$

und für  $\rho > \hat{\rho}$

$$\lim_{\Lambda \rightarrow \infty} \frac{\ln Z(\Lambda, N, \beta)}{\mathbf{m}(\Lambda)} = -\infty.$$

Da Funktionen, die auf einem Intervall konvex sind, dort auch fast überall differenzierbar sind, kann man nun die Gleichungen

$$\beta = \frac{\partial s(\rho, \varepsilon')}{\partial \varepsilon'}, \quad s(\rho, \varepsilon') = \beta^2 \frac{\partial f(\rho, \beta)}{\partial \beta}$$

direkt ablesen, womit wir die üblichen thermodynamischen Relationen der kanonischen Gesamtheit wiedergefunden haben.

## A.5 Die großkanonische Gesamtheit

Während bei der mikrokanonischen und der kanonischen Gesamtheit die Anzahl  $N$  der Teilchen als Konstante betrachtet wird, ist sie bei der großkanonischen Gesamtheit variabel. Sie wird daher durch eine neue Größe ersetzt, das *chemische Potential*  $\mu$ . Wir definieren die *großkanonische Zustandssumme*

$$Y(\Lambda, \mu, \beta) = \sum_{N=0}^{\infty} \frac{e^{N\beta\mu}}{N!} \left( \frac{2\pi m}{\beta} \right)^{Nd/2} \int_{\Lambda} \int_{\Lambda} \cdots \int_{\Lambda} \exp[\beta V(x_1, x_2, \dots, x_N)] dx_1 dx_2 \cdots dx_N$$

sowie für  $0 \leq \rho < \hat{\rho}$  und  $\beta > 0$  den *Druck*

$$p(\mu, \beta) = \sup_{0 \leq \rho < \hat{\rho}} \left\{ \rho \left[ \mu + \frac{d}{2\beta} \ln \frac{2\pi m}{\beta} \right] - f(\rho, \beta) \right\};$$

$\beta p(\mu, \beta)$  ist eine konvexe Funktion von  $\mu$  und  $\beta$ . Für die großkanonische Zustandssumme gilt  $Y \geq 1$ , folglich ist  $p \geq 0$ , und da  $Y$  eine monoton steigende Funktion von  $\mu$  ist, gilt das auch für  $p$ . Die zentrale Aussage lautet nun wie folgt.

**Satz:** Für die großkanonische Gesamtheit gilt

$$\lim_{\Lambda \rightarrow \infty} \frac{\ln Y(\Lambda, \mu, \beta)}{\mathbf{m}(\Lambda)} = \beta p(\mu, \beta).$$

BEWEIS: Es sei  $\delta > 0$ . Dann gibt es ein  $\rho'$  mit  $0 \leq \rho' < \hat{\rho}$  mit

$$\rho' \left[ \beta\mu + \frac{d}{2} \ln \frac{2\pi m}{\beta} \right] \geq \beta [p(\mu, \beta) + f(\rho', \beta)] - \delta.$$

Wählen wir  $N'$  so, daß

$$\lim_{\Lambda \rightarrow \infty} \frac{N'}{\mathbf{m}(\Lambda)} = \rho'$$

gilt, so folgt für genügend großes  $\Lambda$

$$Y(\Lambda, \mu, \beta) \geq \frac{e^{N'\beta\mu}}{N'!} \left( \frac{2\pi m}{\beta} \right)^{N'd/2} \int_{\Lambda} \int_{\Lambda} \cdots \int_{\Lambda} \exp[\beta V(x_1, x_2, \dots, x'_{N'})] dx_1 dx_2 \cdots dx'_{N'}$$

und

$$\exp \left\{ \mathbf{m}(\Lambda) \left[ \rho' \left( \beta\mu + \frac{d}{2} \ln \frac{2\pi m}{\beta} \right) - \beta f(\rho', \beta) - \delta \right] \right\} \geq \exp \{ \mathbf{m}(\Lambda) [\beta p(\mu, \beta) - 2\delta] \};$$

daraus erhält man

$$\lim_{\Lambda \rightarrow \infty} \frac{\ln Y(\Lambda, \mu, \beta)}{\mathbf{m}(\Lambda)} \geq \beta p(\mu, \beta). \quad (\text{A.15})$$

Aus der Stabilitätsbedingung für das Potential folgt außerdem

$$\begin{aligned} & \frac{e^{N\beta\mu}}{N!} \left( \frac{2\pi m}{\beta} \right)^{Nd/2} \int_{\Lambda} \int_{\Lambda} \cdots \int_{\Lambda} \exp[\beta V(x_1, x_2, \dots, x_N)] dx_1 dx_2 \cdots dx_N \\ & \leq \frac{1}{N!} \left[ \left( \frac{2\pi m}{\beta} \right)^{d/2} \mathbf{m}(\Lambda) e^{\beta(\mu+B)} \right]^N \end{aligned}$$

und damit auch

$$\begin{aligned} & \frac{e^{N\beta\mu}}{N!} \left( \frac{2\pi m}{\beta} \right)^{Nd/2} \int_{\Lambda} \int_{\Lambda} \cdots \int_{\Lambda} \exp[\beta V(x_1, x_2, \dots, x_N)] dx_1 dx_2 \cdots dx_N \\ & \leq \left[ \left( \frac{2\pi m}{\beta} \right)^{d/2} \frac{\mathbf{m}(\Lambda)}{N} e^{\beta(\mu+B)+1} \right]^N \equiv \left( \frac{\mathcal{A}}{N} \right)^N. \end{aligned}$$

Summiert man das auf, so erhält man die Abschätzung

$$\begin{aligned} & \sum_{N>2\mathcal{A}} \frac{e^{N\beta\mu}}{N!} \left( \frac{2\pi m}{\beta} \right)^{Nd/2} \int_{\Lambda} \int_{\Lambda} \cdots \int_{\Lambda} \exp[\beta V(x_1, x_2, \dots, x_N)] dx_1 dx_2 \cdots dx_N \\ & \leq \sum_{N>2\mathcal{A}} \left( \frac{\mathcal{A}}{N} \right)^N \leq \sum_{N>2\mathcal{A}} \left( \frac{1}{2} \right)^N. \end{aligned}$$

Wählt man ein  $\delta > 0$ , so gilt gleichzeitig im abgeschlossenen Intervall  $[0, \mathcal{A}/2 \mathbf{m}(\Lambda)]$  für genügend großes  $\Lambda$

$$\begin{aligned} & \frac{e^{N\beta\mu}}{N!} \left( \frac{2\pi m}{\beta} \right)^{Nd/2} \int_{\Lambda} \int_{\Lambda} \cdots \int_{\Lambda} \exp[\beta V(x_1, x_2, \dots, x_N)] dx_1 dx_2 \cdots dx_N \\ & \leq \exp \left\{ \mathbf{m}(\Lambda) \left[ \rho \left( \beta\mu + \frac{d}{2} \ln \frac{2\pi m}{\beta} \right) - \beta f \left( \frac{N}{\mathbf{m}(\Lambda)}, \beta \right) + \delta \right] \right\} \\ & \leq \exp \{ \mathbf{m}(\Lambda) [\beta p(\mu, \beta) + \delta] \}. \end{aligned}$$

Das liefert

$$Y(\Lambda, \mu, \beta) \leq \frac{\mathcal{A}}{2} \exp \{ \mathbf{m}(\Lambda) [\beta p(\mu, \beta) + \delta] \} + \sum_{N>\mathcal{A}/2} \left( \frac{1}{2} \right)^N$$

und weiter

$$\lim_{\Lambda \rightarrow \infty} \frac{\ln Y(\Lambda, \mu, \beta)}{\mathbf{m}(\Lambda)} \leq \beta p(\mu, \beta). \quad (\text{A.16})$$

Aus (A.15) und (A.16) folgt die Behauptung.  $\square$



Wieder folgt aus der Konvexität von  $p$  Differenzierbarkeit fast überall, und man liest die Gleichungen

$$\mu + \frac{d}{2\beta} \ln \frac{2\pi m}{\beta} = \frac{\partial f(\rho', \beta)}{\partial \rho'}, \quad \rho' = \frac{\partial p(\mu, \beta)}{\partial \mu}$$

ab, welche die thermodynamischen Relationen für die großkanonische Gesamtheit darstellen.

# Anhang B

## Der Fock-Raum

Da Fock-Räume in Abschnitt 2.6 im Zusammenhang mit weißem Rauschen an zentraler Stelle aufgetaucht sind und zudem in der Quantenfeldtheorie von großer Bedeutung sind, lohnt es sich, einen formal präzisen Blick darauf zu werfen<sup>1</sup>. Während dabei das mathematische Konzept in der der stochastischen Analysis und der Quantenfeldtheorie identisch ist, unterscheiden sich die physikalischen Interpretationen beträchtlich, denn im ersten Fall ist der Hilbertraum  $\mathcal{H}$  in  $\Gamma(\mathcal{H})$  ein Hilbertraum über  $\mathbb{R}$ , was hier die Bedeutung der Zeitachse hat, wogegen im zweiten Fall das  $\mathbb{R}^4$  in  $\Gamma(\mathcal{L}^2(\mathbb{R}^4))$  für den Konfigurationsraum steht, auf dem sich die betrachtete Quantenfeldtheorie abspielt.

### B.1 Definition und mathematische Eigenschaften

Im folgenden sei  $\mathcal{H}$  ein separabler komplexer Hilbertraum. Aus mathematischer Sicht handelt es sich bei einem Fock-Raum um die direkte Summe

$$\mathcal{F}(\mathcal{H}) = \bigoplus_{n=0}^{\infty} \mathcal{H}^{\otimes n}. \quad (\text{B.1})$$

Seine Elemente sind Folgen der Form  $\{\varphi^{(n)}\}_{n \in \mathbb{N}}$  mit  $\varphi^{(n)} \in \mathcal{H}^{\otimes n}$ . Ist  $\mathcal{H}$  ein reeller Hilbertraum, ersetzt man in der direkten Summe jeweils  $\mathcal{H}^{\otimes n}$  durch  $\mathcal{H}_r^{\otimes n}$ ; ansonsten läuft alles weitere analog.

Mit dem Produkt

$$\varphi \otimes \psi := \sum_{n=0}^{\infty} \varphi^{(n)} \otimes \psi^{(n)} \quad \text{für } \varphi = \{\varphi^{(n)}\}_{n \in \mathbb{N}}, \quad \psi = \{\psi^{(n)}\}_{n \in \mathbb{N}}$$

kann man  $\mathcal{F}(\mathcal{H})$  als Algebra betrachten und nennt das dann die *Tensoralgebra* über  $\mathcal{H}$ ; hierfür schreibt man anstelle von (B.1) üblicherweise

$$T(\mathcal{H}) = \bigoplus_{n=0}^{\infty} T^n(\mathcal{H}).$$

---

<sup>1</sup>Der Begriff des Fockraums wurde von V. Fock 1932 heuristisch eingeführt [174] und erstmals 1953 von J. Cook formal streng definiert [109]. Eine sehr anschauliche physikalische, allerdings nicht mathematisch rigorose Darstellung des Fock-Raum-Konzepts findet man in [202]; vergleiche auch [269]. Informationen aus mathematischer Sicht liefern beispielsweise [242], [318], [398] oder [475].

Aus physikalischer Sicht beschreibt der Fock-Raum Zustände mit variabler Teilchenzahl; dabei können die Elemente der einzelnen Summanden  $\mathcal{H}^{\otimes n}$  als  $n$ -Teilchen-Zustände aufgefaßt werden, weswegen  $\mathcal{H}^{\otimes 0} \equiv \mathfrak{U}^1$  als *Vakuum-Unterraum* und die  $\mathcal{H}^{\otimes n}$  als  *$n$ -Teilchen-Unterräume* bezeichnet werden. Somit können geeignete Übergänge zwischen unterschiedlichen Elementen von  $\Gamma(\mathcal{H})$  als Veränderungen der Teilchenzahl aufgefaßt werden.

Die Beschreibung eines physikalischen Systems durch Elemente eines Fock-Raumes (sogenannte Fock-Zustände) wird daher auch *Besetzungszahl-Darstellung* genannt. Da in der Hochenergiephysik typischerweise Teilchen vernichtet und erzeugt werden, ist der Fock-Raum in der Quantenfeldtheorie von großer Bedeutung. Die in Abschnitt 2.6.3 beschriebenen symmetrischen Fockräume

$$\Gamma(\mathcal{H}) := \bigoplus_{n=0}^{\infty} \mathcal{H}^{\widehat{\otimes} n},$$

kommen dabei zur Beschreibung von Bosonen zum Einsatz; zur Konstruktion antisymmetrischer Fockräume zur Beschreibung von Fermionen geht man etwas anders vor. Zunächst definiert man für alle  $k \in \mathbb{N}$

$$J^k(\mathcal{H}) := \{ f_1 \otimes f_2 \otimes \dots \otimes f_k \mid \text{Es gibt } 1 \leq i, j \leq k, \text{ so daß } f_i = f_j \} \subset \mathcal{H}^{\otimes k}$$

und bildet daraus die direkte Summe

$$J(\mathcal{H}) := \bigoplus_{n=0}^{\infty} J^n(\mathcal{H}).$$

Den fermionischen Fockraum erhält man aus dem allgemeinen Fockraum durch die Quotientenbildung

$$\Lambda(\mathcal{H}) := \mathcal{F}(\mathcal{H})/J(\mathcal{H}),$$

wodurch erreicht wird, daß jeder Zustand, in dem ein Niveau doppelt besetzt ist, verschwindet; das ist natürlich nichts anderes als das Pauli-Prinzip. Betrachtet man  $\Lambda(\mathcal{H})$  als Vektorraum, so gilt

$$\Lambda(\mathcal{H}) \cong \bigoplus_{n=0}^{\infty} \Lambda^n \equiv \bigoplus_{n=0}^{\infty} \mathcal{H}^{\otimes n} \Big/ \bigoplus_{n=0}^{\infty} J^n(\mathcal{H}).$$

Auch hierfür läßt sich ein Produkt definieren. Dazu verwendet man das durch

$$a^{(n)} \wedge b^{(m)} := \frac{n! + m!}{n!m!} a^{(n)} \widehat{\otimes} b^{(m)}, \quad a^{(n)} \in \Lambda^n(\mathcal{H}), \quad b^{(m)} \in \Lambda^m(\mathcal{H})$$

definierte *Dachprodukt*; mit diesem definiert man durch

$$a \wedge b := \sum_{n=0}^{\infty} a^{(n)} \wedge b^{(n)} \quad \text{für } a = \{a^{(n)}\}_{n \in \mathbb{N}}, \quad b = \{b^{(n)}\}_{n \in \mathbb{N}}$$

das Dachprodukt auf  $\Lambda(\mathcal{H})$ . Betrachtet man hiermit  $\Lambda(\mathcal{H})$  als Algebra, so nennt man sie auch *Äußere Algebra* oder *Grassmann-Algebra*<sup>2</sup>.

<sup>2</sup>Benannt nach H. Grassman, der im 19. Jahrhundert die Grundlagen hierfür lieferte [196], [197]. Siehe dazu auch [237] und [536].

## B.2 Basissysteme für Fock-Räume

Orthonormierte Schauderbasisen für Fockräume erhält man folgendermaßen: Man wählt zunächst eine Orthonormalbasis  $\{e_j\}_{j \in \mathfrak{N}}$  von  $\mathcal{H}$  und erhält mit

$$B_n := \left\{ \bigotimes_{i=1}^n e_{k_i} \mid i, k_i \in \{1, 2, \dots, n\} \right\}$$

eine Orthonormalbasis für  $\mathcal{H}^{\otimes n}$ . Trivialerweise gilt dabei  $B_0 = \{1\}$ . Eine Orthonormalbasis für  $\mathcal{F}(\mathcal{H})$  konstruieren wir dann ganz einfach in Gestalt des Systems

$$\mathcal{B} = \{(0, 0, \dots, b_n, 0, \dots) \mid b_n \in B_n, n \in \mathfrak{N}\}, \quad (\text{B.2})$$

aller Folgen, die jeweils nur an der  $(n+1)$ -ten Stelle einen nichtverschwindenden Eintrag in Form eines Elementes von  $B_n$  aufweisen, wobei  $n$  ganz  $\mathfrak{N}$  durchläuft. Entsprechend verfahren wir bei symmetrischen Fockräumen.

$$\widehat{B}_n := \left\{ \sqrt{\frac{n!}{n_0! n_1! \dots n_k!}} e_0^{\otimes n_0} \widehat{\otimes} e_1^{\otimes n_1} \widehat{\otimes} \dots \widehat{\otimes} e_k^{\otimes n_k} \mid n_0 + n_1 + \dots + n_k = n \right\}$$

ist eine Orthonormalbasis für  $\mathcal{H}^{\widehat{\otimes} n}$ , wieder gilt  $\widehat{B}_0 = \{1\}$ , und wir erhalten mit dem analog zu (B.2) aufgebauten System

$$\widehat{\mathcal{B}} = \{(0, 0, \dots, \widehat{b}_n, 0, \dots) \mid \widehat{b}_n \in \widehat{B}_n, n \in \mathfrak{N}\}$$

eine Orthonormalbasis für  $\Gamma(\widehat{H})$ .

Die Basis-Elemente des Fockraums, die sogenannten *Fock-Zustände*, können als Zustände mit *definierter Teilchenzahl* aufgefaßt werden<sup>3</sup>; das Gegenteil hierzu bilden die *kohärenten Zustände*<sup>4</sup>. Jeder Vektor  $f \in \mathcal{H}$  definiert einen kohärenten Zustand

$$\phi_f := \left( 1, f, \frac{f^{\otimes 2}}{2!}, \dots, \frac{f^{\otimes n}}{n!}, \dots \right) \in \Gamma(\mathcal{H}).$$

Kohärente Zustände sind Zustände mit *völlig unbestimmter Teilchenzahl*<sup>5</sup>. Ist  $\mathcal{B} \subset \mathcal{H}$  ein vollständiges System linear unabhängiger Vektoren, so ist auch das System  $\mathcal{K}_{\mathcal{B}} = \{\phi_f \mid f \in \mathcal{B}\}$  der zugehörigen kohärenten Zustände linear unabhängig und sogar übervollständig, das heißt, bereits eine echte Teilmenge von  $\mathcal{K}_{\mathcal{B}}$ , bei geeigneter Wahl

<sup>3</sup>Beispiele für Systeme, die durch Fock-Zustände beschrieben werden können, sind wie erwähnt Teilchenzustände in der Elementarteilchenphysik, aber auch Systeme aus einzelnen Photonen. Man spricht dabei auch von nichtklassischem Licht.

<sup>4</sup>Der Begriff der kohärenten Zustände wurde 1963 von R. J. Glauber eingeführt [191]. Ausführliche Informationen dazu stehen in [406].

<sup>5</sup>Kohärente Zustände nennt man auch *quasiklassische Zustände*, da Systeme, die durch sie beschrieben werden, in ihrem Verhalten klassischen Systemen so nah wie physikalisch möglich kommen. Dieses quasiklassische Verhalten gilt jedoch nur für Erwartungswerte; die Quantenobjekte selbst bleiben natürlich beliebig nichtklassisch. Man erhält beispielsweise im Rahmen der Quantisierung des elektromagnetischen Feldes mit kohärenten Zuständen ein sehr tragfähiges Modell für klassisches Licht, wie es etwa bei einem Laserstrahl vorkommt.

sogar eine abzählbare Teilmenge, spannt den kompletten Raum  $\Gamma(\mathcal{H})$  auf<sup>6</sup>. Solche echte Teilmengen bilden damit ebenfalls Schauderbasen für  $\Gamma(\mathcal{H})$ , allerdings keine orthonormierten, denn für beliebige  $f_1, f_2, f \in \mathcal{H}$  gilt

$$(\phi_{f_1}, \phi_{f_2})_{\Gamma(\mathcal{H})} = e^{\langle f_1, f_2 \rangle}$$

und

$$\|\phi_f\|_{\Gamma(\mathcal{H})} = e^{\|f\|^2/2}.$$

Zusätzliche Bedeutung kommt den kohärenten Zuständen im Zusammenhang mit dem in Abschnitt 2.6.3 mehrfach aufgetretenen Wiener-Itô-Segal Isomorphismus  $\Gamma(\mathcal{E}) \cong \mathcal{L}^2(\mathcal{E}, \mathcal{B}(\mathcal{E}), \mu)$  zu, denn dieser ist eindeutig bestimmt durch die für alle  $f \in \mathcal{E}$  gegebene Zuordnung

$$\mathfrak{g}_f \longleftrightarrow \phi_f,$$

mit

$$\mathfrak{g}_f(x) = e^{\langle x, f \rangle - \|f\|^2/2} \in \mathcal{L}^2(\mathcal{E}, \mathcal{B}(\mathcal{E}), \mu)$$

und

$$\phi_f = \left( 1, f, \frac{f^{\otimes 2}}{2!}, \dots, \frac{f^{\otimes n}}{n!}, \dots \right) \in \Gamma(\mathcal{E}).$$

---

<sup>6</sup>Auch hierzu finden sich die Details in [406].

# Anhang C

## Lösungen der Aufgaben

### Kapitel 3

#### Abschnitt 3.3

1a)  $M = 1200, N_1 = 39 \implies N_1^M = 39^{1200} \approx (10^{1,5911})^{1200} \approx 10^{1909}$

b)  $M = 1200, N_2 = 40 \implies N_2^M = 40^{1200} \approx (10^{1,6021})^{1200} \approx 10^{1923}$

Die Anzahl der Möglichkeiten steigt um das  $10^{23}$ -fache.

2a) Der Stein könnte sich oder die Umgebung oder beides abkühlen, die dabei freiwerdende Energie in Bewegungsenergie und diese wiederum in potentielle Energie im Gravitationsfeld umwandeln.

b) Bei diesem Vorgang müßten sich ein oder mehrere Systeme spontan abkühlen, das heißt, die Anzahl der Mikrozustände des Makrozustandes vor dem Hochspringen wäre sehr viel größer als die Anzahl der Mikrozustände des Makrozustandes nach dem Hochspringen. Folglich ist der beschriebene Vorgang extrem unwahrscheinlich.

3a) Spezifische Wärmekapazität von Wasser bei  $17^\circ\text{C} \dots 37^\circ\text{C}$  :  $c = 4180 \text{ J/kg K}$ ,

$T =$  mittlere Temperatur

$$\Delta S = \frac{\Delta Q}{T} = \frac{c m \Delta T}{T} = \frac{4,180 \text{ kJ/kg K} \cdot 1000 \text{ kg} \cdot 10 \text{ K}}{295 \text{ K}} = 14169 \text{ J/K}$$

b)  $\Delta S = \frac{\Delta Q}{T} = \frac{c m \Delta T}{T} = \frac{4180 \text{ J/kg K} \cdot 1000 \text{ kg} \cdot 10 \text{ K}}{315 \text{ K}} = 13270 \text{ J/K}$

c) Sollen diese Prozesse spontan ablaufen, so muß Wärme aus der Umgebung auf das Wasser übergehen. Beim ersten Prozeß bedeutet dies, daß Wärme von einem heißeren auf ein kälteres System und beim zweiten, daß Wärme von einem kälteren auf ein heißeres System übergeht. Gemäß dem zweiten Hauptsatz kann daher nur der erste, nicht aber der zweite Prozeß von selbst ablaufen.

### Abschnitt 3.4

- 1a) Spezifische Schmelzwärme von Wasser:  $w = 334 \text{ kJ/kg}$

$$\Delta S = \frac{\Delta Q}{T} = \frac{w m}{T} = \frac{334000 \text{ J/kg} \cdot 100 \text{ kg}}{273 \text{ K}} = 122300 \text{ J/K}$$

Spezifische Wärmekapazität von Wasser bei  $0^\circ\text{C} \dots 37^\circ\text{C}$ :  $c = 4,218 \text{ kJ/kg K}$ ,

$$\Delta S = \frac{\Delta Q}{T} = \frac{c m \Delta T}{T} = \frac{4218 \text{ J/kg K} \cdot 100 \text{ kg} \cdot 1 \text{ K}}{273,5 \text{ K}} = 1542 \text{ J/K}$$

- b) Spezifische Verdampfungswärme von Wasser bei  $100^\circ\text{C}$  und  $1013 \text{ hPa}$ :

$$v = 2257 \text{ kJ/kg}$$

$$\Delta S = \frac{\Delta Q}{T} = \frac{v m}{T} = \frac{2257000 \text{ J/kg} \cdot 100 \text{ kg}}{373 \text{ K}} = 605094 \text{ J/K}$$

Spezifische Wärmekapazität von Wasserdampf bei  $100^\circ\text{C}$  und  $1013 \text{ hPa}$ :

$$c = 2,043 \text{ kJ/kg K}$$

$$\Delta S = \frac{\Delta Q}{T} = \frac{c m \Delta T}{T} = \frac{2043 \text{ J/kg K} \cdot 100 \text{ kg} \cdot 1 \text{ K}}{373 \text{ K}} = 548 \text{ J/K}$$

- c) In allen vier Fällen ist die Entropieänderung positiv, folglich laufen die Prozesse von selbst ab. Da beim Schmelzen beziehungsweise Verdampfen jeweils sehr viel mehr Entropie entsteht wie bei der reinen Temperaturerhöhung, nimmt dabei die Anzahl der Mikrozustände sehr viel stärker zu. Folglich sind Phasenübergänge mit einer sehr viel stärkeren Zunahme der inneren Unordnung verbunden als Temperaturänderungen, die ohne Phasenübergang, aber bei vergleichbaren Temperaturen stattfinden.

- 2)  $\eta = 1 - T_2/T_1 = 1 - 277 \text{ K}/298 \text{ K} = 0,705$

Das bedeutet, daß im Idealfall von jedem J thermischer Energie, die vom wärmeren auf das kältere Meereswasser übergeht, maximal  $0,705 \text{ J}$  mechanische Arbeit gewonnen werden kann.

- 3)  $\eta = 1 - T_2/T_1 = 1 - 300 \text{ K}/1000 \text{ K} = 0,7$

$\implies$  Es können maximal  $700 \text{ J}$  in mechanische Arbeit umgewandelt werden.

- 4)  $\eta = 1 - T_2/T_1 \implies T_1 = T_2/(1 - \eta)$

$$\eta = 0,4 \implies T_1 = 283 \text{ K}/0,6 = 472 \text{ K}$$

$$\eta = 0,7 \implies T_1 = 283 \text{ K}/0,3 = 943 \text{ K}$$

$\implies$  Die Temperatur des oberen Wärmereservoirs muß um  $471 \text{ K}$  erhöht werden.



### Abschnitt 3.5

- 1) Eine Kältemaschine entnimmt einem Wärmereservoir mit der tieferen Temperatur  $T_k$  die Wärmemenge  $Q_k$  und gibt diese an ein zweites Wärmereservoir mit der höheren Temperatur  $T_w$  ab; dazu muß ihr die Arbeit  $W$  zugeführt werden. Die Leistungszahl einer Kältemaschine ist der Quotient  $\varepsilon_L = Q_k/W$ . Nach dem ersten Hauptsatz der Thermodynamik ist dabei  $W = Q_w - Q_k$  und folglich

$$\varepsilon_L = \frac{Q_k}{Q_w - Q_k} = \frac{Q_k/Q_w}{1 - Q_k/Q_w}.$$

Da es sich um eine ideale Kältemaschine handeln soll, gilt außerdem  $Q_k/Q_w = T_k/T_w$  und somit

$$\varepsilon_L = \frac{1}{T_w/T_k - 1} = \frac{T_k}{T_w - T_k}.$$

Für  $T_k = 0$  wird  $\varepsilon_L = 0$  und folglich die aufzubringende Arbeit  $W$  unendlich.

- 2) Die Entropie ist ein Maß für die Anzahl der unterschiedlichen mikroskopischen Möglichkeiten, die zu makroskopisch nicht unterscheidbaren Zuständen des betrachteten Systems führen. Am absoluten Nullpunkt befinden sich alle Teilchen, aus welchen das System besteht, an festen Orten und in ihren energetisch tiefstmöglichen Zuständen. Jede Änderung würde daher zumindest eine Temperaturerhöhung nach sich ziehen, sodaß es nur einen einzigen möglichen mikroskopischen Zustand des Systems gibt. Am absoluten Nullpunkt ist folglich eine weitere Absenkung der Entropie nicht möglich.

### Abschnitt 3.6

1a) Abgebremstes Auto:  $W = \frac{1}{2} m v^2 = \frac{1}{2} \cdot 1000 \text{ kg} \cdot 100 \text{ m}^2/\text{s}^2 = 50000 \text{ J}$

Wärmeübertragung:  $\eta = 1 - T_k/T_w = 1 - 3/4 = 0,25$

$\Rightarrow W = Q - W = Q(1 - \eta) = Q(1 - 1/4) = 3Q/4 = 40000 \text{ J} \cdot 3/4 = 30000 \text{ J}$

Beim Bremsvorgang geht folglich mehr nutzbare Energie verloren.

b) Abgebremstes Auto:  $\Delta S = Q/T = W/T = 50000 \text{ J}/300 \text{ K} = 166,667 \text{ J/K}$

Wärmeübertragung:  $\Delta S = Q(1/T_k - 1/T_w) = 40000 \text{ J}(1/300 \text{ K} - 1/400 \text{ K})$   
 $= 33,33 \text{ J/K}$

Der Prozeß, bei dem mehr nutzbare Energie verloren geht, erzeugt mehr Entropie.

- 2) Zwischen der an das Heizwasser übertragenen Wärmemenge  $\Delta Q$  und der dazu vom Dieselaggregat aufgewendeten Energie  $\Delta E_D$  besteht die Relation

$$\Delta Q \leq \Delta E_D T_U / (T_H - T_U)$$

folglich gilt

$$\Delta Q \leq \Delta E_D \cdot 273 \text{ K} / (313 \text{ K} - 273 \text{ K}) = 6,825 \Delta E_D.$$

Die Wärmepumpe ist somit im Idealfall 6,825 mal so effektiv wie eine Dieselheizung.

- 3a) Da dem Kühlraum thermische Energie entzogen wird, nimmt dessen Entropie ab.
- b) Eine Kühlanlage arbeitet mit einer als Kältemaschine verwendeten Wärmepumpe, die dem zu kühlenden Raum bei der tieferen Temperatur  $T_k$  die Wärmemenge  $Q_k$  entzieht und diese der Umgebung bei der höheren Temperatur  $T_u$  zuführt. Dabei verringert sich die Entropie im Kühlraum um den Betrag  $\Delta S_k = Q_k/T_k$ , während sie sich in der Umgebung um  $\Delta S_u = Q_k/T_u$  erhöht. Für die gesamte Entropieänderung folgt daraus

$$\Delta S = Q_k \left( \frac{1}{T_u} - \frac{1}{T_k} \right) < 0,$$

das heißt, die gesamte Entropie scheint abzunehmen. Das ist nach dem zweiten Hauptsatz der Thermodynamik nicht möglich, sodaß ein weiterer Prozeß erforderlich ist, der zu einer Entropieerhöhung führt und damit die Entropieverringerng mindestens kompensiert. Dieser Prozeß muß eine zusätzliche Wärmemenge  $\Delta Q$  an die Umgebung abgeben, sodaß insgesamt

$$\frac{\Delta Q}{T_u} + \Delta S \geq 0$$

gilt. Dazu muß mindestens die Arbeit

$$W = \Delta Q = T_u \Delta S = Q_k \left( 1 - \frac{T_u}{T_k} \right)$$

aufgewendet werden.

- c) Die zu verrichtende Arbeit  $W$  ist um so größer, je größer die Differenz zwischen  $T_u$  und  $T_k$  ist. Deshalb sollte man nie tiefer kühlen als unbedingt erforderlich und dabei die Temperatur der Umgebung so wenig wie möglich überschreiten.
- d) Oben wurde die Relation

$$\varepsilon_L = \frac{T_k}{T_w - T_k}$$

hergeleitet. Mit  $\eta_C = 1 - T_k/T_w =$  folgt daraus

$$\varepsilon_L = \frac{T_k}{\eta_C T_w}.$$

## Kapitel 4

### Abschnitt 4.5.2

- 1) Elektronen:  $v = 10^7 \text{ m/s}$ ,  $m = 9,11 \cdot 10^{-31} \text{ kg}$ ,  
 $\lambda_{\text{dB}} = 6,626 \cdot 10^{-34} \text{ Js} / 9,11 \cdot 10^{-31} \text{ kg} \cdot 10^7 \text{ m/s} = 7,27 \cdot 10^{-11} \text{ m}$ .  
 Auto:  $v = 40 \text{ m/s}$ ,  $m = 1000 \text{ kg}$ ,  
 $\lambda_{\text{dB}} = 6,626 \cdot 10^{-34} \text{ Js} / 1000 \text{ kg} \cdot 40 \text{ m/s} = 1,66 \cdot 10^{-38} \text{ m}$ .  
 Erde:  $v = 29780 \text{ m/s}$ ,  $m = 6 \cdot 10^{24} \text{ kg}$ ,  
 $\lambda_{\text{dB}} = 6,626 \cdot 10^{-34} \text{ Js} / 6 \cdot 10^{24} \text{ kg} \cdot 29780 \text{ m/s} = 3,71 \cdot 10^{-63} \text{ m}$ .  
 Sonne:  $v = 225000 \text{ m/s}$ ,  $m = 2 \cdot 10^{30} \text{ kg}$ ,  
 $\lambda_{\text{dB}} = 6,626 \cdot 10^{-34} \text{ Js} / 2 \cdot 10^{30} \text{ kg} \cdot 225000 \text{ m/s} = 1,5 \cdot 10^{-69} \text{ m}$ .

- 2) De Broglie-Wellenlänge der Elefanten:  
 $\lambda_{\text{dB}} = h/mv = 6,626 \cdot 10^{-34} \text{ Js} / 12000 \text{ kg} \cdot 25 \text{ m/s} = 2,2 \cdot 10^{-39} \text{ m}$ ,  
 Bedingung für Beugungsmaxima:  $\sin \alpha_n = n\lambda_{\text{dB}}/g$   
 $\Rightarrow$  Ablenkungswinkel der Maxima erster Ordnung:  
 $\alpha_1 = \arcsin(2,2 \cdot 10^{-39} \text{ m} / 4 \text{ m}) = 3,15 \cdot 10^{-38} \text{ }^\circ$ .

- 3) De Broglie-Wellenlänge der Elektronen:  
 $\lambda_{\text{dB}} = 6,626 \cdot 10^{-34} \text{ Js} / 9,11 \cdot 10^{-31} \text{ kg} \cdot 3,25 \cdot 10^7 \text{ m/s} = 2,24 \cdot 10^{-11} \text{ m}$ ,  
 Bedingung für Maxima bei Bragg-Reflexion:

$$d = \frac{n\lambda}{2 \sin \alpha}$$

Bei den Beugungsringen handelt es sich jeweils um Maxima erster Ordnung:

$$d_i = \frac{\lambda}{2 \sin \alpha_i}, \quad i = 1, 2.$$

Für die Abstände der Kristallebenen ergibt das

$$d_1 = 2,24 \cdot 10^{-11} \text{ m} / 2 \sin 5,921^\circ = 1,09 \cdot 10^{-10} \text{ m};$$

$$d_2 = 2,24 \cdot 10^{-11} \text{ m} / 2 \sin 5,921^\circ = 6,4 \cdot 10^{-11} \text{ m}.$$

- 4) Bedingung für Beugungsmaxima:  $\sin \alpha_n = n\lambda_{\text{dB}}/g$   
 $\Rightarrow$  Ablenkungswinkel der Maxima erster Ordnung:  
 $\alpha_1 = \arcsin(1,03 \cdot 10^{-10} \text{ m} / 8 \cdot 10^{-6} \text{ m}) = 7,377 \cdot 10^{-4} \text{ }^\circ$ ,  
 $\Rightarrow$  Ablenkungswinkel der Maxima zweiter Ordnung:  
 $\alpha_2 = \arcsin(2 \cdot 2,223 \cdot 10^{-39} \text{ m} / 8 \cdot 10^{-6} \text{ m}) = 1,4754 \cdot 10^{-3} \text{ }^\circ$ .  
 $\Rightarrow$  Abstand der Maxima erster und zweiter Ordnung voneinander:  
 $\Delta x = d (\tan \alpha_2 - \tan \alpha_1) = 0,64 \text{ m} (\tan 1,4754 \cdot 10^{-3} \text{ }^\circ - \tan 7,377 \cdot 10^{-4} \text{ }^\circ) = 8,24 \text{ } \mu\text{m}$   
 $\Rightarrow$  Prozentuale Abweichung zwischen 0,5 % und 12,6 %.

- 5) Ist  $\alpha$  der Ablenkungswinkel eines Beugungsmaximums  $n$ -ter Ordnung, so folgt aus der Bedingung für Beugungsmaxima am Gitter und der Definition der de Broglie-Wellenlänge zunächst  $v = n h / m g \sin \alpha$ . Für den Abstand des Maximums von der Schirmmitte gilt außerdem  $x = d \tan \alpha$ . Mit der für kleine Winkel in guter Näherung geltenden Relation  $\sin \alpha \approx \tan \alpha$  erhält man  $v = n h d / m g x$  und mit  $n = 1$  weiter

$$v = 6,626 \cdot 10^{-34} \text{ Js} \cdot 1,25 \text{ m} / 1,2 \cdot 10^{-24} \text{ kg} \cdot 10^{-7} \text{ m} \cdot 5 \cdot 10^{-8} \text{ m} = 138000 \text{ m/s.}$$

- 6a) Zunächst passiert nichts. Ab einer bestimmten Wellenlänge mißt man zwischen den Elektroden der Photozelle eine Photospannung, die mit abnehmender Wellenlänge zunimmt.

- b) Die schnellsten Elektronen werden durch violette Photonen ausgelöst; deren Frequenz beträgt  $7,895 \cdot 10^{14}$  Hz.

$$\begin{aligned} \text{Energie:} \quad E &= h\nu - E_A \\ &= 6,626 \cdot 10^{-34} \text{ Js} \cdot 7,895 \cdot 10^{14} \text{ Hz} - 3,104 \cdot 10^{-19} \text{ J} \\ &= 2,187 \cdot 10^{-19} \text{ J} = 1,367 \text{ eV;} \end{aligned}$$

$$\begin{aligned} \text{Geschwindigkeit: } v &= \sqrt{2E/m} = \sqrt{2 \cdot 2,187 \cdot 10^{-19} \text{ J} / 9,11 \cdot 10^{-31} \text{ kg}} \\ &= 6,93 \cdot 10^5 \text{ m/s.} \end{aligned}$$

- c) Neben dem violetten Ende des Spektrums befindet sich dessen ultravioletter Teil. Die zugehörigen Photonen haben höhere Energien als die violetten und können ebenfalls Elektronen aus der Photokathode auslösen.
- d) Neben dem roten Ende des Spektrums befindet sich dessen infraroter Teil. Die zugehörigen Photonen haben geringere Energien als die roten und können keine Elektronen aus der Photokathode auslösen.
- e) Mißt man die Energie der ausgelösten Elektronen in Abhängigkeit von der Frequenz des Lichts, so erhält man eine Gerade, deren Steigung die Plancksche Konstante ist.
- f) Die Energie eines Photons ist proportional zu dessen Frequenz. Die Plancksche Konstante ist der zugehörige Proportionalitätsfaktor.
- h) Die Energie der Photoelektronen hängt nur von der Frequenz des verwendeten Lichts ab, ihre Anzahl nur von dessen Intensität. Das läßt sich deuten, indem man annimmt, daß jedes Photoelektron von einem Photon ausgelöst wird.

### Abschnitt 4.5.3

- 1a) Es gilt in guter Näherung (siehe oben)  $v = n h d / m g x$  und mit  $n = 1$  und  $d = 7,5 \mu\text{m}$  folglich
- $$v = 6,626 \cdot 10^{-34} \text{ Js} \cdot 0,34 \text{ m} / 9,11 \cdot 10^{-31} \text{ kg} \cdot 10^{-6} \text{ m} \cdot 7,5 \cdot 10^{-5} \text{ m}$$
- $$= 3,32 \cdot 10^6 \text{ m/s.}$$
- b) Zunächst treten auf dem Schirm scheinbar zufällig verteilte einzelne Einschläge auf. Nach und nach entsteht ein Interferenzmuster aus hellen und dunklen Streifen, das aus vielen einzelnen Einschlägen besteht. Das Auftreten einzelner Einschläge läßt sich mit einem Teilchen-, nicht aber mit einem Wellenbild erklären; das Auftreten eines Interferenzmusters läßt sich mit einem Wellen-, nicht aber mit einem Teilchenbild erklären. Da beides gleichzeitig beobachtet wird, ist das gesamte Phänomen weder mit einem Teilchen- noch mit einem Wellenbild erklärbar.
- 2a) Wie man in Abbildung 4.2 sieht, können Neutronen, die unter einem bestimmten Winkel  $\alpha$  auf dem Beobachtungsschirm auftreten, an verschiedenen Kristallebenen gestreut worden sein, sodaß sie auf mehreren ununterscheidbaren Wegen zum Schirm gelangen können. Damit ist keine Welcher-Weg-Information zu erhalten, und es tritt ein Interferenzmuster auf.
- b) Die Spin-Änderungen erlauben im Prinzip den Nachweis, an welchem der Kristallionen die betreffenden Neutronen jeweils gestreut wurden, sodaß für diese Neutronen Welcher-Weg-Information erhalten werden könnte und die Interferenzfähigkeit verloren geht. Folglich wird das Interferenzmuster von einem gleichverteilten Untergrund überlagert.
- 3) Beim Michelson-Interferometer gelangen Photonen durch den Strahlteiler S1 entweder über den Spiegel S2 oder den Spiegel S3 auf den Detektor D. Es gibt folglich die beiden ununterscheidbaren Wege Q S1 S2 D und Q S1 S3 D, sodaß keinerlei Welcher-Weg-Information verfügbar ist und am Detektor konstruktive Interferenz auftritt. Da die Quelle die Photonen nicht nur in waagrechter Richtung, sondern auch in axial dazu abweichenden Richtungen emittiert, entsteht am Detektor ein Interferenzmuster in Form konzentrischer Kreise mit einem Maximum im Mittelpunkt. Bringt man Polarisationsfilter in die Strahlengänge, so ist analog wie beim Mach-Zehnder-Interferometer je nach deren Orientierung mehr oder weniger Welcher-Weg-Information möglich, und entsprechend wird das Interferenzmuster undeutlicher oder deutlicher. Bei maximaler Welcher-Weg-Information, wie sie durch gekreuzte Polarisationsfilter erhältlich ist, tritt kein Interferenzmuster auf.
- 4a) Die Idler-Photonen machen die beiden Wege unterscheidbar, denn wenn Detektor D1 anspricht, ist das Signal-Photon über den Weg Q S1 K1 S2 S4 D gelaufen, spricht Detektor D2 an, so hat es den Weg Q K2 S3 S4 D genommen. Folglich liegt maximale Welcher-Weg-Information vor, und bei Detektor D tritt kein Interferenzmuster auf.

- b) Der zusätzliche Strahlteiler S sorgt dafür, daß keine Welcher-Weg-Information mehr zu bekommen ist, da nun bei beiden klassisch möglichen Wegen mit gleicher Wahrscheinlichkeit sowohl Detektor D1 als auch Detektor D2 ansprechen kann. Folglich tritt bei Detektor D ein Interferenzmuster auf<sup>1</sup>.

## Abschnitt 4.5.4

- 1) Zu Grundregel 2:

Tritt bei einem Interferenzmuster mit einzelnen Photonen in einem Mach-Zehnder-Interferometer ein Interferenzmuster auf, so interferiert dabei jedes einzelne Photon nur mit sich selbst. Es liegt also in einem Überlagerungszustand vor, in dem es beide klassisch möglichen Wege gleichzeitig durchläuft. Folglich können ihm klassische Eigenschaften wie diejenigen, einen Ort oder eine Geschwindigkeit zu besitzen, nicht zugeschrieben werden.

Zu Grundregel 3:

Mißt man den Ort eines Photons in einem Mach-Zehnder-Interferometer, indem man etwa einen Detektor in einen von dessen beiden Arme hält, so erhält man zwar ein eindeutiges Meßergebnis, denn man stellt fest, daß der Detektor entweder anspricht oder nicht, das heißt, entweder man findet das Photon in dem ausgewählten Arm oder nicht. Das heißt jedoch nicht, daß das Photon die Eigenschaft, *sich in einem der beiden Arme zu befinden*, unmittelbar vor der Messung überhaupt besessen hat. Wenn es ohne den Detektor zu einem Interferenzmuster gekommen wäre, hatte es diese Eigenschaft unmittelbar vor der Messung definitiv nicht.

Zu Grundregel 4:

Nach einer Messung wie der soeben beschriebenen ist klar, ob sich das Photon nun am Ort des Detektors im ausgewählten Arm des Mach-Zehnder-Interferometers befindet oder nicht, das heißt, sofern das Photon im Detektor überhaupt noch feststellbar ist, liefert eine sofortige erneute Überprüfung von dessen Ort das selbe Resultat. Das Photon befindet sich somit in einem Ortseigenzustand<sup>2</sup>.

- 2) Grundregeln der klassischen Punktmechanik:

1. Der Zustand eines physikalischen Systems wird durch eine Weg-Zeit-Funktion  $x$  vollständig beschrieben.
2. Die dynamischen Größen  $x$  und  $\dot{x} = v$  des Systems sind zu jedem Zeitpunkt simultan definiert.
3. An einem System sind beliebige unterschiedliche physikalische Größen prinzipiell gleichzeitig streuungsfrei meßbar, und wenn bei der Messung einer Größe ein bestimmter Meßwert erhalten wird, hatte das System die gemessene Eigenschaft unmittelbar vor der Messung bereits.

<sup>1</sup>Auch diese Anordnung wird in der Literatur unter der Bezeichnung *Quantenradierer* geführt.

<sup>2</sup>Man beachte jedoch auch hier das in Anmerkung 4 auf Seite 69 sowie in den Abschnitten 2.6.3 und 4.2.2 gesagte.

4. Messungen sind prinzipiell wiederholbar, und bei unmittelbarer Wiederholung von Messungen kommt stets wieder der selbe Meßwert heraus.
5. Sind bei einem System die Anfangswerte von  $x$  und  $\dot{x}$  bekannt, so kann das weitere Verhalten des Systems in alle Ewigkeit exakt vorausberechnet werden.
6. Die Weg-Zeit-Funktion des Systems ist gegeben als Lösung der Bewegungsgleichung

$$m\ddot{x} = F$$

des Systems; dabei ist  $m$  die Masse des Systems und  $F$  die äußere Kraft, die auf das System wirkt.

- 3) Träfe diese Erklärung zu, so lägen die Merkwürdigkeiten der Quantenmechanik allein an der Kleinheit atomarer und subatomarer Objekte. Insbesondere wäre dann der Einfluß von Messungen auf gewöhnliche physikalische Wechselwirkungen reduziert, die durch die entsprechenden dynamischen Gleichungen beschreibbar sein müßten und keineswegs irgendwelche unstetige Veränderungen auslösen könnten. Die spontane Umwandlung von Superpositionen aus Eigenzuständen in Eigenzustände ist damit nicht erklärbar. Es dürfte dann auch keine makroskopische Quantenphänomene geben. Die Quantenmechanik wäre eine verfeinerte Version der klassischen Physik, die erst bei der Beschreibung mikroskopischer Prozesse erforderlich wird. Abgesehen davon ist es einfach nicht wahr, daß klassische Messungen stets von vernachlässigbarer Wirkung auf das gemessene System sind. Das folgende Beispiel mag das veranschaulichen. Ein Fußballspieler, der eine Flanke direkt aus der Luft auf das Tor drischt, nimmt damit physikalisch betrachtet eine Ortsmessung am System Ball vor; von einer vernachlässigbaren Wirkung oder Störung dieser Messung auf den Ball kann jedoch nicht wirklich die Rede sein. Dennoch läßt sich der Vorgang im Rahmen der klassischen Mechanik beschreiben, und es treten keinerlei paradoxe Phänomene auf, höchstens vielleicht aus sportlicher Sicht. Irgendwelche Störungen des gemessenen Systems durch die Messung können also keineswegs die Ursache der Besonderheiten quantenmechanischer Messungen sein.

### Abschnitt 4.5.5

- 1) Wir betrachten den Fußball als Massenpunkt; seine Masse betrage 450 g.  
 $\Delta p \geq \hbar/\Delta x \implies \Delta v \geq \hbar/m \Delta x = 6,626 \cdot 10^{-34} \text{ Js} / 0,45 \text{ kg} \cdot 7,32 \text{ m} = 2 \cdot 10^{-34} \text{ m}.$   
 Unberechenbarkeiten der Ballbewegung in dieser Größenordnung dürften einen guten Torwart schwerlich irritieren.
- 2)  $\Delta v \geq \hbar/m \Delta x = 6,626 \cdot 10^{-34} \text{ Js} / 9,11 \cdot 10^{-31} \text{ kg} \cdot 10^{-10} \text{ m} = 7,3 \cdot 10^6 \text{ m}.$
- 3) Unschärfen sind Streuungen von Meßreihen. Folglich sind Unschärferelationen stets Aussagen über sehr viele Messungen an identischen Systemen. Die Anwendung auf eine einzelne Ortsmessung an einem Elektron ist daher sinnlos.
- 4) Diese Aussage ist vollkommen falsch. Jede quantenmechanische Messung ist abgesehen von technischen Unzulänglichkeiten prinzipiell beliebig genau. Daher machen



Unschärferelationen keinerlei Aussagen über Meßungenauigkeiten. Richtig ist folgendes:

Die Heisenbergsche Unschärferelation besagt, daß es nicht möglich ist, bei sehr vielen Ortsmessungen und sehr vielen Impulsmessungen an identischen Kopien eines quantenmechanischen Systems die Streuungen dieser Messungen beliebig klein zu halten. Je kleiner die Streuung der Ortsmessungen sind, desto größer sind diejenigen der Impulsmessungen und umgekehrt.

### Abschnitt 4.5.6

- 1) Für  $0 \leq x \leq a$  lautet die Schrödingergleichung

$$-\frac{\hbar^2}{2m} \psi'' + V_0 \psi = E \psi.$$

Damit folgt völlig analog zu Abschnitt 4.5.6.2

$$\psi_n(x) = \begin{cases} \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi x}{a}\right) & \text{für } 0 \leq x \leq a \\ 0 & \text{sonst} \end{cases}$$

und

$$E_n = \frac{\hbar^2 \pi^2 n^2}{2m a^2} + V_0, \quad n \in \mathbb{N} \setminus \{0\}.$$

- 2) Bei 22 Elektronen und je zwei Elektronen pro Energieniveau handelt es sich beim tiefsten Übergang um denjenigen vom 11. in das 12. Niveau. Für die Übergangsfrequenzen gilt

$$\nu_{n,m} = \frac{\Delta E_{n,m}}{h} = \frac{E_m - E_n}{h} = \frac{\pi \hbar}{4m a^2} (m^2 - n^2)$$

und damit

$$a^2 = \frac{\pi \hbar}{4m \nu_{n,m}} (m^2 - n^2) = \frac{\pi \hbar \lambda_{m,n}}{4m c} (m^2 - n^2).$$

Die Länge des Moleküls ist folglich

$$a = \sqrt{\frac{\pi \hbar \lambda_{m,n}}{4m c} (m^2 - n^2)} = \sqrt{\frac{\pi \hbar \lambda_{12,11}}{4m c} (12^2 - 11^2)} = 1,774 \text{ nm}.$$

- 3a) Ein harmonischer Oszillator, das heißt ein harmonisch schwingendes System, ist klassisch dadurch definiert, daß die Rückstellkraft, welche die Schwingungsbewegung verursacht, stets proportional zur momentanen Auslenkung ist; es gilt also  $F = kx$  mit einer Proportionalitätskonstanten  $k$ . Die potentielle Energie des Systems gibt an, wieviel Arbeit verrichtet wurde, um es aus der Ruhelage an die momentane ausgelenkte Position zu bringen. Daraus folgt

$$E_{\text{pot}} = \int_0^x F dx' = \int_0^x k x' dx' = \frac{1}{2} k x^2.$$

Als Potential wählt man daher  $V = \frac{1}{2} k x^2$ .

b) Die Schrödingergleichung des eindimensionalen harmonischen Oszillators lautet

$$-\frac{\hbar^2}{2m} \psi'' + \frac{1}{2} k x^2 \psi = E \psi.$$

Zweimaliges Ableiten des Ansatzes

$$\psi_0(x) = B e^{-a x^2}$$

liefert

$$\psi_0'(x) = -2Bax e^{-a x^2}, \quad \psi_0''(x) = -2Ba(1 - 2ax^2) e^{-a x^2}$$

und anschließendes Einsetzen in die Schrödingergleichung weiter

$$\left(\frac{k}{2} - \frac{2\hbar^2 a^2}{m}\right) x^2 + \frac{a\hbar^2}{m} - E = 0.$$

Diese Gleichung kann für beliebige  $x$  nur dann erfüllt sein, wenn

$$\frac{k}{2} - \frac{2\hbar^2 a^2}{m} = \frac{a\hbar^2}{m} - E = 0 \tag{C.1}$$

gilt<sup>3</sup>; mit  $\omega = \sqrt{k/m}$  folgt daraus  $a^2 = \omega^2/4\hbar^2$ , und die Wellenfunktion des Grundzustands lautet zunächst

$$\psi_0(x) = B e^{-\omega x^2/2\hbar}.$$

Diese muß nun normiert werden, das heißt man fordert

$$\int_{-\infty}^{\infty} |\psi_0(x)|^2 dx = 1.$$

Daraus ergibt sich

$$B^2 \int_{-\infty}^{\infty} e^{-2ax^2} dx = B^2 \sqrt{\frac{\pi}{2a}} = 1,$$

das heißt  $B^2 = \sqrt{2a/\pi}$ , und man erhält die Grundzustandswellenfunktion

$$\psi_0(x) = \left(\frac{2a}{\pi}\right)^{1/4} e^{-\omega x^2/2\hbar}.$$

c) Aus (C.1) folgt auch

$$E = \frac{a\hbar^2}{m} = \frac{1}{2} \hbar \sqrt{\frac{k}{m}}$$

und damit

$$E = \frac{1}{2} \hbar \omega.$$

---

<sup>3</sup>Formal gesehen führt man hier wie in den beiden folgenden Teilaufgaben einen Koeffizientenvergleich durch.

d) Zweimaliges Ableiten des Ansatzes

$$\psi_1(x) = B x e^{-ax^2}$$

liefert

$$\psi_1'(x) = B(1 - 2ax^2)e^{-ax^2}, \quad \psi_1''(x) = 2Ba(2ax^3 - x)e^{-ax^2}$$

und Einsetzen in die Schrödingergleichung

$$\left(\frac{k}{2} - \frac{2\hbar^2 a^2}{m}\right) x^3 + \left(\frac{3a\hbar^2}{m} - E\right) = 0,$$

eine Gleichung, die wiederum für beliebige  $x$  nur erfüllt sein kann, wenn

$$\frac{k}{2} - \frac{2\hbar^2 a^2}{m} = \frac{3a\hbar^2}{m} - E = 0$$

gilt. Das liefert  $a^2 = \omega^2/4\hbar^2$  und

$$E = \frac{3a\hbar^2}{m} = \frac{3}{2} \hbar \sqrt{\frac{k}{m}}.$$

Die Energie des ersten angeregten Zustands ist damit

$$E = \frac{3}{2} \hbar^2 \omega.$$

e) Zweimaliges Ableiten des Ansatzes

$$\psi_2(x) = B(2ax^2 - 1)e^{-ax^2/2}$$

liefert

$$\psi_2'(x) = Ba(5x - 2ax^3)e^{-ax^2/2}, \quad \psi_2''(x) = Ba(2a^2x^4 - 11ax^2 + 5)e^{-ax^2/2};$$

setzt man dies in die Schrödingergleichung ein, erhält man

$$\left(ka - \frac{\hbar^2 a^3}{m}\right) x^4 + \left(\frac{11a^2 \hbar^2}{2m} - \frac{k}{2} - 2aE\right) x^2 + E - \frac{5a\hbar^2}{2m} = 0.$$

Diese Gleichung ist für beliebige  $x$  nur erfüllbar, wenn

$$ka - \frac{\hbar^2 a^3}{m} = \frac{11a^2 \hbar^2}{2m} - \frac{k}{2} - 2aE = E - \frac{5a\hbar^2}{2m} = 0 \quad (\text{C.2})$$

gilt. Daraus folgt  $a^2 = km/\hbar^2$  sowie

$$E = \frac{5a\hbar^2}{2m} = \frac{5}{2} \hbar \sqrt{\frac{k}{m}}.$$

Damit lautet die Energie des zweiten angeregten Zustands

$$E = \frac{5}{2} \hbar \omega.$$

- f) Die allgemeine Formel für die Energieeigenwerte des eindimensionalen harmonischen Oszillators lautet

$$E_n = \hbar \omega \left( n + \frac{1}{2} \right).$$

- h) Man kann keinerlei Aussagen über diese Größen machen. Da es sich bei den Oszillatorzuständen um Energieeigenzustände handelt, kann man einem solchen System die Eigenschaft, über einen Ort und einen Impuls zu verfügen, nicht zuschreiben. Die potentielle Energie ist eine Funktion des Ortes und die kinetische Energie eine Funktion der Geschwindigkeit, folglich verfügt das System auch nicht über die Eigenschaft, potentielle oder kinetische Energie zu besitzen.
- 4) Zweites Energieniveau: Zweimaliges Ableiten der Wellenfunktion

$$\psi_2(r) = \frac{1}{\sqrt{8\pi a^3}} \left( 1 - \frac{r}{2a} \right) e^{-r/2a}$$

liefert

$$\psi_2'(r) = \frac{1}{\sqrt{8\pi a^3}} \left( \frac{r}{4a^2} - \frac{1}{a} \right) e^{-r/2a}, \quad \psi_2''(r) = \frac{1}{\sqrt{128\pi a^7}} \left( 3 - \frac{r}{2a} \right) e^{-r/2a},$$

und Einsetzen in die Schrödingergleichung weiter

$$-\frac{5\hbar^2}{4ma^2} + \frac{e^2}{8\pi\epsilon_0 a} - E + \left( \frac{\hbar^2}{16ma^3} + \frac{E}{2a} \right) r + \left( \frac{\hbar^2}{ma} - \frac{e^2}{4\pi\epsilon_0} \right) \frac{1}{r} = 0.$$

Das soll für alle  $r$  gelten, woraus

$$-\frac{5\hbar^2}{4ma^2} + \frac{e^2}{8\pi\epsilon_0 a} - E = \frac{\hbar^2}{16ma^3} + \frac{E}{2a} = \frac{\hbar^2}{ma} - \frac{e^2}{4\pi\epsilon_0} = 0$$

folgt. Die zweite dieser drei Gleichungen liefert

$$E = -\frac{\hbar^2}{8ma^2},$$

die dritte

$$a = \frac{4\pi\epsilon_0 \hbar^2}{me^2},$$

und zusammen erhält man

$$E_2 = -\frac{me^4}{128\hbar^2\pi^2\epsilon_0^2}.$$

Drittes Energieniveau: Zweimaliges Ableiten der Wellenfunktion

$$\psi_3(r) = \frac{1}{\sqrt{27\pi a^3}} \left( 1 - \frac{2r}{3a} + \frac{2r^2}{27a^2} \right) e^{-r/3a}$$

ergibt

$$\psi_3'(r) = -\frac{1}{\sqrt{27\pi a^5}} \left( 1 - \frac{10r}{27a} + \frac{2r^2}{81a^2} \right) e^{-r/3a},$$

$$\psi_3''(r) = \frac{1}{\sqrt{27\pi a^7}} \left( \frac{19}{27} - \frac{14r}{81a} + \frac{2r^2}{243a^2} \right) e^{-r/3a},$$

Einsetzen in die Schrödingergleichung sodann

$$\begin{aligned} -\frac{13\hbar^2}{18ma^2} + \frac{e^2}{6\pi\epsilon_0 a} - E + \left( \frac{\hbar^2}{9ma^3} - \frac{e^2}{54\pi\epsilon_0 a^2} + \frac{2E}{3a} \right) r \\ - \left( \frac{\hbar^2}{243ma^4} - \frac{2E}{27a^2} \right) r^2 + \left( \frac{\hbar^2}{ma} - \frac{e^2}{4\pi\epsilon_0} \right) \frac{1}{r} = 0. \end{aligned}$$

Soll das für alle  $r$  der Fall sein, folgt daraus

$$\begin{aligned} -\frac{13\hbar^2}{18ma^2} + \frac{e^2}{6\pi\epsilon_0 a} - E &= \frac{\hbar^2}{9ma^3} - \frac{e^2}{54\pi\epsilon_0 a^2} + \frac{2E}{3a} \\ &= \frac{\hbar^2}{243ma^4} - \frac{2E}{27a^2} = \frac{\hbar^2}{ma} - \frac{e^2}{4\pi\epsilon_0} = 0. \end{aligned}$$

Die dritte dieser vier Gleichungen liefert

$$E = -\frac{\hbar^2}{18a^2},$$

die vierte

$$a = \frac{4\pi\epsilon_0 \hbar^2}{me^2},$$

und insgesamt erhält man

$$E_3 = -\frac{me^4}{288\pi^2\epsilon_0^2\hbar^2}.$$

$$\begin{aligned} 5a) \nu_{200,201} &= \frac{\Delta E_{200,201}}{h} = \frac{me^4}{64\pi^3\epsilon_0^2\hbar^3} \left( \frac{1}{n_1^2} - \frac{1}{n_2^2} \right) \\ &= \frac{9,11 \cdot 10^{-31} \text{ kg} \cdot (1,6 \cdot 10^{-19})^4 \text{ C}}{64 \cdot \pi^3 \cdot (8,85 \cdot 10^{-12} \text{ C/Vm})^2 \cdot (1,055 \cdot 10^{-34} \text{ Js})^3} \left( \frac{1}{200^2} - \frac{1}{201^2} \right) \\ &= 8,164 \cdot 10^8 \text{ Hz} \end{aligned}$$

$$\Rightarrow \lambda_{200,201} = \frac{c}{\nu_{200,201}} = \frac{3 \cdot 10^8 \text{ m/s}}{8,164 \cdot 10^8 \text{ 1/s}} = 0,3675 \text{ m}.$$

$$\begin{aligned} b) E &= \frac{me^4}{32\pi^2\epsilon_0^2\hbar^2 n^2} = \frac{9,11 \cdot 10^{-31} \text{ kg} \cdot (1,6 \cdot 10^{-19})^4 \text{ C}}{32 \cdot \pi^2 \cdot (8,85 \cdot 10^{-12} \text{ C/Vm})^2 \cdot (1,055 \cdot 10^{-34} \text{ Js})^2 \cdot 400^2} \\ &= 1,36 \cdot 10^{-23} \text{ J} = 8,5 \cdot 10^{-5} \text{ eV} \end{aligned}$$

6) Aluminium hat die Ordnungszahl 13

$$\begin{aligned} \Rightarrow E &= \frac{Zme^4}{32\pi^2\epsilon_0^2\hbar^2} = n \frac{13 \cdot 207 \cdot 9,11 \cdot 10^{-31} \text{ kg} \cdot (1,6 \cdot 10^{-19})^4 \text{ C}}{32 \cdot \pi^2 \cdot (8,85 \cdot 10^{-12} \text{ C/Vm})^2 \cdot (1,055 \cdot 10^{-34} \text{ Js})^2} \\ &= 5,86 \cdot 10^{-15} \text{ J} = 36598 \text{ eV} \end{aligned}$$

$$\Rightarrow \lambda = \frac{hc}{E} = \frac{6,626 \cdot 10^{-34} \text{ Js} \cdot 3 \cdot 10^8 \text{ m/s}}{5,86 \cdot 10^{-15} \text{ J}} = 33,95 \text{ pm}$$

- 7a) Setzt man den Energienullpunkt auf die Erdoberfläche, so gilt für die potentielle Energie eines Teilchens der Masse  $m$  im Gravitationsfeld der Erde in der Höhe  $x$  über dem Erdboden bei kleinen Höhen näherungsweise  $E_{\text{pot}} = mgx$ .

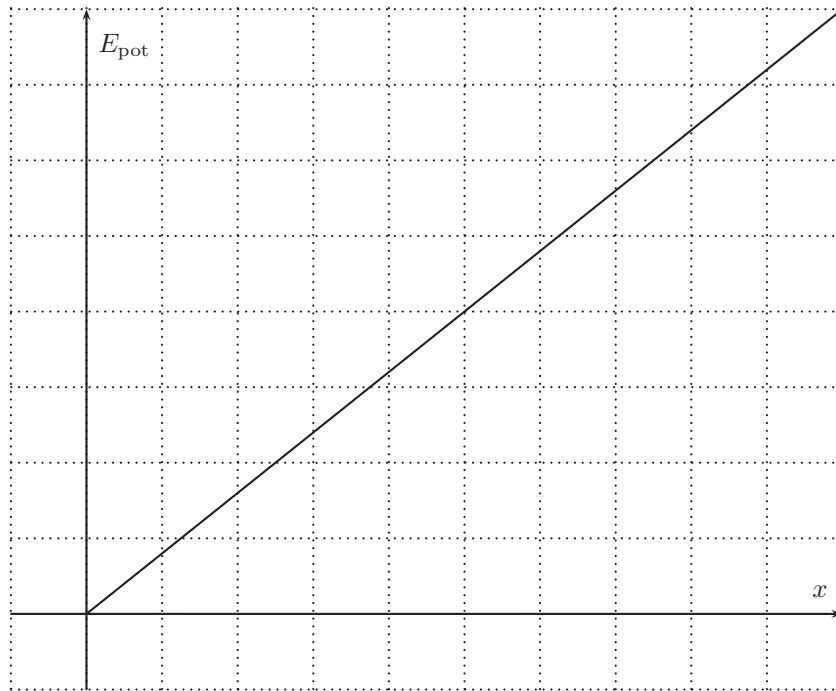


Abbildung 4.25: Gravitationspotential in der Nähe der Erdoberfläche

- b) Mit  $V = mgx$  lautet die Schrödingergleichung

$$-\frac{\hbar^2}{2m} \psi'' + mgx\psi = E\psi.$$

### Abschnitt 4.5.7

$$\begin{aligned} 1) \langle \Psi | \Psi \rangle &= \left\langle \frac{1}{\sqrt{2}} (\langle \uparrow \uparrow | + \langle \rightarrow \rightarrow |) \left| \frac{1}{\sqrt{2}} (|\uparrow \uparrow \rangle + |\rightarrow \rightarrow \rangle) \right. \right\rangle \\ &= \frac{1}{2} (\langle \uparrow \uparrow | \uparrow \uparrow \rangle + \langle \uparrow \uparrow | \rightarrow \rightarrow \rangle + \langle \rightarrow \rightarrow | \uparrow \uparrow \rangle + \langle \rightarrow \rightarrow | \rightarrow \rightarrow \rangle) \\ &= \frac{1}{2} \cdot (1 + 0 + 0 + 1) = \frac{1}{2} \cdot 2 = 1 \end{aligned}$$

- 2) Direktes Ausmultiplizieren zeigt jeweils die Äquivalenz der beiden Darstellungen von  $|\Psi_1\rangle$  beziehungsweise  $|\Psi_2\rangle$ . Folglich repräsentiert  $|\Psi\rangle$  einen Produktzustand, da dieser durch einfaches Zusammenkleben zweier Zustände darstellbar ist.  $|\Psi_2\rangle$

dagegen repräsentiert einen verschränkten Zustand, da seine Wellenfunktion auch nach Umformung eine Summe aus mehreren zusammengeklebten Komponenten ist. Anmerkung: Genaugenommen müßte man zusätzlich zeigen, daß die hier verwendete sogenannte *Schmidt-Zerlegung* eindeutig ist und keine weitere Vereinfachung gestattet. Allgemein gilt für zweikomponentige Systeme: Besteht die Schmidt-Zerlegung des zugehörigen Zustands aus nur einem Summanden, handelt es sich um einen Produktzustand, enthält sie mehr als einen Summanden, ist der Zustand verschränkt. Eine genauere Betrachtung dieses Sachverhalts würde jedoch jeden vernünftigen Rahmen gymnasialen Physikunterrichts bei weitem sprengen<sup>4</sup>.

$$3) |\Psi_1\rangle = \frac{1}{\sqrt{3}} (|\uparrow\downarrow\downarrow\rangle + |\downarrow\uparrow\downarrow\rangle + |\downarrow\downarrow\uparrow\rangle)$$

Dieser Zustand beschreibt zum Beispiel drei verschränkte Elektronen, wobei jeweils zwei mit Spin aufwärts und das dritte mit Spin abwärts polarisiert sind<sup>5</sup>. Eine Spinnmessung an einem System aus in dieser Weise verschränkten Elektronen, etwa mit einem Filter, der Elektronen mit Spin aufwärts durchläßt und solche mit Spin abwärts absorbiert, liefert stets zwei unverschränkte, parallel polarisierte Photonen.

$$|\Psi_2\rangle = \frac{1}{\sqrt{3}} (|\uparrow\uparrow\uparrow\rangle + |\uparrow\uparrow\downarrow\rangle + |\uparrow\downarrow\downarrow\rangle + |\downarrow\downarrow\downarrow\rangle)$$

Dieser Zustand beschreibt unter anderem drei verschränkte Elektronen, wobei sich diese in einer Überlagerung aus dreimal Spin aufwärts, zweimal Spin aufwärts und einmal Spin abwärts, einmal Spin aufwärts und zweimal Spin abwärts sowie dreimal Spin abwärts befinden. Eine Messung des Spins an einem solchen System verschränkter Elektronen mit einem Filter für Elektronen mit Spin aufwärts oder abwärts liefert in statistisch gemischter Weise vollständig absorbierte Elektronen oder ein Elektron mit Spin aufwärts oder zwei oder drei unverschränkte Elektronen mit parallelen Spins.

4) Der Zustand

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|\uparrow\rightarrow\rangle + |\rightarrow\uparrow\rangle)$$

beschreibt beispielsweise zwei verschränkte Photonen, deren Polarisationsrichtungen stets orthogonal zueinander sind. Zwei grundsätzliche Möglichkeiten sind zu unterscheiden:

1. Polarisationsfilter mit senkrechter oder waagrechtlicher Richtung:

Jeweils eines der beiden Komponenten-Photonen befindet sich in einem Eigenzustand des Polarisationsfilters. Folglich absorbiert der Filter eines der

---

<sup>4</sup>Das Verfahren ist aus der linearen Algebra bekannt und geht auf Erhard Schmidt zurück [451]. Genauereres dazu findet man beispielsweise in [27]. Die Schmidt-Zerlegung ist nur bei zweikomponentigen zusammengesetzten Systemen immer möglich; bei Systemen mit mehr als zwei Komponenten gibt es nichts von vergleichbarer Allgemeinheit. Wenigstens kann man für drei Komponenten zeigen, daß eine sogenannte triorthogonale Zerlegung  $|\Psi\rangle = \sum_i a_i |\psi_i\varphi_i\chi_i\rangle$  zwar nicht immer existiert, sie aber, wenn es sie doch gibt, stets eindeutig ist [84].

<sup>5</sup> $|\Psi_1\rangle$  ist ein Beispiel für einen sogenannten *Werner-Zustand* oder *W-Zustand*, benannt nach R. F. Werner, der diese Zustände 1989 entdeckte [520].



Photonen und läßt das andere durch, das anschließend parallel zur Polarisationsrichtung des Filters polarisiert ist.

2. Polarisationsfilter mit beliebiger anderer Richtung:

Keines der beiden Komponenten-Photonen befindet sich in einem Eigenzustand des Polarisationsfilters. Beide Photonen werden vom Polarisationsfilter mit einer Wahrscheinlichkeit durchgelassen, die gleich dem Quadrat des Cosinus des Winkels zwischen den Polarisationsrichtungen des jeweiligen Photons und des Filters ist (Gesetz von Malus). Die Photonen sind nach der Messung nicht mehr verschränkt.

- 5) Gemeint sind hier natürlich Quantenkorrelationen, und diese sind stets Konsequenz verschränkter Zustände. Man betrachte dazu etwa wieder den Zustand

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|\uparrow \rightarrow\rangle + |\rightarrow \uparrow\rangle)$$

sowie eine beliebige Polarisationsmessung am ersten der beiden Photonen, beispielsweise mit einem Polarisationsfilter, dessen Eigenzustand  $|\searrow\rangle$  sei. Jeder polarisierte Photonenzustand kann als Linearkombination der Zustände  $|\searrow\rangle$  und  $|\nearrow\rangle$  geschrieben werden, obiger Zustand läßt sich folglich auch in der Form

$$\begin{aligned} |\Psi\rangle &= \frac{1}{\sqrt{2}} [(\alpha|\searrow\rangle + \beta|\nearrow\rangle) \otimes (\gamma|\searrow\rangle + \delta|\nearrow\rangle) \\ &\quad + (\gamma|\searrow\rangle + \delta|\nearrow\rangle) \otimes (\alpha|\searrow\rangle + \beta|\nearrow\rangle)] \\ &= \sqrt{2} (\alpha\gamma|\searrow\searrow\rangle + \beta\gamma|\nearrow\searrow\rangle + \alpha\delta|\searrow\nearrow\rangle + \beta\delta|\nearrow\nearrow\rangle) \end{aligned}$$

schreiben, mit geeigneten Koeffizienten  $\alpha, \beta, \gamma, \delta$ . Die beschriebene Messung projiziert den Zustand auf den Eigenzustand der gemessenen Größe, das heißt, nach der Messung liegt der Zustand

$$\begin{aligned} |\Phi\rangle &= |\searrow\rangle \langle\searrow|\Psi\rangle \\ &= |\searrow\rangle \otimes (\alpha\gamma|\searrow\rangle + \alpha\delta|\nearrow\rangle). \end{aligned}$$

vor. Die Messung überführt folglich den verschränkten Zustand in einen Produktzustand. Das gilt in völliger Allgemeinheit: Messungen an Teilsystemen zusammengesetzter Systeme, die in verschränkten Zuständen vorliegen, überführen das gesamte System in einen Produktzustand. Damit treten nach der Messung die durch die Verschränkung verursachten Korrelationen nicht mehr auf.

- 6) Angenommen, ein verschränktes System mit dem Zustandsvektor

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|\psi_1\psi_2\rangle + |\varphi_1\varphi_2\rangle),$$

dessen konstituierende Teilsysteme sehr weit voneinander entfernt sind, soll zur Übermittlung von instantanen Botschaften verwendet werden. System 1 werde von Alice, System 2 von Bob beobachtet. Die Verschränkung des Systems bewirkt nun

zwar, daß jede von Alice durchgeführte Messung die einzelnen Resultate der durch Bob ausgeführten Messungen instantan beeinflusst; da Bob jedoch nicht weiß beziehungsweise maximal durch mit Lichtgeschwindigkeit übermittelte Nachrichten erfahren kann, was Alice gemessen hat, stellt er bei seinen Meßreihen unabhängig von Alices Messungen stets lediglich statistische Verteilungen fest. Das heißt, Bob merkt überhaupt nicht, ob Alice Messungen durchführt oder nicht. EPR-Korrelationen können die beiden erst feststellen, wenn sie anschließend ihre beiden Meßreihen vergleichen. Folglich ermöglichen verschränkte Systeme keine Informationsübertragung mit Überlichtgeschwindigkeit.

- 7a) Für ein beliebiges Ensemble von Systemen, die jeweils drei Eigenschaften  $A, B$  und  $C$  eindeutig aufweisen oder nicht aufweisen, gilt stets folgendes:

Die Anzahl der Systeme, welche die Eigenschaften  $A$  und  $B$  besitzen, ist höchstens so groß wie die Anzahl der Systeme, welche die Eigenschaften  $A$  und  $C$  besitzen plus die Anzahl der Systeme, welche die Eigenschaft  $A$  besitzen und die Eigenschaft  $C$  nicht besitzen.

Bezeichnet man die Anzahl der Systeme, welche die Eigenschaften  $X$  und  $Y$  besitzen, mit  $n(X, Y)$ , so läßt sich das in Gestalt der Ungleichung

$$n(A, B) \leq n(A, C) + n(A, \neg C)$$

formulieren – was nichts anderes als eine spezielle Form der Bellschen Ungleichung ist. Ist nun speziell  $A$  die Eigenschaft eines Photons, von einem Polarisationsfilter durchgelassen zu werden, der im Winkel  $\alpha$  eingestellt ist und  $B$  und  $C$  entsprechend für  $\beta$  und  $\gamma$ , so folgt mit der Schreibweise von Abschnitt 4.5.7.4 unmittelbar die Ungleichung

$$\mathfrak{P}_{\alpha\beta}(+1, +1) \leq \mathfrak{P}_{\alpha\gamma}(+1, +1) + \mathfrak{P}_{\beta\gamma}(+1, -1).$$

- b) Aus dieser Ungleichung erhält man für das beschriebene Polarisationsexperiment wie in Abschnitt 4.5.7.4

$$\cos^2(\alpha - \beta) \leq \cos^2(\alpha - \gamma) + 1 - \cos^2(\beta - \gamma)$$

und damit

$$\cos^2(\alpha - \beta) \leq \cos^2(\alpha - \gamma) + \sin^2(\beta - \gamma).$$

Wählt man speziell  $\alpha = 0^\circ, \beta = 30^\circ$  und  $\gamma = 60^\circ$ , so führt das auf

$$\frac{3}{4} \leq \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

und damit auf einen Widerspruch. Das bedeutet eine klare Verletzung der Bellschen Ungleichung.

- 8a) An der Wellenfunktion

$$|\Psi_{ABC}\rangle = \frac{1}{2} (|\uparrow_{2A}\downarrow_{2B}\uparrow_{1C}\rangle + |\downarrow_{2A}\uparrow_{2B}\uparrow_{1C}\rangle + |\uparrow_{2A}\uparrow_{2B}\downarrow_{1C}\rangle + |\downarrow_{2A}\downarrow_{2B}\downarrow_{1C}\rangle)$$

liest man unmittelbar  $s_{2A} s_{2B} s_{1C} = -1$  ab. Vertauschen der Bezeichnungen  $A, B$  und  $C$  liefert entsprechend  $s_{1A} s_{2B} s_{2C} = -1$  und  $s_{2A} s_{1B} s_{2C} = -1$ . Analog erhält man anhand der Wellenfunktion

$$|\Psi_{ABC}\rangle = \frac{1}{2} (|\uparrow_{1A}\uparrow_{1B}\uparrow_{1C}\rangle + |\uparrow_{1A}\downarrow_{1B}\downarrow_{1C}\rangle + |\downarrow_{1A}\uparrow_{1B}\downarrow_{1C}\rangle + |\downarrow_{1A}\downarrow_{1B}\uparrow_{1C}\rangle)$$

durch Ablesen  $s_{1A} s_{1B} s_{1C} = +1$ .

- b) Im Gegensatz zur Quantenmechanik, nach der die Größen  $s_{2A} s_{2B} s_{1C}$ ,  $s_{1A} s_{2B} s_{2C}$ ,  $s_{2A} s_{1B} s_{2C}$  einerseits und die Größe  $s_{1A} s_{1B} s_{1C}$  andererseits nicht gleichzeitig streuungsfrei meßbar sind, ist das in einer klassischen, lokalen Theorie mit verborgenen Variablen sehr wohl möglich; darüberhinaus müssen, da alle Eigenschaften der betrachteten Teilchen schon vor den Messungen und unabhängig von denselben feststehen, die Meßwerte für  $s_{1A} s_{1B} s_{1C}$  aus denjenigen für  $s_{2A} s_{2B} s_{1C}$ ,  $s_{1A} s_{2B} s_{2C}$  und  $s_{2A} s_{1B} s_{2C}$  ableitbar sein. Zunächst erhält man durch Multiplizieren der drei ersten Gleichungen aus Aufgabenteil a)

$$s_{2A} s_{2B} s_{1C} s_{1A} s_{2B} s_{2C} s_{2A} s_{1B} s_{2C} = s_{2A} s_{2A} s_{2B} s_{2B} s_{2C} s_{2C} s_{1A} s_{1B} s_{1C} = -1;$$

definitionsgemäß gilt außerdem

$$s_{2A} s_{2A} = s_{2B} s_{2B} = s_{2C} s_{2C} = 1,$$

und damit folgt

$$s_{1A} s_{1B} s_{1C} = -1.$$

- c) Bei dem Polarisationsexperiment zeigte sich der Unterschied zwischen den Vorhersagen klassischer lokaler Theorien mit verborgenen Variablen und denjenigen der Quantenmechanik erst nach Vergleich der statistischen Verteilung von Meßreihen. Beim hier diskutierten Experiment wird dieser Unterschied im Prinzip bereits an jeder einzelnen Messung deutlich.

# Literaturverzeichnis

- [1] *Bildungsplan Gymnasium*, Amtsblatt des Ministeriums für Kultus und Sport Baden Württemberg, Lehrplanheft 4/1994, Stuttgart 1994
- [2] *Bildungsplan für die Kursstufe des Gymnasiums*, Amtsblatt des Ministeriums für Kultus, Jugend und Sport Baden Württemberg, Lehrplanheft 3/2001, Stuttgart 2001
- [3] *Bildungsplan 2004 Allgemeinbildendes Gymnasium*, Ministerium für Kultus, Jugend und Sport Baden Württemberg, Stuttgart 2004
- [4] S. Albeverio und Z. Brzeźniak, *Journal of Functional Analysis* **113**, 177, 1993
- [5] S. Albeverio, P. Combe, R. Høegh-Krohn, G. Rideau und R. Stora (Hrsg.), *Feynman Path Integrals*, Springer, Berlin 1979
- [6] S. Albeverio, J. E. Fenstad, R. Høegh-Krohn und T Lindstrøm, *Nonstandard Methods in Stochastic Analysis and Mathematical Physics*, Dover, Minneola, New York 2009
- [7] S. Albeverio, G. Guatteri und S. Mazzucchi, *Journal of Mathematical Physics* **43**, 2847, 2002
- [8] S. A. Albeverio und R. J. Høegh-Krohn, *Inventiones mathematicae* **40**, 59, 1977
- [9] S. A. Albeverio und R. Høegh-Krohn, in [5], 3
- [10] S. A. Albeverio, R. J. Høegh-Krohn und S. Mazzucchi, *Mathematical Theory of Feynman Path Integrals. An Introduction*, Springer, Berlin, Heidelberg 2008
- [11] S. Albeverio, G. W. Johnson und Z. M. Ma, *Acta Applicandae Mathematicae* **42**, 267, 1996
- [12] S. Albeverio und S. Mazzucchi, in [50], 73
- [13] R. Allgöwer, *Stuttgarter Zeitung*, 18.10.2004
- [14] M. H. Andersen, J. R. Ensher, M. R. Matthews, C. E. Wieman und E. A. Cornell, *Science* **269**, 198, 1997
- [15] S. Antoci, D.-E. Liebscher, *Annales de la Fondation de Broglie* **21**, 349, 1996

- [16] H. Araki, *Mathematical Theory of Quantum Fields*, Oxford University Press, New York 2009
- [17] M. Arndt und O. Nairz, *Plus Lucis* 3/1999, 5
- [18] M. Arndt, O. Nairz, J. Voss-Andreae, C. Keller, G. van der Zouw und A. Zeilinger, *Nature* **401**, 680, 1999
- [19] N. Aronszajn, *Transaction of the American Mathematical Society* **68**, 337, 1950
- [20] N. Asai, I. Kubo und H.-H. Kuo, in [243], 17
- [21] A. Aspect, *Nature* **398**, 189, 1999
- [22] A. Aspect, in [54], 119
- [23] A. Aspect, P. Grangier und G. Roger, *Physical Review Letters* **47**, 460, 1981
- [24] A. Aspect, P. Grangier und G. Roger, *Physical Review Letters* **49**, 91, 1982
- [25] A. Aspect, J. Dalibard und G. Roger, *Physical Review Letters* **49**, 1804, 1982
- [26] A. Aspect und P. Grangier, in [405], 1
- [27] J. Audretsch, *Verschränkte Systeme. Die Quantenphysik auf neuen Wegen*, Wiley-VCH, Weinheim 2005
- [28] J. Audretsch und V. de Sabbata, *Quantum Mechanics in Curved Space-Time*, Plenum, New York, London 1990
- [29] G. Auletta, *Foundations and Interpretation of Quantum Mechanics. In the Light of a Critical-Historical Analysis and of a Synthesis of the Results*, World Scientific, Singapore, New Jersey, London, Hong Kong 2001
- [30] J. Azéma, P. A. Meyer und M. Yor (Hrsg.), *Séminaire de Probabilités XXV*, Springer, Berlin, Heidelberg 1991
- [31] D. G. Babbitt, *Journal of Mathematical Physics* **4**, 36, 1963
- [32] U. Backhaus und H. J. Schlichting, *Der mathematische und naturwissenschaftliche Unterricht* **34**, 153, 1981
- [33] U. Backhaus und H. J. Schlichting, *Der mathematische und naturwissenschaftliche Unterricht* **34**, 282, 1981
- [34] U. Backhaus und H. J. Schlichting, *Der Physikunterricht* **18/3**, 41, 1984
- [35] U. Backhaus und H. J. Schlichting, *physica didacta* **12/2**, 35, 1985
- [36] F. Bader, *Der mathematische und naturwissenschaftliche Unterricht* **30**, 278, 1977
- [37] F. Bader, *Eine Quantenwelt ohne Dualismus*, Schroedel, Hannover 1996

- [38] F. Bader (Hrsg.), *Dorn-Bader Physik Gymnasium Sekundarstufe II 12/13*, Schroedel, Hannover 2000
- [39] F. Bader (Hrsg.), *Dorn-Bader Physik Gymnasium Sekundarstufe II Kurstufe*, Schroedel, Hannover 2004
- [40] F. Bader, *Physik Journal* **8**, Nr. 8/9, 14, 2009
- [41] S. Banach und H. Steinhaus, *Fundamenta Mathematica* **9**, 50, 1927
- [42] J. D. Barrow, P. C. W. Davies und C. L. Harper jr. (Hrsg.), *Science and Ultimate Reality. Quantum Theory, Cosmology and Complexity*, Cambridge University Press, Cambridge 2005
- [43] L. S. Bartell, *Physical Review D* **21**, 1698, 1980, abgedruckt in [523], 455
- [44] K. Baumann und R. U. Sexl (Hrsg.), *Die Deutungen der Quantentheorie*, Vieweg, Braunschweig, Wiesbaden 1992
- [45] R. Bayer et al., *Knotenpunkte der Naturwissenschaften. Auf dem Weg zu einer gemeinsamen Fachsprache der Naturwissenschaften*, Landesinstitut für Schulentwicklung, Stuttgart 2006
- [46] G. Beck, *Zeitschrift für Physik* **41**, 443, 1927
- [47] J. S. Bell, *Physics* **1**, 195, 1964, abgedruckt in [49], 14 und in [523], 403
- [48] J. S. Bell, *Reviews of Modern Physics* **38**, 447, 1966, abgedruckt in [49], 1 und in [523], 397, deutsche Übersetzung in [44], 193
- [49] J. S. Bell, *Speakable and unspeakable in quantum mechanics*, Cambridge University Press, Cambridge 1993
- [50] F. E. Benth, G. Di Nunno, T. Lindstrøm, B. Øksendahl und T. Zhang, *Stochastic Analysis and Applications*, Springer, Berlin, Heidelberg 2007
- [51] A. Berg, H. Fischler, M. Lichtfeldt, M. Nitzsche, B. Richter und F. Walther, *Einführung in die Quantenphysik*, Pädagogisches Zentrum Berlin, Berlin 1989
- [52] S. Bergman, *The Kernel Function and Conformal Mapping*, American Mathematical Society, Providence, Rhode Island 1992
- [53] J. Bergou und B.-G. Englert, *Annals of Physics* **209**, 479, 1991
- [54] R. A. Bertlmann und A. Zeilinger (Hrsg.), *Quantum [Un]speakables. From Bell to Quantum Information*, Springer, Berlin, Heidelberg, New York 2002
- [55] G. Bimonte und R. Musto, [quant-ph/0208024](#)
- [56] G. Bimonte und R. Musto, [quant-ph/0301017](#)
- [57] G. Bimonte und R. Musto, [quant-ph/0304004](#)

- [58] A. Bivar-Weinholtz und M. L. Lapidus, Proceedings of the American Mathematical Society *110*, 449, 1990
- [59] S. Bochner, *Vorlesungen über Fouriersche Integrale*, Akademie-Verlag, Berlin 1932
- [60] S. Bochner und W. T. Martin, *Several Complex Variables*, Princeton University Press, Princeton 1948
- [61] G. Böhme (Hrsg.), *Theorie-Diskussion Protophysik. Für und wider eine konstruktive Wissenschaftstheorie der Physik*, Suhrkamp, Frankfurt am Main 1976
- [62] A. Bohm, *The Rigged Hilbert Space and Quantum Mechanics*, Springer, Berlin, Heidelberg, New York 1978
- [63] A. Bohm und M. Gadella, *Dirac Kets, Gamow Vectors and Gel'fand Triplets. The Rigged Hilbert Space Formulation of Quantum Mechanics*, Springer, Berlin, Heidelberg, New York 1989
- [64] N. Bohr, Philosophical Magazine **26**, 1, 1913
- [65] N. Bohr, Philosophical Magazine **26**, 476, 1913
- [66] N. Bohr, Philosophical Magazine **26**, 857, 1913
- [67] N. Bohr, Nature **121**, 580, 1928
- [68] N. Bohr, Naturwissenschaften **16**, 245, 1928
- [69] N. Bohr, Physical Review **48**, 696, 1935, abgedruckt in [523], deutsche Übersetzung in [44], 87
- [70] L. Boltzmann, Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften in Wien, Mathematisch-Naturwissenschaftliche Classe **66**, 275, 1872
- [71] L. Boltzmann, Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften in Wien, Mathematisch-Naturwissenschaftliche Classe **76**, 373, 1877
- [72] N. Bourbaki, *Eléments de Mathématiques, Volume VI*, Hermann, Paris 1969
- [73] , D. Bouwmeester, J.-W. Pan, M. Daniell, H. Weinfurter und A. Zeilinger, Physical Review Letters **82**, 1345, 1999
- [74] A. Brachner und R. Fichtner, *Quantenmechanik*, Schroedel, Hannover 1980
- [75] H.-P. Breuer und F. Petruccione, *The Theory of Open Quantum Systems*, Oxford University Press, Oxford 2007
- [76] B. Brezger, L. Hackermüller, S. Uttenthaler, J. Petschinka, M. Arndt und A. Zeilinger, Physical Review Letters **88**, 100404, 2002
- [77] K. Broderix, H. Leschke und P. Müller, Journal of Functional Analysis **212**, 287, 2004



- [78] L. V. de Broglie, *Nature* **112**, 540, 1923
- [79] L. V. de Broglie, *Comptes Rendus* **177**, 507, 1923
- [80] L. V. de Broglie, *Philosophical Magazine* **47**, 446, 1924
- [81] L. V. de Broglie, *Annales de Physique* **3**, 22, 1925
- [82] R. Brown, *Poggendorfer Annalen* **14**, 294
- [83] L. M. Brown (Hrsg.), *Feynman's Thesis. A New Approach to Quantum Theory*, World Scientific, New Jersey, London, Singapore 2005
- [84] J. Bub, *Interpreting the Quantum World*, Cambridge University Press, Cambridge 1997
- [85] W. Büchel, in [61], 235
- [86] M. Bunge (Hrsg.), *Quantum Theory and Reality*, Springer, Berlin, Heidelberg, New York 1967
- [87] P. Busch und P. J. Lahti, *Rivista del Nuovo Cimento* **18**, 1, 1995
- [88] J. Butterfield und T. Placek (Hrsg.), *Proceedings of the NATO Advanced Research workshop on Modality, Probability, and Bell's Theorem*, Kluwer, Dordrecht 2002
- [89] R. H. Cameron, *Journal of Mathematical and Physical Sciences* **39**, 126, 1960
- [90] R. H. Cameron und D. H. Storvick, *Journal of Mathematics and Mechanics* **18**, 517, 1968
- [91] R. H. Cameron und D. H. Storvick, *Journal d'Analyse Mathématique* **38**, 34, 1980
- [92] O. Carnal und J. Mlynek, *Physikalische Blätter* **47**, 379, 1991
- [93] O. Carnal und J. Mlynek, *Physical Review Letters* **66**, 2689, 1991
- [94] P. Cartier und C. DeWitt-Morette, *Journal of Mathematical Physics* **36**, 2137, 1995
- [95] P. Cartier und C. DeWitt-Morette, *Functional Integration: Action and Symmetries*, Cambridge University Press, Cambridge 2006
- [96] P. R. Chernoff, *Journal of Functional Analysis* **2**, 238, 1968
- [97] P. R. Chernoff, *Memoires of the American Mathematical Society* No. 140, 1, 1974
- [98] L. Chetouani und T. Hammann, *Journal of Mathematical Physics* **27**, 2944, 1986
- [99] J.S.R. Chisholm und A.K. Commons (Hrsg.), *Clifford Algebras and their Applications in Mathematical Physics*, Reidel, Dordrecht, Boston 1986
- [100] K. L. Chung, *Lectures from Markov Processes to Brownian Motion*, Springer, New York 1982

- [101] K. L. Chung und S. R. S. Varadhan, *Studia Mathematica* **68**, 249, 1979
- [102] J. F. Clauser, M. A. Horne, A. Shimony und R. A. Holt, *Physical Review Letters* **23**, 880, 1969, abgedruckt in [523], 409
- [103] R. Clausius, *Annalen der Physik* **79**, 368, 1850
- [104] R. Clausius, *Annalen der Physik* **79**, 500, 1850
- [105] R. Clifton, *Perspectives on Quantum Reality. Non-Relativistic, Relativistic and Field-Theoretic*, Kluwer, Dordrecht, Boston, London 1996
- [106] W. G. Cochran, H.-H. Kuo und A. Sengupta, *Infinite Dimensional Analysis, Quantum Probability and Related Topics* **1**, 43, 1998
- [107] D. L. Cohn, *Measure Theory*, Birkhäuser, Boston 1996
- [108] J. Conway und S. Kochen, in [54], 257
- [109] J. Cook, *Transactions of the American Mathematical Society* **74**, 222, 1953
- [110] M. Cunha, C. Drumond, P. Leukert, J. L. Silva und W. Westerkamp, *Annalen der Physik* **4**, 53, 1995
- [111] H. L. Cycon, R. G. Froese, W. Kirsch und B. Simon, *Schrödinger Operators with Applications to Quantum Mechanics and Global Geometry*, Springer, Berlin, Heidelberg, New York 2008
- [112] I. Daubechies und J. R. Klauder, *Journal of Mathematical Physics* **26**, 2239, 1985
- [113] K. J. Davey, *Problema in Applying Mathematics: On the inferential and representational Limits of Mathematics in Physics*, Dissertation, University of Pittsburgh, Pittsburgh 2003
- [114] E. Deotto und G. C. Ghirardi, *Foundations of Physics* **28**, 1, 1998
- [115] C. DeWitt-Morette, *Communications in Mathematical Physics* **28**, 47, 1972
- [116] C. DeWitt-Morette, *Communications in Mathematical Physics* **37**, 63, 1974
- [117] C. DeWitt-Morette, in [5], 227
- [118] C. DeWitt-Morette, in [28], 49
- [119] C. DeWitt-Morette, A. Masheshwari und B. Nelson, *General Relativity and Gravitation* **8**, 581, 1977
- [120] C. DeWitt-Morette, A. Masheshwari und B. Nelson, *Physics Reports* **50**, 266, 1979
- [121] P. A. M. Dirac, *Physikalische Zeitschrift der Sowjetunion* **3**, 64, 1933, abgedruckt in [83], 111
- [122] P. A. M. Dirac, *Review of Modern Physics* **17**, 195, 1945

- [123] P. A. M. Dirac, *The Principles of Quantum Mechanics*, Oxford University Press, Oxford 1995
- [124] A. A. diSessa, *American Journal of Physics* **48**, 365, 1980
- [125] M. D. Donsker und S. R. S. Varadhan, *Communications in Pure and Applied Mathematics* **28**, 525, 1975
- [126] S. Dürr, *Physical Review A* **64**, 042113, 2001
- [127] S. Dürr, T. Nonn und G. Rempe, *Nature* **395**, 33, 1998
- [128] S. Dürr, T. Nonn und G. Rempe, *Physical Review Letters* **26**, 5705, 1998
- [129] S. Dürr und G. Rempe, *Optics Communications* **179**, 323, 2000
- [130] K. J. Düsberg, in [61], 265
- [131] N. Dunford und J. T. Schwartz, *Linear Operators. Part I: General Theory*, Wiley, Hoboken 1988
- [132] I. H. Duru und H. Kleinert, *Physics Letters B* **84**, 30, 1979
- [133] I. H. Duru und H. Kleinert, *Fortschritte der Physik* **30**, 401, 1982
- [134] E. Dynkin, *Matematicheskii Sbornik* **25**, 155, 1949
- [135] D. Ebert, *Eichtheorien – Grundlage der Elementarteilchenphysik*, VCH, Weinheim 1989
- [136] A. Einstein, *Annalen der Physik* **17**, 132, 1905
- [137] A. Einstein, *Annalen der Physik* **17**, 549, 1905
- [138] A. Einstein, *Annalen der Physik* **19**, 371, 1906
- [139] A. Einstein, B. Podolsky und N. Rosen, *Physical Review* **47**, 777, 1935, abgedruckt in [523], deutsche Übersetzung in [44], 80
- [140] A. Elitzur und L. Vaidman, *Foundations of Physics* **23**, 987, 1993
- [141] D. Elworthy und A. Truman, *Annales de l' Institut Henri Poincaré (A) Physique théorique* **41**, 115, 1984
- [142] G. G. Emch, *Algebraic Methods in Statistical Mechanics and Quantum Field Theory*, Wiley-Interscience, New York, London, Sydney, Toronto 1972
- [143] R. Emden, *Nature* **141**, 908, 1938
- [144] K.-J. Engel und R. Nagel, *One-Parameter Semigroups for Linear Evolution Equations*, Springer, New York 2000
- [145] K.-J. Engel und R. Nagel, *A Short Course on Operator Semigroups*, Springer, New York 2006

- [146] B.-G. Englert, *Physical Review Letters* **77**, 2154, 1996
- [147] B.-G. Englert, *Zeitschrift für Naturforschung* **54a**, 11, 1998
- [148] B.-G. Englert, M. O. Scully, G. Süßmann und H. Walther, *Zeitschrift für Naturforschung* **47a**, 1175, 1992
- [149] B.-G. Englert und H. Walther, *Physik in unserer Zeit* **23**, 213, 1992
- [150] B.-G. Englert, H. Walther und M. O. Scully, *Applied Physics* **B 54**, 366, 1992
- [151] B.-G. Englert, M. O. Scully und H. Walther, *Spektrum der Wissenschaft* 2/1995, 50
- [152] C. P. Enz und J. Mehra, *Physical Reality and Mathematical Description*, Reidel, Dordrecht 1974
- [153] B. d'Espagnat, *Conceptual Foundations of Quantum Mechanics*, Addison-Wesley, Reading, Menlo Park, New York 1989
- [154] P. Exner, *Open Quantum Systems and Feynman Integrals*, Kluwer, Dordrecht 1984
- [155] G. Falk, *Theoretische Physik auf der Grundlage einer allgemeinen Dynamik. Band I*, Springer, Heidelberg 1968
- [156] G. Falk, *Theoretische Physik auf der Grundlage einer allgemeinen Dynamik. Band II*, Springer, Heidelberg 1968
- [157] G. Falk, in [158], 5
- [158] G. Falk und F. Herrmann (Hrsg.), *Konzepte eines zeitgemäßen Physikunterrichtes – Heft 3*, Schroedel, Hannover 1979
- [159] G. Falk und F. Herrmann (Hrsg.), *Konzepte eines zeitgemäßen Physikunterrichtes – Heft 5*, Schroedel, Hannover 1982
- [160] G. Falk, F. Herrmann und G. B. Schmid, *American Journal of Physics* **51**, 1074, 1983
- [161] M. de Faria, M. J. Oliveira und L. Streit, *Journal of Mathematical Physics* **46**, 063505, 2005
- [162] M. de Faria, J. Potthoff und L. Streit, *Journal of Mathematical Physics* **32**, 2123, 1991
- [163] J. Feldman, *Transactions of the American Mathematical Society* **108**, 251, 1963
- [164] R. P. Feynman, *The Principle of Least Action in Quantum Mechanics*, Dissertation, Princeton University, Princeton 1942, abgedruckt in [83], 1
- [165] R. P. Feynman, *Review on Modern Physics* **20**, 367, 1948, abgedruckt in [83], 71
- [166] R. P. Feynman, *Physical Review* **76**, 769, 1949

- [167] R. P. Feynman, *QED – Die seltsame Theorie des Lichts und der Materie*, Piper, München 1994
- [168] R. P. Feynman und A. R. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York 1965
- [169] R. P. Feynman, R. B. Leighton und M. Sands, *Feynman Vorlesungen über Physik. Band III: Quantenmechanik*, Oldenbourg, München, Wien 2007
- [170] H. Fischler, *Praxis der Naturwissenschaften - Physik* **41**, 13, 1992
- [171] H. Fischler (Hrsg.), *Quantenphysik in der Schule*, IPN, Kiel 1992
- [172] H. Fischler, in [171], 245
- [173] M. E. Fisher, *Archive for Rational Mechanics and Analysis* **17**, 377, 1964
- [174] V. Fock, *Zeitschrift für Physik* **75**, 622, 1932
- [175] J. Franck und G. Hertz, *Verhandlungen der Deutschen Physikalischen Gesellschaft* **16**, 457, 1914
- [176] J. Franck und G. Hertz, *Verhandlungen der Deutschen Physikalischen Gesellschaft* **16**, 512, 1914
- [177] G. Fubini, *Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche, Matematiche e Naturali. Roma* **16**, 608, 1907
- [178] J. C. Garrison und J. Wong, *Journal of Mathematical Physics* **11**, 2242, 1970
- [179] B. Gaveau, T. Jacobson, M. Kac und L. S. Schulman, *Physical Review Letters* **53**, 419, 1984
- [180] B. Gaveau und L. S. Schulman, *Physical Review D* **36**, 1135, 1987
- [181] B. Gaveau und L. S. Schulman, *Il Nuovo Cimento* **11 D**, 31, 1989
- [182] I. M. Gelfand und A. M. Yaglom, *Journal of Mathematical Physics* **1**, 48, 1960
- [183] I. M. Gelfand und G. E. Schilow, *Verallgemeinerte Funktionen, Band I*, Deutscher Verlag der Wissenschaften, Berlin 1960
- [184] I. M. Gelfand und G. E. Schilow, *Verallgemeinerte Funktionen, Band II*, Deutscher Verlag der Wissenschaften, Berlin 1962
- [185] I. M. Gelfand und G. E. Schilow, *Verallgemeinerte Funktionen, Band III*, Deutscher Verlag der Wissenschaften, Berlin 1964
- [186] I. M. Gelfand und N. J. Wilenkin, *Verallgemeinerte Funktionen, Band IV*, Deutscher Verlag der Wissenschaften, Berlin 1967
- [187] H. Genz, *Gedankenexperimente*, Wiley-VCH, Weinheim 1999

- [188] P. Ghose, [quant-ph/0003037v3](#)
- [189] P. Ghose, [quant-ph/0103126v8](#)
- [190] J. W. Gibbs, *Elementary Principles in Statistical Mechanics*, Charles Scribner's Sons, New York 1902
- [191] R. J. Glauber, *Physical Review* **131**, 2766, 1963
- [192] J. Glimm, in [289], 227
- [193] J. Glimm und A. Jaffe, *Quantum Physics. A Functional Integral Point of View*, Springer, Berlin, Heidelberg, New York 1987
- [194] T. Görnitz, *Quanten sind anders. Die verborgene Einheit der Welt*, Spektrum Akademischer Verlag, Heidelberg, Berlin 1999
- [195] I. Gohberg und M. Kac (Hrsg.), *Topics in Functional Analysis*, Academic Press, New York 1978
- [196] H. Grassmann, *Lineale Ausdehnungslehre oder die Wissenschaft der extensiven Grösse, eine neue mathematische Disciplin*, Verlag Otto Wiegand, Leipzig 1844
- [197] H. Grassmann, *Die Ausdehnungslehre. Vollständig und in strenger Form*, Enslin, Berlin 1862
- [198] D. M. Greenberger, M. A. Horne, A. Shimony und A. Zeilinger, *American Journal of Physics* **58**, 1131, 1990
- [199] D. M. Greenberger, M. A. Horne und A. Zeilinger, in [496], 69
- [200] D. M. Greenberger, M. A. Horne und A. Zeilinger, *American Journal of Physics* **58**, 1131, 1990
- [201] D. M. Greenberger und A. YaSin, *Physics Letters A* **128**, 391, 1988
- [202] W. Greiner und J. Reinhardt, *Theoretische Physik. Band 7A: Feldquantisierung*, Harri Deutsch, Thun, Frankfurt am main 1993
- [203] C. Grosche und F. Steiner, *Journal of Mathematical Physics* **36**, 2355, 1995
- [204] S. Großmann, *Funktionalanalysis im Hinblick auf Anwendungen in der Physik*, Aula, Wiesbaden 1988
- [205] M. Grothaus, D. C. Khandekar, J. L. Silva und L. Streit, *Journal of Mathematical Physics* **38**, 3278, 1997
- [206] M. Grothaus, L. Streit und A. Vogel, [math-ph/0805.3253v1](#)
- [207] S. J. Gustafson und I. M. Sigal, *Mathematical Concepts of Quantum Mechanics*, Springer, Berlin, Heidelberg, New York 2003

- [208] R. Haag, *Local Quantum Physics. Fields, Particles, Algebras*, Springer, Berlin, Heidelberg, New York 1996
- [209] L. Hackermüller, S. Uttenthaler, K. Hornberger, E.Reiger, B. Brezger, A. Zeilinger und M. Arndt, *Physical Review Letters* **91**, 090408, 2003
- [210] M. J. W. Hall, *Journal of Physics A* **37**, 9549, 2004
- [211] H. Halvorson, *Philosophy of Science* **68**, 111, 2001
- [212] L. Hardy, *Physical Review Letters* **68**, 2981, 1992
- [213] L. Hardy, [quant-ph/0101012](#)
- [214] L. Hardy, in [88], 61
- [215] L. Hardy, in [42], 45
- [216] B. O. Haugsby, *An Operator Valued Integral in a Function Space of Continuous Vector Valued Functions*, Dissertation, University of Minnesota, Minneapolis 1972
- [217] H. Hauptmann, *Der Karlsruher Physikkurs für die Sekundarstufe II. Unterrichtskonzepte zur Rotationsmechanik, Elektrodynamik, Schwingungslehre und Physik der Sonne*, Dissertation, Universität Karlsruhe, Karlsruhe 2001
- [218] G. Heiduck, F. Herrmann und G. B. Schmid, *European Journal of Physics* **8**, 41, 1987
- [219] W. Heisenberg, *Zeitschrift für Physik* **43**, 172, 1927
- [220] C. Held, *Die Bohr-Einstein-Debatte. Quantenmechanik und physikalische Wirklichkeit*, Schöningh, Paderborn, München, Wien, Zürich 1998
- [221] M. Henneaux und C. Teitelboim, *Quantization of Gauge Systems*, Princeton University Press, Princeton 1992
- [222] F. Herrmann, in [159], 26
- [223] F. Herrmann, *American Journal of Physics* **57**,707, 1989
- [224] F. Herrmann, *Praxis der Naturwissenschaften – Physik* **44**, 41, 1995
- [225] F. Herrmann, *Karlsruher Physikkurs S I, Band 1*, Aulis, Köln 1998
- [226] F. Herrmann, *Karlsruher Physikkurs S I, Band 2*, Aulis, Köln 1998
- [227] F. Herrmann, *Karlsruher Physikkurs S I, Band 3*, Aulis, Köln 1998
- [228] F. Herrmann, *European Journal of Physics* **21**, 49, 2000
- [229] F. Herrmann und H. Hauptmann, *Karlsruher Physikkurs S II, Band 1*, Aulis, Köln 2008



- [230] F. Herrmann und H. Hauptmann, *Karlsruher Physikkurs S II, Band 2*, Aulis, Köln 2008
- [231] F. Herrmann und H. Hauptmann, *Karlsruher Physikkurs S II, Band 3*, Aulis, Köln 2008
- [232] F. Herrmann und H. Hauptmann, *Karlsruher Physikkurs S II, Band 4*, Aulis, Köln 2008
- [233] F. Herrmann und G. Job, *Altlasten der Physik*, Aulis, Köln 2002
- [234] F. Herrmann und P. Schmälzle, Der mathematische und naturwissenschaftliche Unterricht **37**, 476, 1984
- [235] W. Herzog, Der mathematische und naturwissenschaftliche Unterricht **60**, 500, 2007
- [236] D. Hestenes, in [99], 1
- [237] D. Hestenes, in [463], 191
- [238] D. Hestenes und G. Sobczyk, *Clifford Algebra to Geometric Calculus*, Kluwer, Dordrecht 1992
- [239] H. Heuser, *Funktionalanalysis. Theorie und Anwendung*, Teubner, Stuttgart 1986
- [240] S. Heusler und H. J. Schlichting, Physik und Didaktik in Schule und Hochschule **1**, 9, 2008
- [241] T. Hida, Journal of the Korean Mathematical Society **38**, 275, 2001
- [242] T. Hida, H.-H. Kuo, J. Potthoff und L. Streit, *White Noise. An Infinite Dimensional Calculus*, Kluwer, Dordrecht 1995
- [243] T. Hida und K. Saito (Hrsg.), *Quantum Information II*, World Scientific, Singapore 2000
- [244] T. Hida und K. Saito (Hrsg.), *Quantum Information IV*, World Scientific, Singapore 2002
- [245] T. Hida und L. Streit, Stochastic Processes and their Applications **16**, 55, 1983
- [246] D. Hilbert, *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen*, Teubner, Leipzig, Berlin 1912
- [247] E. Hille und R. S. Phillips, *Functional analysis and semi-groups*, American Mathematical Society Colloquium Publications, volume 31, Providence 1957
- [248] A. Hillmer, Hamburger Abendblatt, 19./20.5.2004
- [249] F. Hirzebruch und W. Scharlau, *Einführung in die Funktionalanalysis*, Spektrum Akademischer Verlag, Heidelberg, Berlin, Oxford 1996

- [250] P. D. Hislop und I. M. Sigal, *Introduction to Spectral Theory. With Applications to Schrödinger Operators*, Springer, New York, Berlin, Heidelberg 1996
- [251] R. Ho und A. Inomata, *Physical Review Letters* **48**, 231, 1982
- [252] L. Hörmander, *Acta Mathematica* **127(1)**, 79, 1971
- [253] C. A. Hooker (Hrsg.), *The Locigo-Algebraic Approach to Quantum Mechanics, Vol. 1*, Reidel, Dordrecht 1975
- [254] L. Van Hove, *Physica* **15**, 951, 1949
- [255] A. Huber, *Die Computerprogramme Polfilter und Quantum Eraser*, Zulassungsarbeit, Ludwigs-Maximilians-Universität München, München 1999
- [256] T. Ichinose, *Proceedings of the Japan Academy, Series A, Mathematical Sciences* **58**, 290, 1982
- [257] T. Ichinose, *Duke Mathematical Journal* **51**, 1, 1984
- [258] T. Ichinose, *Sugaku Expositions* **6**, 13, 1993
- [259] T. Ichinose und H. Tamura, *Journal of Mathematical Physics* **25**, 1810, 1984
- [260] T. Ichinose und H. Tamura, *Progress in Theoretical Physics Supplements* **92**, 144, 1987
- [261] T. Ichinose und H. Tamura, *Journal of Mathematical Physics* **29**, 103, 1988
- [262] R. Inhetveen, *Konstruktive Geometrie. Eine formentheoretische Begründung der euklidischen Geometrie*, Bibliographisches Institut, Mannheim 1983
- [263] A. Inomata, *Physics Letters A* **87**, 387, 1982
- [264] A. Inomata, *Physics Letters A* **101**, 253, 1984
- [265] K. Itô, *Journal of the Mathematical Society of Japan* **3**, 157, 1951
- [266] K. Itô, in [394], 227
- [267] K. Itô, in [344], 145
- [268] K. Itô und H. P. McKean, *Diffusion Processes and their Sample Paths*, Springer, Berlin, Heidelberg, New York 1996
- [269] C. Itzykson und J.-B. Zuber, *Quantum Field Theory*, McGraw-Hill, Singapore 1985
- [270] M. Jäger, *Die Philosophie des Konstruktivismus auf dem Hintergrund des Konstruktionsbegriffs*, Olms, Hildesheim, Zürich, New York 1998
- [271] P. Janich, *Zweck und Methode der Physik aus philosophischer Sicht*, Universitätsverlag, Konstanz 1973

- [272] P. Janich, in [61], 83
- [273] P. Janich, *Das Maß der Dinge. Protophysik von Raum, Zeit und Materie*, Suhrkamp, Frankfurt am Main 1997
- [274] W. Janke und A. Pelster (Hrsg.), *Path Integrals. New Trends and Perspectives*, World Scientific, Singapore, New Jersey, London, Hong Kong 2008
- [275] J. M. Jauch, *Foundations of Quantum Mechanics*, Addison-Wesley, Reading, 1977
- [276] E. T. Jaynes, *Physical Review* **106**, 620, 1957
- [277] B. Jefferies, *Evolution Processes and the Feynman-Kac-Formula*, Kluwer, Dordrecht 1996
- [278] U. C. Ji und N. Obata, *Infinite Dimensional Analysis, Quantum Probability and Related Topics* **6**, 167, 2003
- [279] C. Jönsson, *Zeitschrift für Physik A* **161**, 454, 1961
- [280] T. R. Johansen, *The Bochner-Minlos Theorem for nuclear spaces and an abstract white noise space*, Louisiana State University, Baton Rouge 2003
- [281] G. W. Johnson, *Journal of Mathematical Physics* **23**, 2090, 1982
- [282] G. W. Johnson, *Existence Theorems for the Analytic Operator-Valued Feynman Integral*, University of Sherbrooke Monograph, No. 20, Sherbrooke, Québec 1988
- [283] G. W. Johnson und M. L. Lapidus, *Generalized Dyson Series, generalized Feynman Diagrams, the Feynman integral, and Feynman's operational calculus*, American Mathematical Society, Providence, Rhode Island 1986
- [284] G. W. Johnson und M. L. Lapidus, *The Feynman Integral and Feynman's Operational Calculus*, Clarendon, Oxford 2002
- [285] G. W. Johnson und D. L. Skoug, *Nagoya Mathematical Journal* **60**, 93, 1976
- [286] E. Joos, in [287], 41
- [287] E. Joos, H. D. Zeh, C. Kiefer, D. Giulini, J. Kupsch und I.-O. Stamatescu, *Decoherence and the Appearance of a Classical World in Quantum Theory*, Springer, Berlin, Heidelberg, New York 2003
- [288] J. D. Jost, J. P. Home, J. M. Amini, D. Hanneke, R. Ozeri, C. Langer, J. J. Bollinger, D. Leibfried und D. J. Wineland, *Nature* **459**, 683, 2009
- [289] R. Jost, *Local Quantum Theory*, Academic Press, New York, London 1969
- [290] M. Kac, *Transactions of the American Mathematical Society* **65**, 1, 1949
- [291] M. Kac, in [393], 189

- [292] M. Kac, *Probability and Related Topics in the Physical Sciences*, Wiley, New York 1959
- [293] M. Kafatos (Hrsg.), *Bell Theorem, Quantum Theory and Conceptions of the Universe*, Kluwer, Dordrecht 1989
- [294] M. Kaku, *Quantum Field Theory – A Modern Introduction*, Oxford University Press, New York, Oxford 1993
- [295] G. Kallianpur und C. Bromley, in [407], 217
- [296] G. Kallianpur, D. Kannan und R. L. Karandikar, *Annales de l'Institute Henri Poincaré* **21**, 323, 1985
- [297] B. Kanitscheider (Hrsg.), *Moderne Naturphilosophie*, Königshausen & Neumann, Würzburg 1984
- [298] I. Karatzas und S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer, Berlin 2004
- [299] T. Kato, *Transactions of the American Mathematical Society* **70**, 195, 1951
- [300] T. Kato, *Proceedings of the Japan Academy* **35**, 467, 1959
- [301] T. Kato, *Israel Journal of Mathematics* **13**, 135, 1973
- [302] T. Kato, in [152], 193
- [303] T. Kato, in [195], 185
- [304] T. Kato, *Perturbation Theory for Linear Operators*, Springer, Berlin, Heidelberg 1995
- [305] T. Kato und K. Masuda, *Journal of the Mathematical Society of Japan* **30**, 169, 1978
- [306] M. A. Kayed und A. Inomata, *Physical Review Letters* **53**, 107, 1984
- [307] J. R. Klauder, *Physical Review D* **19**, 2349, 1979
- [308] J. R. Klauder und I. Daubechies, *Physical Review Letters* **52**, 1161, 1984
- [309] J. R. Klauder und E. C. G. Sudarshan, *Fundamentals of Quantum Optics*, Dover, Minneola, New York 2006
- [310] H. Kleinert, *Physics Letters A* **120**, 361, 1987
- [311] H. Kleinert, *Modern Physics Letters A* **4**, 2329, 1989
- [312] H. Kleinert, *Physics Letters B* **236**, 315, 1990
- [313] H. Kleinert, *Pfadintegrale in Quantenmechanik, Statistik und Polymerphysik*, Bibliographisches Institut, Mannheim, Leipzig, Wien Zürich 1993

- [314] M. Klohr, Ludwigsburger Kreiszeitung, 6.10.2004
- [315] S. Kochen, in [105], 237
- [316] S. Kochen und E. Specker, Journal of Mathematics and Mechanics **17**, 59, 1967, abgedruckt in [253], 293
- [317] V. Kolokoltsov, Methodology and Computing in Applied Probability **1**, 349, 1999
- [318] Y. G. Kondratiev, P. Leukert, J. Pothoff, L. Streit und W. Westerkamp, Journal of Functional Analysis **141**, 301, 1996
- [319] Y. G. Kondratiev und L. Streit, Ukrainian Mathematical Journal **44**, 832, 1992
- [320] Y. G. Kondratiev und L. Streit, Reports on Mathematical Physics **33**, 341, 1993
- [321] I. Kubo und S. Takenaka, Proceedings of the Japan Academy **56A**, 376, 1980
- [322] I. Kubo und S. Takenaka, Proceedings of the Japan Academy **56A**, 411, 1980
- [323] I. Kubo und S. Takenaka, Proceedings of the Japan Academy **57A**, 433, 1981
- [324] I. Kubo und S. Takenaka, Proceedings of the Japan Academy **58A**, 186, 1982
- [325] J. Küblbeck, *Modellbildung in der Physik am Beispiel der Quantenphysik, ein Unterrichtsversuch im Grundkurs Physik der Klasse 13*, Staatliches Seminar für Schulpädagogik, Stuttgart 1995
- [326] J. Küblbeck, *Modellbildung in der Physik insbesondere in der Quantenphysik mit einem Unterrichtsvorschlag für Grund- und Leistungskurs*, Heft PH22, Landesinstitut für Erziehung und Unterricht, Stuttgart 1997
- [327] J. Küblbeck, Praxis der Naturwissenschaften - Physik **8/49**, 22, 2000
- [328] J. Küblbeck und R. Müller, *Die Wesenszüge der Quantenphysik. Modelle, Bilder, Experimente*, Aulis, Köln 2002
- [329] F. Kühnemund und W. Wacker, Semigroup Forum **60**, 478
- [330] W. Kuhn (Hrsg.), *Physik 2*, Westermann, Braunschweig 2001
- [331] W. Kuhn und J. Strnad, *Quantenfeldtheorie. Photonen und ihre Deutung*, Vieweg, Braunschweig, Wiesbaden 1995
- [332] T. Kuna, L. Streit und W. Westerkamp, Journal of Mathematical Physics **39**, 4476, 1998
- [333] H.-H. Kuo, *White Noise Distribution Theory*, CRC Press, Boca Raton, New York, London, Tokyo 1996
- [334] H.-H. Kuo, in [244], 1

- [335] H.-H. Kuo. *Gaussian Measures in Banach Spaces*, Springer, Berlin, Heidelberg, New York 2006
- [336] H.-H. Kuo, *Introduction to Stochastic Integration*, Springer, New York 2006
- [337] J. Kupsch, in [287], 317
- [338] T. G. Kurtz und M. Pierre, *Journal of Differential Equations* **52**, 407, 1984
- [339] P. Kustaanheimo und E. Stiefel, *Journal für Reine und Angewandte Mathematik* **218**, 204, 1965
- [340] L. D. Landau und R. Peierls, *Zeitschrift für Physik* **69**, 56, 1931
- [341] M. L. Lapidus, *Journal of Functional Analysis* **63**, 261, 1985
- [342] A. Laschek, P. Leukert, L. Streit und W. Westerkamp, *Reports on Mathematical Physics* **33**, 221, 1993
- [343] D. Laugwitz, *Zahlen und Kontinuum. Eine Einführung in die Infinitesimalmathematik*, B. I. Wissenschaftsverlag, Mannheim, Leipzig, Wien, Zürich 1994
- [344] L. M. Le Cam und J. Neyman (Hrsg.), *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 2: Contributions to Probability Theory, Part 1*, University of California Press, Berkeley 1967
- [345] J. Leisen, D. Fries und J. Luggen-Hölscher, *Energie und Entropie. Handreichung zum neuen Lehrplan Physik in der S II*, Pädagogisches Zentrum Rheinland-Pfalz, Bad Kreuznach 2000
- [346] K. Loo, *Journal of Mathematical Physics* **40**, 64, 1999
- [347] K. Loo, *Journal of Physics A* **33**, 9205, 2000
- [348] K. Loo, *Journal of Physics A* **33**, 9215, 2000
- [349] P. Lorenzen, in [61], 40
- [350] P. Lorenzen, *Grundbegriffe technischer und politischer Kultur. Zwölf Beiträge*, Suhrkamp, Frankfurt am Main 1985
- [351] G. Ludwig, *Einführung in die Grundlagen der Theoretischen Physik. Band 3: Quantentheorie*, Vieweg, Braunschweig 1976
- [352] G. Ludwig, *An Axiomatic Basis for Quantum Mechanics. Volume 1: Derivation of Hilbert Space Structure*, Springer, Berlin, Heidelberg, New York 1985
- [353] G. Ludwig, *An Axiomatic Basis for Quantum Mechanics. Volume 2: Quantum Mechanics of Macrosystems*, Springer, Berlin, Heidelberg, New York 1987
- [354] B. Lukić, S. Jeney, C. Tischer, A. J. Kulik, L. Forró und E.-L. Florin, *Physical Review Letters* **95**, 160601, 2005

- [355] G. W. Mackey, *The Mathematical Foundations of Quantum Mechanics*, W. A. Benjamin, New York, Amsterdam 1963
- [356] L. Mandel, *Optics Letters* **16**, 1882, 1991
- [357] H. Margenau und L. Cohen, in [86], 71
- [358] J. Martinez-Linares und D. A. Harmin, [quant-ph/0306057](#)
- [359] J. Martinez-Linares und J. Vargas Medina, [quant-ph/0309006](#)
- [360] V. P. Maslov, *Méthodes opératorielles*, Edition MIR, Moskau 1987
- [361] V. P. Maslov und A. M. Chebotarev, in [5], 58
- [362] K. Maurin, *General Eigenfunction Expansions and Unitary Representations of Topological Groups*, Polish Scientific Publishers, Warschau 1968
- [363] H. P. McKean, *Journal of Mathematical Physics* **18**, 1277, 1977
- [364] E. Merzbacher und D. Park, *American Journal of Physics* **60**, 946, 1992
- [365] P.-A. Meyer und J.-A. Yan, in [30], 61
- [366] G. Mie, *Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Classe der Kaiserlichen Akademie der Wissenschaften* **107**, 1113, 1898
- [367] R. A. Minlos, *Selected Translations in Mathematics, Statistics and Probability* **3**, 291, 1963
- [368] M. M. Mizrahi, *Journal of Mathematical Physics* **17**,
- [369] M. M. Mizrahi, *Journal of Mathematical Physics* **19**, 298, 1978
- [370] C. Monroe, D. M. Meekhof, B. E. King und D. J. Wineland, *Science* **272**, 1131, 1996
- [371] P. Del Moral, *Feynman-Kac Formulae. Genealogical and Interacting Particle Systems with Applications*, Springer, New York 2004
- [372] H. G. J. Moseley, *Philosophical Magazine* **26**, 1025, 1913
- [373] H. G. J. Moseley, *Nature* **92**, 554, 1913
- [374] H. G. J. Moseley, *Philosophical Magazine* **27**, 703, 1914
- [375] H. G. J. Moseley, *British Association Reports* 299, 1914
- [376] Heinz Muckenfuß, *Lernen im sinnstiftenden Kontext. Entwurf einer zeitgemäßen Didaktik des Physikunterrichts*, Cornelsen, Berlin 1995
- [377] R. Müller, *Quantenphysik in der Schule*, Logos, Berlin 2003
- [378] V. F. Müller, *Quantenmechanik*, Oldenbourg, München 2000



- [379] H. Müller-Ebhardt, H. Rehbein, R. Schnabel, K. Danzmann und Y. Chen, *Physical Review Letters* **100**, 013601, 2008
- [380] M. Nagasawa, *Schrödinger Equations and Diffusion Theory*, Birkhäuser, Basel, Boston, Berlin 1993
- [381] O. Nairz, M. Arndt und A. Zeilinger, *Journal of Modern Optics* **47**, 2811, 2000
- [382] T. Nakamura, *Journal of Mathematical Physics* **32**, 457, 1991
- [383] T. Nakamura, *Journal of Mathematical Physics* **38**, 4052, 1997
- [384] T. Nakamura, *Journal of Mathematical Physics* **41**, 5209, 2000
- [385] M. Namiki, S. Pascazio und H. Nakazato, *Decoherence and Quantum Measurements*, World Scientific, Singapore, New Jersey, London, Hong Kong 1999
- [386] E. Nelson, *Journal of Mathematical Physics* **5**, 332, 1964
- [387] B. Nelson und B. Sheeks, *Communications in Mathematical Physics* **84**, 515, 1982
- [388] W. Nernst, *Nachrichten der Königlichen Gesellschaft der Wissenschaften zu Göttingen* **1**, 1, 1906
- [389] W. Nernst, *Berichte der Königlich Preussischen Akademie der Wissenschaften* **52**, 933, 1906
- [390] A. Neumaier, [quant-ph/0001011v2](#)
- [391] A. Neumaier, [quant-ph/0706.0155v3](#)
- [392] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*, Berlin, Heidelberg, New York 1996
- [393] J. Neyman (Hrsg.), *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, Berkeley 1951
- [394] J. Neyman (Hrsg.), *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume II: Contributions to Probability Theory*, University of California Press, Berkeley, Los Angeles 1961
- [395] O. Nikodým, *Fundamenta Mathematicae* **15**, 31, 1930
- [396] O. M. Nikodým, *The Mathematical Apparatus for Quantum-Theories. Based on the Theory of Boolean Lattices*, Springer, Berlin, Heidelberg, New York 1966
- [397] E. Noether, *Nachrichten der Königlichen Gesellschaft der Wissenschaften zu Göttingen, mathematisch-physikalische Klasse*, 235, 1918
- [398] N. Obata, *White Noise Calculus and Fock Space*, Berlin, New York 2008

- [399] M. F. Ong, *Die Feynman-Kac-Formel für unbeschränkte Potentiale und allgemeine Anfangsbedingungen*, Dissertation, Technische Universität Kaiserslautern, Kaiserslautern 2004
- [400] R. Opitz, *Didaktische Aspekte der physikalischen Fachsprache*, Dissertation, Universität-Gesamthochschule Duisburg, Duisburg 2002
- [401] K. Osterwalder und R. Schrader, *Communications in Mathematical Physics* **31**, 83, 1973
- [402] J.-W. Pan, D. Bouwmeester, M. Daniell, H. Weinfurter und A. Zeilinger, *Nature* **404**, 515, 2000
- [403] J.-W. Pan, M. Daniell, S. Gasparoni, G. Weihs und A. Zeilinger, *Physical Review Letters* **86**, 4435, 2001
- [404] J.-W. Pan und A. Zeilinger, in [54], 225
- [405] R. Penrose and C. J. Isham (Hrsg.), *Quantum Concepts in Space and Time*, Oxford University Press, Oxford 1986
- [406] A. Perelomov, *Generalized Coherent States and their Applications*, Springer, Berlin, Heidelberg, New York 1986
- [407] M. A. Pinsky (Hrsg.), *Stochastic Analysis and Applications*, Marcel Dekker, New York 1984
- [408] C. Piron, *Foundations of Quantum Physics*, Benjamin, New York 1976
- [409] M. Planck, *Verhandlungen der Deutschen Physikalischen Gesellschaft* **2**, 202, 1900
- [410] M. Planck, *Physikalische Zeitschrift* **9**, 828, 1908
- [411] M. Planck, *Vorlesungen über Thermodynamik*, Veit, Leipzig 1917
- [412] D. Plappert, in [158], 59
- [413] D. Plappert, *Alles strömt!?*, Staatliches Seminar für Schulpädagogik (Gymnasien) Freiburg, Freiburg 2002
- [414] D. Plappert, *Praxis der Naturwissenschaften – Physik* **7/52**, 2, 2003
- [415] D. Plappert, *Praxis der Naturwissenschaften – Physik* **4/53**, 14, 2004
- [416] D. Plappert, *Praxis der Naturwissenschaften – Physik* **1/54**, 6, 2005
- [417] D. Plappert, *Praxis der Naturwissenschaften – Physik* **3/54**, 13, 2005
- [418] D. Plappert, *Praxis der Naturwissenschaften – Physik* **6/55**, 5, 2006
- [419] J. Potthoff und L. Streit, *Journal of Functional Analysis* **101**, 212, 1991
- [420] E. M. Purcell, *Nature* **178**, 1449, 1956

- [421] R. Quezada, *Universitatis Iagellonicae Acta Mathematica*, Fasciculus XXX, 69, 1993
- [422] H. Rauch, W. Treimer und U. Bonse, *Physics Letters* **47 A**, 369, 1974
- [423] R. Rebolledo, J. Rezende und J.-C. Zambrini, *Stochastic Analysis and Mathematical Physics*, World Scientific, Singapore 2002
- [424] M. Redhead, *Incompleteness, Nonlocality, and Realism. A Prolegomenon to the Philosophy of Quantum Mechanics*, Clarendon, Oxford 1992
- [425] M. Redhead, *Foundations of Physics* **25**, 123, 1995
- [426] M. Reed und B. Simon, *Methods of Modern Mathematical Physics. I Functional Analysis*, Academic Press, New York 1980
- [427] M. Reed und B. Simon, *Methods of Modern Mathematical Physics. II Fourier Analysis, Self-Adjointness*, Academic Press, New York 1975
- [428] M. Reed und B. Simon, *Methods of Modern Mathematical Physics. III Scattering Theory*, Academic Press, New York 1970
- [429] M. Reed und B. Simon, *Methods of Modern Mathematical Physics. IV Analysis of Operators*, Academic Press, New York 1978
- [430] H. Reeh und S. Schlieder, *Il Nuovo Cimento* **22**, 1051, 1961
- [431] J. Rezende, *Communications in Mathematical Physics* **101**, 187, 1985
- [432] O. W. Richardson, *Philosophical Magazine* **23**, 594, 1912
- [433] O. W. Richardson, *Philosophical Magazine* **24**, 570, 1912
- [434] H. P. Robertson, *Physical Review* **34**, 163, 1929
- [435] A. Robinson, *Non-standard Analysis*, Princeton University Press, Princeton 1996
- [436] G. Roepstorff, *Pfadintegrale in der Quantenphysik*, Vieweg, Braunschweig, Wiesbaden 1992
- [437] D. Ruelle, *Helvetica Physica Acta* **36**, 183, 1963
- [438] D. Ruelle, *Il Nuovo Cimento* **61 A**, 655, 1969
- [439] D. Ruelle, *Statistical Mechanics. Rigorous Results*, World Scientific, Singapore 1999
- [440] D. Ruelle, *Thermodynamic Formalism. The Mathematical Structure of Equilibrium Statistical Mechanics*, Cambridge University Press, Cambridge 2004
- [441] L. H. Ryder, *Quantum Field Theory*, Cambridge University Press, Cambridge, 1985
- [442] C. A. Sackett, D. Kielpinski, B. E. King, C. Langer, V. Meyer, C. J. Myatt, M. Rowe, Q. A. Turchette, W. M. Itano, D. J. Wineland und C. Monroe, *Nature* **404**, 256, 2000

- [443] W. Salm, *Zugänge zur Quantentheorie*, Aulis, Köln 1999
- [444] H. H. Schaefer, *Topological Vector Spaces*, Springer, New York, Berlin, Heidelberg 1999
- [445] P. A. Schilpp (Hrsg.), *Albert Einstein als Philosoph und Naturforscher*, Kohlhammer, Stuttgart 1955
- [446] H. J. Schlichting, *Praxis der Naturwissenschaften – Physik* **2/49**, 7, 2000
- [447] H. J. Schlichting und U. Backhaus, *Physikunterricht 5-10*, Beltz, München 1981
- [448] H. J. Schlichting und U. Backhaus, *Der Physikunterricht* **18/3**, 24, 1984
- [449] M. Schlosshauer, *Decoherence and the Quantum-To-Classical Transition*, Springer, Berlin, Heidelberg 2007
- [450] G. B. Schmid, *Physics Education* **17**, 212, 1982
- [451] E. Schmidt, *Mathematische Annalen* **63**, 433, 1906
- [452] P. Schmüser, *Feynman-Graphen und Eichtheorien für Experimentalphysiker*, Springer, Berlin, Heidelberg 1995
- [453] R. Schnabel, H. Müller-Eberhardt und H. Rehbein, *Physik in unserer Zeit* **39**, 234, 2008
- [454] E. Schrödinger, *Annalen der Physik* **79**, 361, 1926
- [455] E. Schrödinger, *Annalen der Physik* **79**, 489, 1926
- [456] E. Schrödinger, *Annalen der Physik* **79**, 734, 1926
- [457] E. Schrödinger, *Annalen der Physik* **80**, 437, 1926
- [458] E. Schrödinger, *Annalen der Physik* **81**, 109, 1926
- [459] E. Schrödinger, *Sitzungsberichte der preussischen Akademie der Wissenschaften, Physicalisch–Mathematische Klasse*, 144, 1931
- [460] E. Schrödinger, *Die Naturwissenschaften* **23**, 807, 1935, abgedruckt in [44], 98
- [461] E. Schrödinger, *Die Naturwissenschaften* **23**, 823, 1935, abgedruckt in [44], 108
- [462] E. Schrödinger, *Die Naturwissenschaften* **23**, 844, 1935, abgedruckt in [44], 118
- [463] G. Schubring (Hrsg.), *Hermann Günther Grassmann: Visionary Mathematician, Scientist and Neohumanist Scholar*, Kluwer, Dordrecht, Boston 1996
- [464] L. S. Schulmann, *Techniques and Applications of Path Integration*, Dover, Minneola, New York 2001
- [465] L. Schwartz, *Radon measures on arbitrary topological spaces and cylindrical measures*, Oxford University Press, Oxford 1973

- [466] P. D. D. Schwindt, P. G. Kwiat und B.-G. Englert, *Physical Review A* **60**, 4285, 1999
- [467] J. Schwinger, *Physical Review* **82**, 664, 1951
- [468] J. Schwinger, *Physical Review* **82**, 914, 1951
- [469] J. Schwinger, *Proceedings of the National Academy of Sciences U.S.A.* **46**, 570, 1960
- [470] M. O. Scully, B.-G. Englert und H. Walther, *Nature* **351**, 111, 1991
- [471] D. Sen, A. N. Basu and S. Sengupta, *Helvetica Physica Acta* **67**, 785, 1994
- [472] C. E. Shannon, *The Bell System Technical Journal* **27**, 379, 1948
- [473] C. E. Shannon, *The Bell System Technical Journal* **27**, 623, 1948
- [474] C. E. Shannon, *The Bell System Technical Journal* **30**, 50, 1951
- [475] J. L. da Silva, *Studies in non-Gaussian Analysis*, Dissertation, University of Madeira, Funchal, Madeira 1998
- [476] J. L. Silva und L. Streit, in [423], 285
- [477] B. Simon, *Bulletin of the American Mathematical Society* **7**, 447, 1982
- [478] B. Simon, *Bulletin of the American Mathematical Society* **11**, 426, 1984
- [479] B. Simon, *Conference Proceedings of the Canadian Mathematical Society* **28**, 317, 2000
- [480] B. Simon, *Functional Integration and Quantum Physics*, American Mathematical Society, Providence, Rhode Island 2005
- [481] J. E. Sipe, *Physical Review A* **52**, 1875, 1995
- [482] A. Sloan, *Journal of Mathematical Physics* **18**, 2495, 1977
- [483] A. Sloan, *Pacific Journal of Mathematics* **79**, 207, 1978
- [484] M. von Smoluchowski, *Annalen der Physik* **21**, 756, 1906
- [485] A. Sommerfeld, *Annalen der Physik* **51**, 1, 1916
- [486] A. Sommerfeld, *Annalen der Physik* **51**, 125, 1916
- [487] A. Sommerfeld, *Vorlesungen über Theoretische Physik. Band IV: Optik*, Thun, Frankfurt am Main 1989
- [488] A. Sommerfeld, *Vorlesungen über Theoretische Physik. Band V: Thermodynamik und Statistik*, Thun, Frankfurt am Main 1988
- [489] E. Specker, *Dialectica* **14**, 239, 1960, abgedruckt (auf Englisch) in [253], 135

- [490] E. Starauschek, Zeitschrift für Didaktik der Naturwissenschaften **8**, 7, 2002
- [491] E. Starauschek, Physik und Didaktik in Schule und Hochschule **1/2**, 39, 2003
- [492] J. Stark, Physikalische Zeitschrift **10**, 902, 1909
- [493] F. Steiner, Physics Letters **A 106**, 363, 1984
- [494] M. Stöckler, in [297], 223
- [495] R. F. Streater, *Lost Causes in and beyond Physics*, Springer, Berlin, Heidelberg, New York 2007
- [496] R. F. Streater und A. S. Wightman, *PCT, Spin and Statistics, and all that*, Princeton University Press, Princeton, Oxford 2000
- [497] L. Streit und T. Hida, Stochastic Processes and their Applications **16**, 55, 1983
- [498] J. Strnad, European Journal of Physics **21**, L33, 2000
- [499] F. Strocchi, *Selected Topics on the General Properties of Quantum Field Theory*, World Scientific, Singapore, New Jersey, London, Hong Kong 1993
- [500] F. Strocchi, *An Introduction to the Mathematical Structure of Quantum Mechanics. A Short Course for Mathematicians*, World Scientific, Singapore, New Jersey, London, Hong Kong 2005
- [501] E. C. G. Sudarshan und T. Rothmann, American Journal of Physics **59**, 592, 1991
- [502] S. J. Summers und R. Werner, Physics Letters **110A**, 257, 1985
- [503] K. Symanzik, in [289], 152
- [504] L. A. Takhtajan, *Quantum Mechanics for Mathematicians*, American Mathematical Society, Providence, Rhode Island, 2008
- [505] W. Thirring, *Lehrbuch der Mathematischen Physik. Band 4: Quantenmechanik großer Systeme*, Springer, Wien, New York 1980
- [506] P. A. Tipler, *Physik*, Spektrum akademischer Verlag, Heidelberg, Berlin, Oxford 1995
- [507] H. Triebel, *Höhere Analysis*, Harri Deutsch, Thun, Frankfurt am Main 1980
- [508] H. F. Trotter, Pacific Journal of Mathematics **8**, 887, 1958
- [509] H. F. Trotter, Proceedings of the American Mathematical Society **10**, 545, 1959
- [510] A. Truman, Journal of Mathematical Physics **17**, 1852, 1976
- [511] A. Truman, Journal of Mathematical Physics **18**, 1499, 1977
- [512] A. Truman, in [5], 73

- [513] L. Vaidman, `quant-ph/0102049v1`
- [514] J. W. Warren, *Physics Education* **18**, 209, 1983
- [515] H. J. Wendel, *Moderner Relativismus*, Mohr Siebeck, Tübingen 1990
- [516] H. J. Wendel, *Die Grenzen des Naturalismus*, Mohr Siebeck, Tübingen 1997
- [517] G. Wentzel, *Zeitschrift für Physik* **22**, 193, 1924
- [518] G. Wentzel, *Zeitschrift für Physik* **40**, 574, 1926
- [519] G. Wentzel, *Zeitschrift für Physik* **41**, 828, 1927
- [520] R. F. Werner, *Physical Review A* **40**, 4277, 1989
- [521] W. Westerkamp, *Recent Results in Infinite Dimensional Analysis and Applications to Feynman Integrals*, Dissertation, Universität Bielefeld, Bielefeld 1995
- [522] H. Weyl, *Die Naturwissenschaften* **12**, 561, 1924
- [523] J. A. Wheeler und W. H. Zurek (Hrsg.), *Quantum Theory and Measurement*, Princeton University Press, Princeton 1983
- [524] N. Wiener, *Proceedings of the National Academy of Science (US)* **7**, 294, 1921
- [525] N. Wiener, *Journal of Mathematical and Physical Sciences* **2**, 132, 1923
- [526] N. Wiener, *Proceedings of the London Mathematical Society* **22**, 454, 1924
- [527] N. Wiener, *American Journal of Mathematics* **60**, 897, 1938
- [528] R. M. Wilcox, *Journal of Mathematical Physics* **8** 962, 1967
- [529] J. Wloka, *Grundräume und verallgemeinerte Funktionen*, Springer, Berlin, Heidelberg, New York 1969
- [530] W. K. Wootters und W. H. Zurek, *Physical Review D* **19**, 473, 1979, abgedruckt in [523], 443
- [531] A. Wünschmann, *Der Weg zur Quantenmechanik*, Studien-Verlag Wünschmann, Kirchheimbolanden 2009
- [532] Y. Yamasaki, *Measures on Infinite Dimensional Spaces*, World Scientific, Singapore 1985
- [533] C. N. Yang und T. D. Lee, *Physical Review* **87**, 404, 1952
- [534] I. Yoo, *Applied Mathematics & Optimization* **30**, 225, 1994
- [535] R. A. Young, *American Journal of Physics* **44**, 1043, 1976



- [536] A. Zaddach, *Grassmanns Algebra in der Geometrie. Mit Seitenblicken auf verwandte Strukturen*, B. I. Wissenschaftsverlag, Mannheim, Leipzig, Wien, Zürich 1994
- [537] T. Zastawniak, Bulletin of the Polish Academy of Sciences – Mathematics **36**, 341, 1988
- [538] T. Zastawniak, Journal of Mathematical Physics **30**, 1354, 1989
- [539] T. Zastawniak, Universitatis Iagellonicae Acta Mathematica **28**, 187, 1991
- [540] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena*, Clarendon, Oxford 2008
- [541] M. Zukowski, Physival Review A **61**, 022109, 2000



