

Diu / Guthmann / Lederer / Roulet  
**Grundlagen der Statistischen Physik**  
Ein Lehrbuch mit Übungen



---

Bernard Diu / Claudine Guthmann  
Danielle Lederer / Bernard Roulet

---

# Grundlagen der Statistischen Physik

Ein Lehrbuch mit Übungen

Übersetzt aus dem  
Französischen  
von  
Freimut Marschner



Walter de Gruyter • Berlin • New York 1994

### **Titel der Originalausgabe**

Bernard Diu, Claudine Guthmann,  
Danielle Lederer, Bernard Roulet

Éléments de Physique Statistique  
(Collection Enseignement des Sciences, 37)

© Copyright 1989 by Hermann Éditeurs des Sciences et des Arts,  
293 rue Lecourbe, 75015 Paris

### **Autoren**

Bernard Diu, Claudine Guthmann, Bernard Roulet  
Professoren an der Universität Paris VII

Danielle Lederer  
Dozentin an der Universität Paris VII

### **Deutschsprachige Ausgabe**

Privatdozent Dr. Freimut Marschner  
Organisch-Chemisches Institut  
Technische Universität Berlin  
Straße des 17. Juni 135  
D-10623 Berlin

Das Buch enthält 162 Abbildungen.

### *Die Deutsche Bibliothek – CIP-Einheitsaufnahme*

**Grundlagen der statistischen Physik** : ein Lehrbuch mit Übungen /  
Bernard Diu ... Übers. aus dem Franz. von Freimut Marschner. –  
Berlin ; New York : de Gruyter, 1994  
Einheitssacht.: Éléments de physique statistique <dt.>  
ISBN 3-11-013593-0  
NE: Diu, Bernard; Marschner, Freimut [Übers.]; EST

∞ Gedruckt auf säurefreiem Papier, das die US-ANSI-Norm über Haltbarkeit erfüllt.

© Copyright 1994 by Walter de Gruyter & Co., D-10785 Berlin.

Dieses Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen. Printed in Germany.  
Druck: Karl Gerike GmbH, Berlin. Buchbinderische Verarbeitung: Lüderitz & Bauer GmbH, Berlin. Einbandentwurf: Hansbernd Lindemann, Berlin.

## Vorwort zur deutschen Übersetzung

Das Buch « Statistische Physik » ist vom Umfang und der Form der Darstellung einzigartig. Auf vier Ebenen werden hier Physik, Mathematik und Physikalische Chemie betrieben : im Haupttext (Lehrtext), im Fußnotentext, in den Ergänzungstexten und in den Übungen - vier Ebenen des Lehrens und des Lernens. Diese Form von Lehre, Vertiefung, Ergänzung und Übung wird in sechs Buchteilen durchgeführt.

Das Autorenteam hat sich darum bemüht einen fundamentalen Zweig der Physik umfassend didaktisch so aufzubereiten, daß nicht nur Studenten der Physik sondern auch Nicht-Physiker, die an physikalischen, also naturwissenschaftlichen Fragen interessiert sind, hier eine Basis zum Verständnis des Verhaltens makroskopischer und mikroskopischer Systeme finden. Daß dazu auch die Quantenmechanik gehört, ist selbstverständlich; so finden sich auch in den vertiefenden Teilen der Fußnoten bzw. Ergänzungen stets eindeutige quantenmechanische Begründungen und Herleitungen.

Dem Stil des französischen Bildungssystems entsprechend wird in diesem Werk eine deutliche Präferenz für die mathematische Herleitung und Begründung, für den Beweis physikalischer Sätze, sowie für das Aufstellen von Regeln, Gesetzen oder Theoreme gegeben. Deutlich wird der approximative Charakter einzelner Modellvorstellungen herausgearbeitet. Hinweise auf offene Fragestellungen machen die Lektüre noch spannender.

Die Übersetzung wurde unter folgenden Gesichtspunkten erstellt :

- Deutschsprachigen Studenten wird eine Hilfe angeboten sich nicht der Doppelbelastung durch das zusätzliche Erlernen des fremdsprachlichen Fachvokabulars aussetzen zu müssen. Der Zugang zur Erarbeitung der physikalischen Grundlagen wird somit erleichtert.
- Die Nähe zu dem Originaltext soll dem Leser so direkt wie möglich das Physik-Verständnis der Autoren vermitteln. Ihn könnte dies zum Überdenken der Fachbegriffe und intendierten Argumentationsweisen anregen und zur kritischen Auseinandersetzung mit dem Inhalt ermutigen.
- Die Verbindung von Sprache und mathematischem Ausdruck wird von den Autoren intensiv betrieben. Die Übersetzung versucht, diese Intensität nachzuvollziehen.
- Die Bearbeitung der Formeln erfolgte durch kritisches Abschreiben und Kontrolle mit dem Original. Dadurch konnten auch die Formeln im Original überprüft werden. Zum Satzdruck wurde das  $\text{\LaTeX}$ -Programm von Eberhard Mattes benutzt, das auf dem  $\text{\TeX}$ -Programm

von Donald E. Knuth beruht. Die dadurch bedingten Abweichungen vom Original ändern nicht die Eindeutigkeit der Aussagen von Formeln und Text.

Ganz besonders herzlich danke ich Frau Dipl.-Phys. Dorle Streubel und Herrn Prof. Dr. Wolfgang Muschik für das kritische Gegenlesen der Übersetzung und Herrn Dr. Rudolf Weber und seinen Mitarbeiterinnen und Mitarbeitern vom Verlag Walter de Gruyter für die ausgezeichnete Zusammenarbeit.

Berlin, im Juli 1994

Freimut Marschner

**Struktur des Buches :**  
 « Hinweise für den Leser »

Dieses Buch besteht aus *zwei zusammenhängenden Teilen* : den Hauptteilen (hier kurz als Teil bezeichnet) und den Ergänzungen. Die sechs *Hauptteile* umfassen einen *Unterrichts-Lehrgang*, in dem in progressiver Reihenfolge die grundlegenden Notationen, Methoden und Ergebnisse dargestellt werden; ihr Verständnis erfordert nicht die Kenntnis irgend-einer Ergänzung. Die *Ergänzungen* sind kurze *Monographien* zur Illustration oder Vertiefung des Lehrgangs, sie sind (ohne Ausnahme) voneinander *unabhängig*; jede Ergänzung kann von dem voranstehenden Teil her verstanden werden; die Ergänzungen, die jedem Hauptteil folgen, können in beliebiger Reihenfolge bearbeitet werden (ein kurzer zusammenfassender Text unterstützt die Auswahl); es ist auf keinen Fall ratsam, alle zu lesen bevor der nächste Hauptteil begonnen wird. Jeder Teil oder jede Ergänzung wird in der Form eines *Haupttextes* dargestellt, der von *Fußnoten* begleitet wird. Das Ziel ist eine « *Lektüre auf zwei Ebenen* » anzubieten und zu ermöglichen : Der Haupttext, ohne die Fußnoten gelesen, vermittelt eine erste umfassende Idee vom Thema; wird ein tieferes Verständnis gewünscht, beginnt man noch einmal mit der Lektüre, diesmal unterstützt durch die Fußnoten. Auf keinen Fall versuche man die Fußnoten gleich beim ersten Mal aufzunehmen : Die meisten sind einfach (die Rechnungen und technischen Ausführungen sind beispielsweise systematisch in den Fußnoten enthalten), aber einige wenden sich auch an erfahrenere Leser. Die *Übungen*, die nach jeder Einheit Hauptteil-Ergänzungen vorgeschlagen werden, sind nach Themen gruppiert und innerhalb einer Gruppe nach wachsendem Schwierigkeitsgrad angeordnet.

Das Symbol  $\hat{=}$  bedeutet « gleich per Definition » \*.

Die Initialen C.D.L. beziehen sich auf C. Cohen-Tannoudji, B. Diu und F. Laloë, *Mécanique quantique*, Hermann (1985). Eine Übersetzung dieses Werkes ist im de Gruyter-Verlag in Vorbereitung.

---

\* Anmerkung des Übersetzers : In der deutschen Übersetzung wird statt  $\hat{=}$  das Symbol  $\hat{=}$  verwendet. Außerdem wird das französische mathematische Symbol  $\wedge$  durch  $\times$  und das z.T. benutzte Multiplikationszeichen  $\times$  durch  $\cdot$  ersetzt.



# Inhaltsverzeichnis

Vorwort zur deutschen Übersetzung . . . . .	V
Hinweise für den Leser . . . . .	VII
Vorwort zur französischen Ausgabe . . . . .	XIII

## TEIL I

<b>Grundbegriffe und elementare Hilfsmittel der statistischen Mechanik</b>	<b>1</b>
I. Einleitung : Makroskopischer und mikroskopischer Bereich . .	3
II. Wahrscheinlichkeiten, Information und statistische Entropie .	11
III. Zustände eines makroskopischen Systems . . . . .	16
IV. Ergodenhypothese und statistisches Ensemble . . . . .	36

<b>Ergänzungen zu Teil I</b>	<b>47</b>
I.A. Größenordnungen . . . . .	48
I.B. Statistische Verteilung einer Funktion von Zufallsvariablen .	72
I.C. Die Binomialverteilung im Grenzfall großer Zahlen . . . . .	84
I.D. Zufallsbewegung und Brownsche Bewegung . . . . .	96
I.E. Zustandsdichte eines idealen Gases . . . . .	110
I.F. Elektronenzustände in Festkörpern . . . . .	114
I.G. Information und Entropie . . . . .	131
I.H. Dichteoperator . . . . .	155
ÜBUNGEN I . . . . .	179

## TEIL II

<b>Isoliertes System im Gleichgewicht – Mikrokanonische Verteilung</b>	<b>197</b>
I. Mikrokanonische Verteilung . . . . .	199
II. Eigenschaften und Verhalten eines isolierten Systems . . . . .	209
III. Gleichgewicht zwischen Teilsystemen eines isolierten Systems	217

<b>Ergänzungen zu Teil II</b>	<b>236</b>
II.A. Kopplung zwischen makroskopischen Systemen . . . . .	237
II.B. Der ideale paramagnetische Kristall. Negative Temperaturen	246
II.C. Das klassische ideale Gas : Mischungsentropie und Gibbs-Paradoxon . . . . .	260
II.D. Thermodynamische Umwandlungen und Entropieänderungen . . . . .	275
II.E. Makroskopische Bewegung eines völlig isolierten Systems im statistischen Gleichgewicht . . . . .	304
II.F. Mikrokanonische Verteilung und mikrokanonische Entropie in der klassischen Mechanik . . . . .	325
ÜBUNGEN II . . . . .	333
<b>TEIL III</b>	
<b>System im Gleichgewicht mit einem Thermostaten. Kanonische Beschreibung</b>	<b>347</b>
I. Kanonischer Zustand und kanonische Verteilung . . . . .	351
II. Zustandssumme und freie Energie . . . . .	368
III. Der kanonische Formalismus im thermodynamischen Grenzfall	386
IV. Kanonische Beschreibung von Systemen aus wechselwirkungsfreien Teilchen . . . . .	393
V. Kanonische Verteilung in der klassischen Mechanik . . . . .	411
<b>Ergänzungen zu Teil III</b>	<b>423</b>
III.A. Paramagnetismus und Diamagnetismus von Systemen wechselwirkungsfreier Teilchen . . . . .	425
III.B. Mehratomige ideale Gase . . . . .	452
III.C. Kinetische Gastheorie . . . . .	481
III.D. $T$ - $p$ -Ensemble . . . . .	501
III.E. Spezifische Wärme von Festkörpern . . . . .	518
III.F. Verdünnte Lösungen . . . . .	547
III.G. Van der Waals-Gleichung und Übergang gasförmig-flüssig .	570
III.H. Systeme mit geladenen Teilchen . . . . .	594
III.J. Ferromagnetismus . . . . .	608
III.K. Ordnungs-Unordnungs-Übergang in binären Legierungen .	653
III.L. Wärme und Arbeit. Der « zweite Hauptsatz » der Thermodynamik . . . . .	679
ÜBUNGEN III . . . . .	691

## TEIL IV

<b>Entwicklung zum Gleichgewicht und Irreversibilität</b>	<b>737</b>
I. Mastergleichung . . . . .	739
II. Haupteigenschaften der Mastergleichung . . . . .	751

<b>Ergänzungen zu Teil IV</b>	<b>773</b>
-------------------------------	------------

IV.A. Genaue Begriffsbildung zur Ergodenhypothese . . . . .	774
IV.B. Zufällige Störungen eines « isolierten » Systems : Physikalische Diskussion und Größenordnungen bei mehratomigen idealen Gasen . . . . .	789
IV.C. Eigenschaften der Mastergleichung . . . . .	803
IV.D. « Entwicklungssatz » und Schrödinger-Gleichung . . . . .	814
IV.E. Entwicklung eines verdünnten klassischen Gases : Die Boltzmann-Gleichung . . . . .	834
IV.F. Allgemeine Entwicklungsgleichungen eines klassischen Fluids	858
ÜBUNGEN IV . . . . .	870

## TEIL V

<b>System im Gleichgewicht mit einem Teilchenreservoir. Großkanonische Beschreibung</b>	<b>893</b>
---	------------

I. Großkanonischer Zustand und großkanonische Verteilung . . .	895
II. Große Zustandssumme und großkanonisches Potential . . . . .	906
III. Der großkanonische Formalismus im thermodynamischen Grenzfalle . . . . .	917
IV. System von unabhängigen, nicht unterscheidbaren Teilchen .	923

<b>Ergänzungen zu Teil V</b>	<b>933</b>
------------------------------	------------

V.A. Adsorption . . . . .	935
V.B. Makroskopisches Gleichgewichtssystem in einem äußeren Kraftfeld . . . . .	950
V.C. Reale Gase und Virialentwicklung . . . . .	972
V.D. Fundamentalpostulat in « generalisierter » Form . . . . .	990
V.E. Dichteoperator-Formalismus . . . . .	1006
V.F. Schwankungen in einem Gleichgewichtssystem . . . . .	1028
ÜBUNGEN V . . . . .	1045

## TEIL VI

<b>Ideale Quantengase</b>	<b>1069</b>
---------------------------	-------------

I. Wechselwirkungsfreies Fermionen-Gas . . . . .	1075
II. Wechselwirkungsfreies Bosonen-Gas . . . . .	1101
III. Photonengas. Thermodynamik der Strahlung . . . . .	1134

<b>Ergänzungen zu Teil VI</b>	<b>1152</b>
VI.A. Elektroneneigenschaften in Metallen und Halbleitern . . . .	1154
VI.B. Weiße Zwerge . . . . .	1190
VI.C. Chemisches Potential und spezifische Wärme eines freien Bosonen-Gases in der Umgebung der Bose-Temperatur . . .	1208
VI.D. Helium bei tiefen Temperaturen . . . . .	1214
VI.E. Emission und Absorption von Atomstrahlung. Einstein- Koeffizienten . . . . .	1252
VI.F. Absorption und Emission von Strahlung eines makroskopi- schen Körpers. Kirchoff-Gesetz . . . . .	1259
VI.G. Die Kosmische Strahlung bei drei Kelvin . . . . .	1269
ÜBUNGEN VI . . . . .	1277
<b>Anhang</b>	<b>1309</b>
Anhang I Elementare Begriffe der Wahrscheinlichkeitstheorie . .	1311
Anhang II Quantenbeschreibung eines physikalischen Systems .	1335
Anhang III Klassische Beschreibung eines physikalischen Systems	1358
Anhang IV Die Methode der Lagrange-Multiplikatoren . . . . .	1369
Index . . . . .	1373

## Vorwort der französischen Ausgabe

Dieses Buch ist aus dem Unterricht von statistischer Mechanik und Thermodynamik hervorgegangen, den wir über mehrere Jahre hinweg auf dem Niveau der Maîtrise (Diplom bzw. 1. wissenschaftliches Examen für das Lehramt) in Physik gehalten haben. Wir sind zu Beginn sehr stark von dem Scriptum beeinflusst worden, das Philippe Nozières redigiert hat, als er nach den Reformen der Universitätsprogramme 1967, den Hochschulunterricht in der naturwissenschaftlichen Fakultät der Universität von Paris begann.

Leitgedanke unserer Vorgehensweise ist, zunächst die elementare statistische Mechanik darzustellen, um anschließend die Thermodynamik daraus herzuleiten. Die statistische Mechanik stellt mit der Quantenmechanik und der Relativitätstheorie einen der « Pfeiler » der modernen Physik dar; es ist wichtig, daß die Studenten sich dieser fundamentalen Theorie bewußt werden, und sie nicht als einfachen Versuch betrachten, um die Thermodynamik *im nachhinein* zu rechtfertigen. Die Thermodynamik ihrerseits gewinnt im Rückblick ein genaueres und tieferes Verständnis ihrer « Hauptsätze » und Methoden. Das vorliegende Buch stellt die statistische Mechanik dar und zeigt ihre Verbindung zur Thermodynamik auf, aber es ersetzt keinen eigenen Kurs in Thermodynamik; dieser muß offensichtlich, um den größtmöglichen Nutzen aus den von den Studenten erworbenen Kenntnissen zu ziehen, nicht mehr auf der Darlegung der traditionellen « Hauptsätze » beruhen, sondern auf den fundamentalen Eigenschaften der Entropie, wie sie in der statistischen Mechanik vorgeführt wird.

Die Ausarbeitung dieses Werkes wird von zwei Gesichtspunkten begleitet. An erster Stelle haben wir uns bemüht, alle behandelten Themen so ausführlich und klar wie möglich darzustellen, um sie den Studenten auf dem Niveau der Maîtrise in Physik zugänglich zu machen. Anschließend haben wir versucht, nach und nach die Theorie durch physikalische Anwendungen zu veranschaulichen. Offensichtlich sind diese beiden Ansprüche auf den ersten Blick kaum miteinander kompatibel, aber unsere Feinarbeit hatte stets zum Ziel, klar und deutlich auf elementarem Niveau, auf das wir uns festgelegt haben, die physikalischen Phänomene zu erklären, die wir für wichtig und interessant halten.

In technischer Hinsicht vereinigt das Buch zwei verschiedene Teile, die miteinander verzahnt werden : den Hauptteil und die Ergänzungen. Andererseits haben wir bei der Redaktion des Buches systematisch Gebrauch von Fußnoten gemacht, um den Hauptteil stofflich zu entlasten. Der Grund dieser Vorgehensweise und die « Gebrauchsanleitung », die daraus folgt, werden auf der vorangestellten Sonderseite VII angegeben. Die aufmerksame

Lektüre dieser Seite vor der Benutzung des Buches wird dringend empfohlen.

Was den Inhalt anbelangt, so haben wir uns für eine Darstellung der statistischen Mechanik entschieden, die sehr klar zwischen den verschiedenen statistischen Verteilungen (mikrokanonisch, kanonisch, großkanonisch, ...) der mikroskopischen Zustände eines Systems im Gleichgewicht unterscheidet, gemäß der physikalischen Situation, in welcher das System sich befindet (isoliert, im Kontakt mit einem Thermostaten, oder mit einem Teilchen-Reservoir, ...). Offenkundig geben die verschiedenen Beschreibungen die gleichen Ergebnisse, wenn das System genügend groß ist (thermodynamischer Grenzfall), was bei den meisten Anwendungen zutrifft. Aber die fundamentalen Hypothesen und die spezifischen Methoden der statistischen Mechanik werden günstiger bei einer solchen Darstellung erhellt; sie lassen sich jedenfalls besser von den Studenten gedanklich durchdringen, wenn sie mit stufenweiser Steigerung und großer Ausführlichkeit dargestellt werden.

Die statistische Mechanik hat zum Ziel, das Verhalten von makroskopischen Systemen aus ihren mikroskopischen Eigenschaften zu erklären. Generell beschreibt die Quantenmechanik die Eigenschaften und die Entwicklungen physikalischer Systeme im mikroskopischen Maßstab. Die statistische Mechanik ist daher auf dieser quantenmechanischen Beschreibung aufgebaut. Aus pädagogischen Gründen wollten wir jedoch vermeiden, daß die physikalischen Ideen und Überlegungen, die für die statistische Mechanik typisch sind, nicht durch Verständnisschwierigkeiten verdunkelt werden, die mit der Quantenmechanik zusammenhängen. Deshalb stellen wir die statistische Mechanik in einem vereinfachten Formalismus dar (Wahrscheinlichkeiten der Eigenzustände des Hamiltonoperators), der nur elementare Kenntnisse der Quantenmechanik erfordert. Einzig einige Ergänzungen, die den Formalismus der Dichteoperatoren einführen und anschließend entwickeln, und einige Fußnoten benutzen fortgeschrittenere theoretische Erklärungen. Der einfache Formalismus, den wir übrigens durchgängig benutzen, ist im wesentlichen dem des Dichteoperators gleichwertig; dieser Dichteoperator erweist sich nur in einigen technisch komplizierten Problemen als notwendig, wie die Bereitstellung von besonders weit entwickelten Näherungsmethoden oder die strenge Herleitung allgemein theoretischer Ergebnisse. Andererseits bietet der von uns gewählte Rahmen dem Studenten am besten eine kontinuierliche Hinführung zu den physikalischen Grundlagen der statistischen Mechanik.

Dieses Buch wendet sich in erster Linie an Studenten der Physik im 3. bzw. 4. Studienjahr (Second Cycle de physique des Universités): Sie bestimmen das wissenschaftliche und technische Niveau als potentielle Leser. Einige Ergänzungen sind jedoch eher auf dem Niveau eines Studiums für Fortgeschrittene, Diplomanden und Doktoranden (Troisième Cycle de physique des

Universités, etwa Graduierten-Kolleg) angesiedelt, wobei sie natürlich auf der Grundlage der voranstehenden Hauptteile verständlich sind. Die Überlagerung und der Einbau von Hauptteilen und Ergänzungen, von Haupttext und Fußnoten, von theoretischen Vorstellungen und physikalischen Anwendungen, sowie die Diversifizierung der Strukturen (sehr hierarchisch oder im Gegensatz dazu « ausufernd ») haben zum Ziel, die verschiedenen Nutzungsmöglichkeiten zu erleichtern : Wir hoffen, - daß der Student alleine den Lernstoff versteht, indem er dem Ablauf der Hauptteile folgt und einige Ergänzungen dazu studiert, - daß der Dozent einen Kurs oder eine Vorlesung im Hauptstudiengang oder nach dem Diplom (Maîtrise oder D.E.A., diplôme des études approfondies, nach 5-jährigem Studium) aufbauen kann, indem er aus den Hauptteilen und Ergänzungen schöpft, - und daß der Anfänger in der Forschung mit Erfolg dieses Buch zu seinem jeweils besonderen Problem befragen kann.

Wir haben nicht versucht, eine vollständige Bibliographie aufzustellen : Die Bücher der statistischen Physik sind sehr zahlreich. Begnügen wir uns damit, hier einige Bücher zu zitieren, die wir auf dem Niveau des Physik-Hauptdiploms (Maîtrise de physique) für besonders nützlich halten.

In Französisch :

R. Balian, « Du microscopique au macroscopique », Cours de Physique Statistique de l'École Polytechnique, tomes 1 et 2. Ellipses (1982).  
Kurs in statistischer Physik für Polytechnische Hochschulen, Bände 1 und 2.

L. Landau und E. Lifchitz, Physique Statistique, Édition M.I.R. ,Moscou (1967).

In Englisch :

F. Reif, Fundamentals of Statistical and Thermal Physics, McGraw-Hill, New York (1965);

R. Kubo, Statistical Mechanics, North-Holland, Amsterdam (1971).

D. McQuarrie, Statistical Mechanics, Harper and Row, New York (1976).

S. K. Ma, Statistical Mechanics, World Scientific, Singapore (1985).

G. Wannier, Statistical Physics, Wiley, New York (1966).

K. Huang, *Statistical Mechanics*, Wiley, New York (1967).

C. Kittel and H. Kroemer, *Thermal Physics*, Freeman, San Francisco (1980).

In Deutsch\* :

F. Reif, *Statistische Physik und Theorie der Wärme*, 3. Aufl., de Gruyter, Berlin (1987).

C. Kittel, *Physik der Wärme*, Oldenbourg, München (1973).

W. Macke, *Thermodynamik und Statistik*, 3. Aufl., Akadem. Verlagsges., Leipzig (1967).

A. Sommerfeld, *Thermodynamik und Statistik*, 3. Aufl., Geest und Portig, Leipzig (1965).

R. Becker, *Theorie der Wärme* (Heidelberger Taschenbücher Nr. 10), Springer-Verlag, Berlin (1966).

J. I. Frenkel, *Statistische Physik*, Akademie-Verlag, Berlin (1957).

K. Huang, *Statistische Mechanik*, (B.-I.-Hochschultaschenbücher Nr. 68, 69, 70), Bibliogr. Institut, Mannheim (1964).

A. J. Khinchin, *Mathematische Grundlagen der statistischen Mechanik* (B.-I.-Hochschultaschenbücher Nr. 58/58a), Bibliogr. Institut, Mannheim (1964).

L. D. Landau, E. M. Lifschitz, *Statistische Physik*, 2. Aufl., Akademie-Verlag, Berlin (1969).

A. Münster, *Statistische Thermodynamik*, Springer-Verlag, Berlin (1956).

A. Münster, *Prinzipien der statistischen Mechanik*, in *Handbuch der Physik*, Bd. III/2, S. 176-412, Springer-Verlag, Berlin (1959).

E. Schrödinger, *Statistische Thermodynamik*, Barth-Verlag, Leipzig (1952).

J. Wilks, *Der dritte Hauptsatz der Thermodynamik*, Vieweg Verlag, Braunschweig (1963).

---

\*Anmerkung des Übersetzers : Diese kurze Zusammenstellung basiert auf einer Zusammenstellung in dem Buch von Reif.

*Danksagung :*

Wir danken besonders herzlich den Kollegen, mit denen wir in den vergangenen Jahren gemeinsam unterrichtet haben, insbesondere Françoise Babilbar, Odile Betbeder-Matibet und Laurence Viennot, sowie Christiane Caroli : Ihre Bemerkungen, ihre Vorschläge und ihre Ermutigungen waren uns sehr wertvoll. Wir danken ebenfalls Patricia Méralli, Stéphane Dupin, Laure-Anne Némirouvsy und Marie-Claire Bégot, die mit viel Geduld unser Manuskript getippt haben und Romain Lécuille, der bereitwillig die Abbildungen gezeichnet hat.



# Teil I

# Übersicht von Teil I

---

## I. Einleitung : Makroskopischer und Mikroskopischer Bereich

- A. « Definitionen » und kurzer historischer Abriss
  - 1. Die makroskopische Physik
  - 2. Die « Atomhypothese » und ihre Entwicklungen
  - 3. Größenordnungen und « Definitionen »
- B. Die statistische Mechanik
  - 1. Übergang vom mikroskopischen zum makroskopischen Bereich
  - 2. Die statistische Methode
  - 3. Statistische Mechanik und Quantenmechanik

- a. Summation über die Zustände und Summation über die Energien
  - b. Übergang zu einem Integral über die Energie
- 2. Beispiel : Freie Teilchen in einem makroskopischen Kasten
    - a. Einteilchensystem
    - b. System von  $N$  freien und unabhängigen Teilchen
  - 3. Verallgemeinerung zu einem System von  $N$  schwach gekoppelten Elementen

## C. Beschreibung eines klassischen Systems

- 1. Klassischer mikroskopischer Zustand
- 2. Makroskopischer Zustand

## II. Wahrscheinlichkeiten, Information und statistische Entropie

- A. Definition der statistischen Entropie
- B. Haupteigenschaften
  - 1. Positivität und Symmetrie
  - 2. Minimum
  - 3. Maximum (feste Anzahl möglicher Ereignisse)
  - 4. Variation mit der Anzahl möglicher Ereignisse
  - 5. Additivität

## IV. Ergodenhypothese und statistisches Ensemble

- A. Zeitliche Entwicklung eines makroskopischen Systems
  - 1. Untersuchung eines einfachen Beispiels
    - a. Entwicklung auf mikroskopischem Niveau : « thermische Bewegung » und Fluktuationen
    - b. Entwicklung auf makroskopischem Niveau
  - 2. Verallgemeinerung zu einem beliebigen makroskopischen System
- B. Mittelwert über die mikroskopischen Fluktuationen
  - 1. Zeitlicher Mittelwert
    - a. Definition
    - b. Beschreibung der makroskopischen Entwicklung
  - 2. Mittelwert über ein Ensemble
    - a. Statistisches Ensemble
    - b. Charakterisierung eines statistischen Ensembles
    - c. Entwicklung des Ensembles in der Zeit
  - 3. Diskussion : Die Ergodenhypothese

## III. Zustand eines makroskopischen Systems

- A. Quantenbeschreibung des Systems
  - 1. Der mikroskopische Zustand : ein « reiner Zustand »
  - 2. Der makroskopische Zustand : eine « statistische Mischung »
    - a. Definition einer statistischen Mischung von reinen Zuständen
    - b. Charakterisierung des makroskopischen Zustandes
- B. Dichte von (mikroskopischen) Zuständen
  - 1. Definition und Nützlichkeit der Dichte von Zuständen

# Teil I

## Grundbegriffe und elementare Hilfsmittel der statistischen Mechanik

### I. Einleitung : Makroskopischer und mikroskopischer Bereich

Die statistische Mechanik, deren Begründer Ludwig Boltzmann (1844 - 1906) war, gestattet zu verstehen, wie sich die Eigenschaften von Körpern im gängigen Maßstab, den wir als makroskopisch qualifizieren, aus den Eigenschaften ihrer mikroskopischen Bestandteile (Atome, Moleküle, Elektronen, Ionen) erklären lassen. Vor Einführung der Begriffe und Hilfsmittel, die als Grundlage zum Studium der statistischen Mechanik dienen, werden wir diesen ersten Abschnitt dazu verwenden, das allgemeine Ziel zu formulieren und - zumindest im Prinzip - die fundamentalen Ideen aufzuzeigen, um dieses Ziel zu erreichen.

#### A. « Definitionen » und kurzer historischer Abriss

##### 1. Die makroskopische Physik

Bis in die zweite Hälfte des neunzehnten Jahrhunderts beschäftigte sich die Wissenschaft mit der Untersuchung der durch die menschlichen Sinne direkt wahrnehmbaren Phänomene, selbst wenn die Logik dieser Untersuchung die Wissenschaftler dazu führte, die direkte sinnliche Wahrnehmung durch objektive Messungen zu ersetzen, die mit Hilfe von ständig verfeinerten Meßapparaten durchgeführt wurden. Auf dem Gebiet der Physik untersuchte man solchermaßen die Mechanik, die Elektrizität, den Magnetismus, die Optik, die Akustik, aber auch die Thermodynamik und die Zustandsänderungen der Materie, über die wir in Einzelheit im folgenden noch sprechen werden.

Es ist wichtig anzumerken, daß die verschiedenen Forschungsgebiete, die wir eben aufgezählt haben, im neunzehnten Jahrhundert praktisch voneinander unabhängig blieben, denn die verschiedenen Aspekte der Realität, für die man sich interessierte, schienen keine Verbindung zu gestatten; insbesondere waren die Thermodynamiker weit davon entfernt sich vorzustellen (der erbitterte Widerstand der Mehrzahl unter ihnen gegen die Ideen von Boltzmann zeigt dies offenkundig), daß die Theorie, die sie aufbauten, an irgendeines der anderen Gebiete, insbesondere an das der Mechanik, geknüpft werden könnte.

Wohlgermerkt, die makroskopische Physik, das heißt die Physik im menschlichen Maßstab oder in einem noch größeren Maßstab (Astronomie, Himmelsmechanik), hat das neunzehnte Jahrhundert überlebt. Sie stellt und löst auch heute noch äußerst interessante Probleme, und ihre Entwicklung ist ohne Zweifel noch nicht abgeschlossen.

## **2. Die « Atomhypothese » und ihre Entwicklungen**

Aber das neunzehnte Jahrhundert erlebte auch den Aufstieg auf wissenschaftliches Niveau einer sehr alten philosophischen Idee, die bis in die Antike zurückreicht. Die quantitative Untersuchung chemischer Reaktionen enthüllte Gesetze (Gesetz der « konstanten Proportionen », Gesetz der « multiplen Proportionen »), die sich auf sehr überzeugende Weise im Rahmen der « Atomhypothese » deuten ließen, nach der die verschiedenen Reagenzien Anhäufungen oder Aggregate mikroskopischer Bestandteile (Atome und Moleküle) sind. Diese Hypothese, die zunächst als eine Methode die Dinge darzustellen betrachtet wurde (« alles ereignet sich, wie wenn ... »), denn es schien undenkbar, die Atome direkt sichtbar zu machen, eroberte im Verlauf des neunzehnten Jahrhunderts nach und nach die Wissenschaftsgebiete und setzte sich schließlich zu Beginn des zwanzigsten Jahrhunderts in der Form einer nunmehr universell anerkannten « Atomtheorie » durch.

Trotz ihrer Erfolge war die makroskopische Physik somit dazu verurteilt, früher oder später ihre fundamentale Bedeutung zu Gunsten einer mikroskopischen Physik zu verlieren : Die in unserem Maßstab beobachteten Phänomene sind mehr oder weniger direkte oder als solche zu erkennende Folgeerscheinungen von tiefer liegenden Phänomenen, die sich auf die mikroskopischen Bestandteile der betrachteten Objekte beziehen; so sind es konsequenterweise diese mikroskopischen Bestandteile und ihre Eigenschaften, die es sich lohnt zu studieren, wenn man nach einem fundamentalen Verständnis der Gesamtheit der physikalischen Welt forscht.

Geboren zu Beginn dieses Jahrhunderts, hat sich die mikroskopische Physik als äußerst reichhaltig erwiesen und hat ebenso spektakuläre wie unvor-

hergesehene Entwicklungen gekannt : Beweis der Struktur des Atoms, danach der des Kerns, Entdeckung der Quantenmechanik, Untersuchung der mikroskopischen Eigenschaften von Feststoffen und Flüssigkeiten, dann erst jüngst die Erforschung der Elementarteilchen ...

### 3. Größenordnungen und « Definitionen »

Versuchen wir jetzt, zumindest halbquantitativ, den Umfang dieser beiden Gebiete Makrokosmos und Mikrokosmos zu erfassen.

Der Übergang vom Mikrokosmos zum Makrokosmos erfolgt mittels der Avogadro-Konstante (früher Loschmidt-Zahl)

$$N_A = 6,02 \cdot 10^{23} \text{ mol}^{-1}, \quad (\text{I.1})$$

die folgendermaßen eingeführt wird. Die quantitative Untersuchung chemischer Reaktionen führt dazu, jedem Reinstoff eine charakteristische Masse zuzuordnen, welche wir heute als ein *Mol*<sup>1</sup> des betrachteten Reinstoffs bezeichnen : So hat ein Mol Wasser die Masse 18 g, ein Mol Wasserstoff (molekular) 2 g, etc. Ein Mol ist also eine makroskopische Größe. Aber wie wir im Abschnitt 2 zeigten, ist die Möglichkeit, solche charakteristischen Massen mit den verschiedenen Reinstoffen zu verknüpfen, eine Bestätigung ihrer « körnigen » Eigenschaft, das heißt der Existenz von Molekülen : Das Verhältnis 18 g/2 g zwischen den Massen eines Mols Wasser und eines Mols Wasserstoff ist ganz einfach das, welches zwischen den entsprechenden Molekülen H<sub>2</sub>O und H<sub>2</sub> auftritt. Die Avogadro-Konstante (I.1) ist also für irgendeinen reinen Stoff das Verhältnis zwischen der Masse eines Mols und der Masse eines Moleküls dieses Stoffes, das heißt der Anzahl von Molekülen<sup>2</sup>, die ein Mol bilden.

Die Avogadro-Konstante ist « phantastisch groß ». Dies erklärt, warum die Existenz von Atomen und Molekülen schwierig zu beweisen war und Anlaß zu leidenschaftlicher Polemik gegeben hat. Die gewaltige Größe der Avogadro-Konstanten beinhaltet auch eine *saubere Trennung der Größenordnungen* zwischen dem makroskopischen und dem mikroskopischen Gebiet : Man kann zum Beispiel ohne Schwierigkeiten einen Bruchteil eines Mols definieren, der genügend klein ist, um bei einer Messung der Masse auf makroskopischen Niveau (sagen wir 10<sup>-8</sup> oder 10<sup>-9</sup> Mol) nicht mehr wahrgenommen zu werden, der aber genügend groß ist, um noch sehr viele

<sup>1</sup>Das Mol ist vom Isotop 12 des Kohlenstoffs her definiert : Per Definition ist die Masse eines Mols <sup>12</sup>C genau 12 g.

<sup>2</sup>Der Begriff des Mols und die Bedeutung der Avogadro-Konstanten läßt sich auch auf die Fälle übertragen, in denen die Struktur des betrachteten Stoffes nicht molekular sondern ionisch ist.

Moleküle zu enthalten ( $10^{14}$  bis  $10^{15}$  in diesem Beispiel).

Die Untersuchung auf jedem einzelnen Gebiet der Physik ist von einer *Auswahl der dem Gebiet angemessenen Einheiten* begleitet : Wir sagen, daß eine Einheit als angemessen betrachtet wird, wenn die Zahlen, die die entsprechende physikalische Größe messen, meistens zwischen  $10^{-4}$  und  $10^{+4}$  liegen (eventuell mit gelegentlichen « Spitzen » bis zu  $10^{-6}$  oder  $10^{+6}$ ); wenn diese Zahlen systematisch sehr groß oder sehr klein werden, so ändere man die Einheit, um sie auf vernünftige Werte zu bringen. Dieser Begriff von angemessenen Einheiten gestattet uns auf halb-quantitative Art, die Terme « makroskopisch » und « mikroskopisch » zu « definieren ».

Im Folgenden werden die Phänomene und Größen als « makroskopisch » bezeichnet, für die die angemessenen Einheiten im wesentlichen die des festgelegten M.K.S.A.-Systems sind, oder eventuell ihre Vielfachen oder Bruchteile : Somit werden die makroskopischen Längen in « Meter » (oder in Zentimeter, oder sogar in Millimeter), die makroskopischen Energien in *Joule* (oder Millijoule) angegeben. Man befindet sich erst recht im makroskopischen Gebiet, wenn man Einheiten benutzen muß, die größer sind als die vorhergehenden. Im Gegensatz dazu ordnen wir die Phänomene und Größen als *mikroskopisch* ein, die auf dem Niveau von Atomen oder Molekülen angesiedelt sind, oder eventuell auf noch niedrigerem Niveau : Hier sind die angemessenen Einheiten das *Elektronenvolt* ( $1 \text{ eV} = 1,6 \cdot 10^{-19} \text{ J}$ ) für die Energien<sup>3</sup>, und das *Angström* ( $1 \text{ \AA} = 10^{-10} \text{ m}$ )\* für die Längen.

Diese Fragen der Größenordnungen werden im Einzelnen in der Ergänzung I.A wieder aufgegriffen.

## B. Die statistische Mechanik

### 1. Übergang vom mikroskopischen zum makroskopischen Bereich

Es gibt sicher keinen Zweifel daran - ohne daß dies in irgendeiner Weise ein Werturteil beinhaltet -, daß *die mikroskopische Physik grundlegender ist als die makroskopische Physik* : Die eigentlichen Ursachen der Phänomene müssen auf mikroskopischem Niveau erforscht werden; im wesentlichen erkennt man auf makroskopischem Niveau nur gewisse Konsequenzen des Verhaltens von mikroskopischen Teilchen. Ein makroskopisches Gesetz (wie das Ohmsche Gesetz oder das Gesetz von Boyle-Mariotte) kann nur ein « Abbild » mikroskopischer Gesetze sein, und man muß suchen, « woher es

<sup>3</sup>Man bleibt im Mikrokosmos, wenn die Energien das MeV ( $10^6 \text{ eV}$ ) erreichen oder sogar das GeV ( $10^9 \text{ eV}$ ) : Es gibt tatsächlich einen Faktor  $10^{19}$  zwischen dem Elektronenvolt und dem Joule. Dennoch betritt man mit den MeV und den GeV relativistisches Gebiet, das wir im Folgenden ein wenig streifen wollen.

\*Anmerkung des Übersetzers : Das Angström ist keine gesetzliche Einheit.

kommt ». In gewissen Fällen sind die mikroskopischen Eigenschaften in ihrem makroskopischen Auftreten *direkt erkennbar*. Zum Beispiel erklärt das Energiespektrum des  $\text{Cu}^{2+}$ -Ions in überzeugender Weise die blaue Farbe der Kupfersalze. Auch ist die Kraft, die zwischen dem Elektron und dem Proton des Wasserstoffatoms wirksam ist, die gleiche, die von Coulomb auf makroskopischem Niveau entdeckt wurde<sup>4</sup>. Dies bedeutet nicht, daß man nun nicht mehr zu untersuchen braucht, wie der Übergang vom Mikroskopischen zum Makroskopischen erfolgt : Es muß erklärt werden, warum diese radikale Änderung der Größenordnungen nicht grundlegend gewisse Gesetze oder Eigenschaften verändert. Aber man erwartet eine einfache Erklärung. Besonders in solchen Fällen deuten auf den tieferen mikroskopischen Ursprung *charakteristische Effekte* hin, die ihren Platz nicht in einer rein makroskopischen Theorie haben : Zum Beispiel verursacht auf mikroskopischem Niveau die thermische Bewegung der Elektronen in den Leitern ein « Rauschen » in den elektrischen Schaltkreisen (das heißt einen zufälligen Strom, der ständig die Richtung ändert, von schwacher Intensität ist, sich aber sicher nachweisen läßt).

Man könnte naiverweise aus dem Voranstehendem schließen, daß allein die mikroskopische Physik von Bedeutung sei, denn ihr Verständnis würde sofort das der makroskopischen Physik beinhalten. Allerhöchstens - aber das wäre schon eine wesentliche Einschränkung - würde man zugestehen, daß eine detaillierte quantitative Erklärung eines solchen makroskopischen Phänomens zu lange Rechnungen selbst mit Hilfe von Rechnern erfordern könnte, um wirklich zu einem guten Ergebnis zu führen. Situationen dieser Art sind in der Physik geläufig : Zum Beispiel glaubt man die Struktur des Bleiatoms (82 Elektronen) prinzipiell verstanden zu haben, sobald man die Theorie des Wasserstoffatoms entwickelt hat.

Aber eine *solche Schlußfolgerung wäre grundlegend falsch*. Es gibt vielmehr zahlreiche makroskopische Phänomene, die sich nicht direkt aus den mikroskopischen Gesetzen, wenn auch nur qualitativ erklären lassen. Dies ist zum Beispiel bei der Verflüssigung eines Gases oder beim Betrieb einer

---

<sup>4</sup>Wenn man eine Kugel von makroskopischer Dimension elektrisiert, bringt man Elektronen in der Größenordnung von  $10^{10}$  ins Spiel. Die Kraft zwischen zwei Kugeln dieses Typs ist schließlich ungefähr  $10^{20}$ -mal intensiver als zwischen einem Elektron und einem Proton bei gleichem Abstand; dies macht die Kraft makroskopisch messbar, ändert aber nichts an ihrer Natur oder ihren Eigenschaften. Wohlgermerkt, in dieser Situation macht ein Elektron mehr oder weniger auf  $10^{10}$  keinen wahrnehmbaren Unterschied aus : Auf makroskopischem Niveau tritt die Quantifizierung der Ladung (das heißt ihre "körnige" Eigenschaft) völlig unbemerkt auf, und man kann eine makroskopische Theorie der Elektrostatik und der Elektrodynamik entwickeln, in der die Quantitäten der Elektrizität als kontinuierlich behandelt werden (man spricht von « infinitesimalen » Quantitäten der Elektrizität  $dq$ , die trotzdem sehr groß gegenüber der Elementarladung sind).

Dampfmaschine der Fall. Hier sind die makroskopischen Strukturen nicht ohne weiteres aus den mikroskopischen Eigenschaften des Systems erklärbar; es erscheint sogar völlig unerwartet : *Der Übergang vom Mikrokosmos zum Makrokosmos stellt sich als nicht trivial heraus* und erfordert eine besondere Untersuchung.

## 2. Die statistische Methode

Wie kann man das gestellte Problem lösen? Die so verwendete Methode bedient sich der Tatsache, daß beim Übergang vom Mikroskopischen zum Makroskopischen sehr große Zahlen eine Rolle spielen.

Um uns davon eine Vorstellung zu machen, betrachten wir  $1 \text{ mm}^3$  eines Gases, wie die Luft, bei Atmosphärendruck und gewöhnlicher Temperatur. Er enthält ungefähr<sup>5</sup>

$$N \simeq 3 \cdot 10^{16} \quad (\text{I.2})$$

Moleküle. Nehmen wir der Einfachheit halber an, daß die (mikroskopischen) Bewegungen dieser Moleküle durch die Newtonsche Mechanik beschrieben werden können (siehe Ergänzung I.A, §III.2.b). Um den Zustand eines einzelnen Moleküls zu einem bestimmten Zeitpunkt zu charakterisieren, müssen nunmehr mindestens 6 Zahlen angegeben werden (3 Orts- und 3 Geschwindigkeitskoordinaten). Um also zu einem bestimmten Zeitpunkt den Zustand eines Systems, das diese  $\text{mm}^3$  Gas darstellen, zu bestimmen, müßte man ungefähr  $2 \cdot 10^{17}$  Zahlen kennen. Dies ist *vollständig ausgeschlossen* : Man bedenke, daß zum Beispiel ein Rechner, der 40 Zahlen pro Zeile und eine Zeile pro cm schreibt, 50 Milliarden km Papier verbrauchen würde, um diesen momentanen Zustand zu registrieren! Selbst wenn dies in Kauf genommen würde, wäre die Masse an entsprechender Information strenggenommen *unauswertbar*<sup>6</sup>.

Diese großen Zahlen müssen der beschränkten Anzahl von makroskopischen Eigenschaften gegenüber gestellt werden, die man an dieser Gasprobe messen kann : Volumen, Druck, Temperatur, spezifische Wärme, Viskosität, Brechungszahl . . . Wenn man zusammenzählt, bleibt man weit von den vorherigen  $10^{17}$  entfernt.

<sup>5</sup>Unter < Normalbedingungen > von Temperatur und Druck ( $0^\circ \text{ C}$  und 1 Atmosphäre) nimmt ein Mol eines Gases näherungsweise ein Volumen von  $22,4 \text{ l}$  ein. Daraus folgt

$$N = \frac{N_A}{22,4 \cdot 10^6 \text{ mm}^3} \simeq 3 \cdot 10^{16} / \text{mm}^3.$$

<sup>6</sup>Die leistungsfähigsten Rechner sind gegenwärtig in der Lage die Bewegung von Systemen zu < verfolgen >, die alles in allem einige tausend Teilchen enthalten. Man bleibt damit sehr weit unterhalb der Zahl  $N$  der Gleichung (I.2).

Man versteht nunmehr, daß der Übergang vom Mikrokosmos zum Makrokosmos sich nicht immer direkt durchführen läßt, wie das im Abschnitt 1 betrachtet wurde, das heißt « vollständig » das Verhalten von Systemen « zahlreicher » mikroskopischer Objekte zu berechnen. *Die individuellen Eigenschaften* dieser Objekte *sind nicht mit einem Mal zugänglich* (denn es ist unfaßbar, sie registrieren zu können) *und ohne Interesse*, um die makroskopischen Eigenschaften der Systeme zu erklären. Man begreift auch, daß diese Eigenschaften *qualitativ verschieden* von denen eines Ensembles von 2, 3, 10 oder 100 mikroskopischen Objekten sein können, und daß sich daher die Erklärung gewisser makroskopischer Phänomene aus den Gesetzen der mikroskopischen Physik offensichtlich als nicht trivial erweisen.

Doch legt die Tatsache, daß man es mit großen Zahlen zu tun hat, den *Gebrauch von Methoden der Wahrscheinlichkeitsrechnung* nahe. Sind die mikroskopischen Eigenschaften solch großer Systeme unzugänglich, so heißt das, daß die makroskopischen Messungen « grob » sein müssen : Sie sind gegenüber dem individuellen Verhalten der mikroskopischen Objekte nicht empfindlich genug, sondern nur gegenüber einem *Mittelwert* (eines Typs, der noch genau zu definieren ist) aus dieser großen Anzahl von Objekten<sup>7</sup>. Die wahrscheinlichkeitstheoretischen Methoden sind dazu bestimmt, die Verknüpfung zwischen dem Mikrokosmos und dem Makrokosmos herzustellen und begründen die *statistische Mechanik*. Der Untersuchung dieser Methoden und ihrer Anwendung in konkreten Situationen ist dieses Buch gewidmet (die Begriffe und wesentlichen Resultate der Wahrscheinlichkeitstheorie werden im Anhang I dargestellt; die Ergänzungen I.B, I.C und I.D entwickeln einige etwas mehr technische Gesichtspunkte).

Man könnte darüber enttäuscht sein, daß sich der Einsatz der mikroskopischen Gesetze auf makroskopischem Niveau « nur » auf probabilistische

---

<sup>7</sup>Um diesen fundamentalen Punkt zu veranschaulichen, geben wir noch ein Beispiel. Wir betrachten einen Leitungsdraht als Teil eines elektrischen Stromkreises. Die Leitungselektronen, die er enthält, bewegen zufällig sich in alle Richtungen mit Geschwindigkeiten von mehr als Hunderten von km/s. In Abwesenheit der Stromquelle in diesem Stromkreis, das heißt in Abwesenheit eines elektrischen Feldes in diesem Draht, ist der Strom Null, denn « im Mittel » gibt es zu jedem Augenblick ebenso viele Elektronen, die sich in die eine wie in die entgegengesetzte Richtung bewegen. Wenn nun eine Stromquelle in diesem Draht ein elektrisches Feld unterhält, ist dieses Gleichgewicht der elektrischen Kraft wegen unterbrochen, die auf alle Elektronen in der gleichen Richtung wirkt; es gibt jetzt « im Mittel » mehr Elektronen, die sich in Richtung der Kraft bewegen als in umgekehrter Richtung, und es ist diese Differenz, die sich in Form des elektrischen Stromes zeigt : Die Gesamtheit der « Wolke » von Elektronen verschiebt sich insgesamt in dem Draht. Wohlgermerkt, die thermische Bewegung hört insofern nicht auf. Man muß übrigens hervorheben, daß die Geschwindigkeit der gesamten « Wolke » (mittlere Geschwindigkeit der Elektronen) nur einige mm/s beträgt, und daher beträchtlich kleiner ist als individuellen Geschwindigkeiten; dennoch bestimmt sie den makroskopischen Strom.

Art bewerkstelligen läßt, was den Eindruck von Approximativität und Unvollständigkeit erweckt. Dennoch ist die *Riesengröße der auftretenden Zahlen* in dieser Hinsicht beruhigend : Das « Gesetz der großen Zahlen », das die Anwendungen der wahrscheinlichkeitstheoretischen Methoden bestimmt, mündet *praktisch in gewisse Aussagen*, das heißt in eine *praktisch exakte Beschreibung* des makroskopischen Niveaus.

Um ein wenig genauer zu sein, greifen wir nochmal das Beispiel der oben erwähnten Gasprobe auf. Nehmen wir an, daß wir einem Raum  $1 \text{ mm}^3$  Luft entnehmen. Man erwartet, daß die Probe  $N = 3 \cdot 10^{16}$  Moleküle enthält (Gleichung (I.2)). Dennoch weiß man, daß die unaufhörliche Bewegung der Luftmoleküle in dem Raum zu *Fluktuationen* der Anzahl derjenigen Moleküle führt, die zu einem bestimmten Augenblick in dem betrachteten Volumen sind : In gewissen Augenblicken ist diese Anzahl niedriger als die erwartete Anzahl  $N$ , in anderen Augenblicken ist sie größer. Diese Anzahl wird wahrscheinlich zwischen  $N - \sqrt{N}$  und  $N + \sqrt{N}$  liegen (die Fluktuationen sind von der Ordnung  $\sqrt{N}$ . Siehe Anhang I, §II.1.c.γ). Man muß also festhalten, obgleich  $\sqrt{N} \cong 1,7 \cdot 10^8$  selbst groß ist, daß die *relative Fluktuation*  $\sqrt{N}/N$  in der Größenordnung von  $0,6 \cdot 10^{-8}$  ist, das heißt unmöglich durch makroskopische Messungen der Dichte nachzuweisen ist (die Dichteänderungen, die man von einem Punkt zum anderen in einem Raum feststellen kann, haben ihre Ursachen völlig unabhängig von diesen statistischen Fluktuationen; es sind dies : Inhomogenität von Temperatur und Druck, Einfluß der Gewichtskraft, ...).

### 3. Statistische Mechanik und Quantenmechanik

Ein Gesichtspunkt ist in dieser Diskussion der Verknüpfungen zwischen dem Mikroskopischen und dem Makroskopischen noch aufzuhellen. Als im Abschnitt 2 von der Bewegung der Moleküle eines Gases die Rede war, haben wir angenommen, daß sie durch die Newtonsche Mechanik korrekt beschrieben wird. Das ist in der Tat der Fall, wie es die Ergänzung I.A zeigt. Aber diese Situation ist sehr speziell : Generell weiß man, daß es angebracht ist, *auf mikroskopischem Niveau die Quantenmechanik zu benutzen*. Wir werden das im Folgenden tun, außer in den Sonderfällen, die dem Vorhergehenden analog sind. Die Kenntnisse der Quantenmechanik, die wir wirklich benötigen, bleiben auf sehr elementarem Niveau. Sie werden im Anhang II wiederholt. Ein wesentliches Charakteristikum der Quantenmechanik ist, daß sie selbst nur Vorhersagen von probalistischer Art macht. Es ist dennoch angebracht, zu sehen, daß dieses neuerliche Eingreifen der Wahrscheinlichkeiten grundsätzlich in seinem Ursprung von dem verschieden ist, was wir weiter oben diskutiert haben. In der Tat *ist die Quantenbeschreibung eines Sy-*

stems, selbst eines mikroskopischen, *probalistisch in seinem Prinzip* : Selbst wenn man über ein System jede Information besäße, die theoretisch zu erhalten möglich wäre (durch die gleichzeitige Messung eines vollständigen Ensembles von vertauschbaren Observablen), würden die Vorhersagen, die die Quantenmechanik bezüglich der Ergebnisse zukünftiger Messungen liefert, die an diesem System durchgeführt werden können, in Wahrscheinlichkeitstermen ausgedrückt. Die statistische Mechanik kommt anschließend dazu, darauf ihre eigenen wahrscheinlichkeitstheoretischen Methoden zu überlagern, denn *es ist nicht möglich, den Quantenzustand eines makroskopischen Systems vollständig zu kennen*. Die Wahrscheinlichkeiten greifen also auf zwei Niveaus ein : Auf mikroskopischem Niveau, denn der wahre Zustand des Systems muß generell durch die Quantenmechanik beschrieben werden; auf makroskopischem Niveau, weil der wahre Zustand nicht bekannt ist und man nur über bruchstückhafte Kenntnisse über das, was sein kann, verfügt.

## II. Wahrscheinlichkeiten, Information und statistische Entropie

Die statistische Mechanik basiert auf der Verwendung von *Wahrscheinlichkeiten*, was durch die enorme Größe der Avogadro-Konstanten gerechtfertigt ist (§I.B.2).

Behandelt man mit der Technik der Wahrscheinlichkeiten die möglichen Resultate eines Experimentes, so kennt man nun aber im allgemeinen Fall nicht mit Gewißheit den Ausgang im voraus : Man besitzt nicht jede Information, die erforderlich wäre, um das Resultat dieses Experimentes exakt vorherzusagen<sup>8</sup>. Die *Informationstheorie*, die sich im Laufe der letzten dreißig Jahre entwickelt hat, lehrt, wie *die Information quantitativ gemessen* werden kann, nämlich mit Hilfe der *statistischen Entropie*, die wir hier definieren werden; diese Theorie, schon von Boltzmann entdeckt, ist Grundlage der *mikroskopischen Interpretation der thermodynamischen Entropie*.

Die Informationstheorie leitet den Ausdruck für die statistische Entropie aus (völlig einleuchtenden) Hypothesen über die Eigenschaften der Funktion ab, die zum Messen der Information gebraucht wird. Wir werden hier umgekehrt vorgehen : Wir setzen diesen Ausdruck voraus und erklären anschließend seine Haupteigenschaften, um somit zu zeigen, daß es legitim ist, anzugeben, daß die statistische Entropie als Maß für die Information

---

<sup>8</sup>Dieser Mangel an Information kann auf zufällig eingetretenen Umständen beruhen (wenn diese Umstände sich ändern, könnte man über jede zur exakten Vorhersage des Resultates notwendige Information verfügen), oder auf fundamentalen Prinzipien, wie es in der Quantenmechanik der Fall ist (§I.B.3 und Anhang II).

gelten kann. Die Ergänzung I.G gibt einige zusätzliche Präzisierungen zur Information und statistischen Entropie.

## A. Definition der statistischen Entropie

Wir betrachten eine diskrete endliche Menge<sup>9</sup> von möglichen Ereignissen  $e_m$  ( $m = 1, 2, \dots, M$ ) der entsprechenden Wahrscheinlichkeiten  $P_m$  ( $m = 1, 2, \dots, M$ ), die natürlich die Bedingungen (II.1) erfüllen :

$$0 \leq P_m \leq 1 \quad \forall m, \quad (\text{II.1.a})$$

$$\sum_{m=1}^M P_m = 1. \quad (\text{II.1.b})$$

Wenn eine dieser Wahrscheinlichkeiten  $P_{m_0}$  gleich 1 ist, sind alle anderen Null :  $e_{m_0}$  ist sicher, alle anderen sind unmöglich. In diesem Fall ist die Information über unser Experiment vollständig : Es wird der sichere Treffer  $e_{m_0}$  sein. Im Vergleich zu diesem sehr speziellen Fall entspricht jede andere Verteilung von Wahrscheinlichkeiten  $\{P_m\}$  einem gewissen *Mangel an Information*, den wir beziffern wollen. Dies ermöglicht die statistische Entropie  $S(P_1, P_2, \dots, P_m)$ , die definiert wird durch<sup>10</sup>

$$S(P_1, P_2, \dots, P_M) := -k \sum_{m=1}^M P_m \ln P_m, \quad (\text{II.2})$$

mit<sup>11</sup>

$$P_m \ln P_m = 0 \quad , \quad \text{wenn} \quad P_m = 0. \quad (\text{II.3})$$

<sup>9</sup>Wir erstrecken diese Definition auch auf den Fall einer Menge von möglichen unendlich zählbaren Ereignissen, unter der Bedingung, daß die Summe der Formel (II.2) konvergiert. Der Übergang zum Kontinuum stellt hier spezielle Probleme (vgl. Ergänzung I.G, §III) dar, aber wir werden im folgenden die Formel nur für den diskreten Fall benutzen.

<sup>10</sup>Die statistische Entropie erlaubt ebenso die Größe der in einer Operation gewonnenen oder verlorenen Information zu messen; dies erklärt, daß sie Grundlage der Theorie der Information und der Kommunikation ist.

<sup>11</sup> $\ln P_m$  ist nicht definiert für  $P_m = 0$ , aber

$$\lim_{P_m \rightarrow 0} P_m \ln P_m = 0,$$

was die Wahl von (II.3) rechtfertigt. Somit tragen die unmöglichen Ereignisse nicht zur statistischen Entropie bei. Diese Eigenschaft ist aus Gründen der Kohärenz notwendig : Man kann stets die Menge von möglichen Ereignissen einer beliebigen Anzahl von Ereignissen der Wahrscheinlichkeit Null hinzufügen; die mangelnde Information darf bei einer solchen Manipulation nicht zu einer Wertänderung führen.

Diese Definition enthält einen *positiven* Faktor  $k$ , den man passend wählen kann. In der statistischen Mechanik verwendet man für  $k$  die *Boltzmann-Konstante*

$$k = 1,38 \cdot 10^{-23} \text{ J/K}. \quad (\text{II.4})$$

$J$  bezeichnet hier das *Joule* und  $K$  das *Kelvin*, die Einheit der Temperatur. Die mit der Wahl von (II.4) erhaltene Entropie besitzt also eine Einheit (die von  $k$ ). Der Vorteil dieser Wahl wird später noch einleuchten: Die thermodynamische Entropie ist in gewissen Sonderfällen, die wir noch präzisieren werden, exakt identisch mit der so definierten statistischen Entropie.

## B. Haupteigenschaften

Es ist leicht, sich an Hand der Definition (II.2) davon zu überzeugen, daß die statistische Entropie alle die Eigenschaften besitzt, die man zu Recht von einem *Maß des Mangels an Information*, das man über ein System hat, erwartet.

### 1. Positivität und Symmetrie

Wir bemerken zuallererst, daß die erste Bedingung von (II.1), verifiziert durch die Wahrscheinlichkeiten  $P_m$ , beinhaltet, daß  $S$  *positiv definit* ist:

$$S \geq 0 \quad \forall \{P_m\}. \quad (\text{II.5})$$

Andererseits ist  $S$  eine *totalsymmetrische* Funktion seiner  $M$  Argumente: zum Beispiel

$$S(P_2, P_1, P_3, \dots, P_M) = S(P_1, P_2, P_3, \dots, P_M). \quad (\text{II.6})$$

### 2. Minimum

Die statistische Entropie erreicht ihr Minimum, das *Null* ist (vgl. Ungleichung (II.5)), wenn eine der Wahrscheinlichkeiten den Wert 1 hat, die anderen sind dann konsequenterweise Null:

$$S(0, 0, \dots, 1, \dots, 0) = 0. \quad (\text{II.7})$$

In der Tat ist der hier betrachtete Fall der gleiche, wie wir ihn am Anfang von Abschnitt A als Träger einer vollständigen Information dargestellt haben, und es ist daher natürlich, daß der Mangel an Information hier den Wert Null annimmt.

### 3. Maximum (feste Anzahl möglicher Ereignisse)

Es läßt sich leicht zeigen<sup>12</sup>, daß für ein festes  $M$  die Funktion  $S(P_1, P_2, \dots, P_M)$  ein Maximum ergibt, wenn die  $M$  Ereignisse gleich wahrscheinlich sind, das heißt, wenn

$$P_1 = P_2 = \dots = P_M = \frac{1}{M} \quad (\text{II.8})$$

(die letzte Gleichheit stammt aus der Normierungsbedingung (II.1)). Dieses Maximum hat den Wert

$$S_{\max} = k \ln M, \quad (\text{II.9})$$

---

<sup>12</sup>Um das Maximum der Funktion  $S(P_1, P_2, \dots, P_M)$  für ein festes  $M$  zu suchen, muß man berücksichtigen, daß die  $M$  Variablen  $P_m$  nicht voneinander unabhängig sind, denn sie müssen die Normierungsbedingung in der zweiten Bedingung von (II.1) erfüllen. Ist eine einfache Beziehung gegeben, so kann man eine der  $M$  Variablen eliminieren, indem man sie als Funktion der  $(M - 1)$  anderen, die nunmehr unabhängig voneinander sind, ausdrückt (Übung I.8). Aber es gibt eine allgemeine Methode, um diese Art von Problemen zu behandeln, es ist die Methode der Lagrange-Multiplikatoren (Anhang A.IV), und diese wenden wir hier an.

Wir suchen die Extrema der Funktion  $S(P_1, P_2, \dots, P_M)$  und wissen, daß die  $M$  Variablen durch die Relation

$$\sum_{m=1}^M P_m - 1 = 0, \quad (1)$$

verknüpft sind. Die Methode von Lagrange besteht nun darin, die Extrema der Funktion

$$\tilde{S}_a(P_1, P_2, \dots, P_M) := S(P_1, P_2, \dots, P_M) - a \left[ \sum_{m=1}^M P_m - 1 \right] \quad (2)$$

derart zu suchen, als wenn die  $M$  Variablen unabhängig voneinander wären;  $a$  ist ein Parameter ( $\leftarrow$  Multiplikator  $\rightarrow$ ), der anschließend bestimmt wird, indem man die gefundene Lösung der Bedingung (1) unterwirft. Man kann zeigen, daß diese Lösung sehr wohl die des Anfangsproblems ist (siehe Anhang A.IV).

Leiten wir nunmehr die Funktion (2) nach jeder ihrer Variablen ab; unter Berücksichtigung des Ausdrucks (II.2) für  $S$  erhalten wir

$$\frac{\partial}{\partial P_n} \tilde{S}_a = -k \left[ \ln P_n + 1 + \frac{a}{k} \right]; \quad n = 1, 2, \dots, M. \quad (3)$$

Null setzen dieser  $M$  Ableitungen ergibt sofort:

$$\ln P_n = - \left( 1 + \frac{a}{k} \right) \quad (4)$$

Daraus leitet man sofort ab, daß die  $M$  Wahrscheinlichkeiten untereinander gleich sind. Wendet man schließlich die Beziehung (1) an, findet man genau das Ergebnis (II.8). Diese Überlegung zeigt nur, daß die Verteilung der Wahrscheinlichkeiten (II.8) zu einem Extremum in  $S$  führt. Da aber dieses Extremum das einzige ist, und wir Verteilungen kennen, die  $S$  kleiner als den entsprechenden Wert (II.9) machen (siehe Abschnitt 2), kann es sich nur um ein Maximum handeln.

wie man sofort sieht, wenn man die Werte (II.8) in die Definition (II.2) einsetzt.

Tatsächlich ist im Fall, in dem alle Wahrscheinlichkeiten gleich sind, der Mangel an Information am größten : Es ist nicht mehr möglich anzugeben, welches Ereignis die größere Wahrscheinlichkeit besitzt.

#### 4. Variation mit der Anzahl möglicher Ereignisse

Um die Variation der statistischen Entropie als Funktion der Anzahl  $M$  möglicher Ereignisse zu charakterisieren, muß man eine besondere Verteilung von Wahrscheinlichkeiten wählen, die den Vergleich zwischen verschiedenen Werten von  $M$  nicht verfälscht. Die gleichmäßige Verteilung (II.8), die wir im vorhergehenden Abschnitt 3 erhalten haben, erfüllt diese Forderungen, und darum ist es interessant festzustellen, daß *der maximale Wert der statistischen Entropie mit der Anzahl möglicher Ereignisse wächst*, wie es die Formel (II.9) zeigt.

Auch hier noch geht diese Eigenschaft mit der intuitiven Idee konform, die man sich von einer Messung des Mangels an Information macht : Wenn man mit einer Menge von gleichwahrscheinlichen Ereignissen konfrontiert wird, ist die Information, die man über das System besitzt, offensichtlich um so geringer, je größer die Anzahl der möglichen Ereignisse ist.

#### 5. Additivität

Wir betrachten zwei voneinander *unabhängige* Mengen möglicher Ereignisse,

$$\{e_m; m = 1, 2, \dots, M\}; \{e'_{m'}; m' = 1, 2, \dots, M'\}, \quad (\text{II.10})$$

und die entsprechenden Wahrscheinlichkeiten

$$\{P_m; m = 1, 2, \dots, M\}; \{P'_{m'}; m' = 1, 2, \dots, M'\}, \quad (\text{II.11})$$

geeigneterweise normiert :

$$\sum_{m=1}^M P_m = 1; \quad \sum_{m'=1}^{M'} P'_{m'} = 1. \quad (\text{II.12})$$

Die möglichen Ergebnisse eines « doppelten » Experimentes, in dem eines der Ereignisse  $e_m$  und eines von  $e'_{m'}$  eintritt, haben die Wahrscheinlichkeiten (nach dem Gesetz (15) im Anhang A.I)

$$P_{m,m'} := P(e_m \text{ und } e'_{m'}) = P_m \cdot P'_{m'}. \quad (\text{II.13})$$

Nach der Definition (II.2) wird der Mangel an Information bezüglich des doppelten Experimentes durch die statistische Entropie

$$\begin{aligned} S(P_{m,m'}; m = 1, 2, \dots, M; m' = 1, 2, \dots, M') \\ = -k \sum_{m=1}^M P_m \sum_{m'=1}^{M'} P_{m,m'} \ln P_{m,m'} \end{aligned} \quad (\text{II.14})$$

gemessen. Wenn man in diesen Ausdruck die Werte (II.13) der zusammengesetzten Wahrscheinlichkeiten  $P_{m,m'}$  einsetzt, erhält man

$$\begin{aligned} S(P_{m,m'}) &= -k \sum_{m=1}^M \sum_{m'=1}^{M'} P_m P'_{m'} \ln P_m P'_{m'} \\ &= -k \sum_{m=1}^M P_m \sum_{m'=1}^{M'} P'_{m'} [\ln P_m + \ln P'_{m'}] \\ &= -k \left\{ \sum_{m'=1}^{M'} P'_{m'} \sum_{m=1}^M P_m \ln P_m + \sum_{m=1}^M P_m \sum_{m'=1}^{M'} P'_{m'} \ln P'_{m'} \right\} \\ &= S(P_m) + S(P'_{m'}) \end{aligned} \quad (\text{II.15})$$

(wir haben für die statistische Entropie die vereinfachte Schreibweise  $S(P_m)$  eingeführt, und im letzten Schritt die Normierungsbedingungen (II.12) benutzt).

*Der Mangel an Information über ein doppeltes Experiment ist somit die Summe der Informationsmängel über jede der beiden unabhängigen Mengen von Ereignissen.*

### III. Zustände eines makroskopischen Systems

Die Wahrscheinlichkeiten, die wir kurz im Anhang I angeben, werden in der statistischen Mechanik benutzt, um den makroskopischen Zustand von Systemen zu beschreiben. Wir betrachten zuerst den allgemeinen Fall, bei dem sich die Beschreibung des Systems auf die Quantenmechanik gründet (§A), und führen den Begriff der « Zustandsdichte » ein (§B); der Abschnitt C untersucht anschließend den Fall, in dem (auch mikroskopische) Eigenschaften des Systems gut durch die klassische Newtonsche Mechanik beschrieben werden.

## A. Quantenbeschreibung des Systems

### 1. Der mikroskopische Zustand : Ein « reiner Zustand »

Man lernt in der elementaren Quantenmechanik (siehe Anhang II), daß der Zustand eines beliebigen Systems durch einen Ket-Vektor charakterisiert wird, der zum Raum der Zustände des Systems gehört. Dieses Ket  $|\psi(t)\rangle$  hängt von der Zeit ab, und seine Entwicklung im Laufe der Zeit wird durch die Schrödinger-Gleichung gegeben

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = H |\psi(t)\rangle, \quad (\text{III.1})$$

in der  $H$  der Hamiltonoperator des betrachteten Systems ist.

Wenn der Hamiltonoperator  $H$  zeitunabhängig ist (was sehr häufig der Fall ist), sagt man, daß man es mit einem *konservativen System* zu tun hat. Die zeitliche Entwicklung eines beliebigen Zustandes eines Systems wird nun sehr einfach erhalten, wenn man zuvor *die Eigenwertgleichung des Hamiltonoperators*

$$H |\varphi_\ell\rangle = E_\ell |\varphi_\ell\rangle. \quad (\text{III.2})$$

löst. In dieser Gleichung sind die Eigenwerte  $E_\ell$  die *möglichen Energiewerte* des Systems. Die zugehörigen Eigenkets  $|\varphi_\ell\rangle$  werden « *stationäre Zustände* » des Systems genannt : Es läßt sich leicht zeigen (Anhang II, §II.2), daß das System, wenn es sich zu einem gegebenen Zeitpunkt  $t_0$  in einem solchen Zustand  $|\varphi_\ell\rangle$  befindet, darin unbegrenzt verweilt und seine Eigenschaften von der Zeit unabhängig werden. Diese Betrachtungen sind gültig für jedes beliebige System, ob groß oder klein. Die Zustände, die durch ein Ket des Zustandsraumes dermaßen charakterisiert sind, werden « *reine Zustände* » genannt (im Gegensatz zu der statistischen Mischung, die wir im Abschnitt 2 betrachten werden). Erinnern wir uns : Um ein System in einem reinen Zustand zu präparieren, muß man im allgemeinen an diesem System eine gleichzeitige Messung eines « vollständigen Ensembles von Observablen, die kommutieren » (C.D.L., S. 235), durchführen : Die Gesamtheit der in einem solchen Experiment erhaltenen Resultate genügt, um *das* Ket (auf einen Faktor ohne physikalische Bedeutung) genau zu bestimmen, das den Zustand des Systems nach der Messung beschreibt.

Unterstreichen wir einen wichtigen Punkt. Ein möglicher Wert  $E_\ell$  der Energie ist im allgemeinen « *entartet* », das heißt, daß im allgemeinen mehrere voneinander linear unabhängige Eigenvektoren existieren, die dem gleichen Eigenwert  $E_\ell$  in der Gleichung (III.2) entsprechen. Die Anzahl  $g(E_\ell)$  dieser unabhängigen mit der gleichen Energie  $E_\ell$  assoziierten Eigenkets wird « *Entartungsgrad des Niveaus  $E_\ell$*  » genannt (für  $g(E_\ell) = 1$  ist das Niveau

$E_\ell$  « nicht entartet »). In den Lehrbüchern der Quantenmechanik fügt man des öfteren einen zusätzlichen Index an die Eigenkets, um dieser Tatsache Rechnung zu tragen (Anhang II, §II.2) :

$$H |\varphi_\ell^i\rangle = E_\ell |\varphi_\ell^i\rangle; \quad i = 1, 2, \dots, g(E_\ell). \quad (\text{III.3})$$

Wir verwenden hier eine *andere Konvention*, um die Schwerfälligkeit infolge von Mehrfachindizierungen zu vermeiden : *Der Index  $\ell$  kennzeichnet die stationären Zustände*; wenn mehrere Zustände  $|\varphi_\ell\rangle$  zur gleichen Energie gehören, dann bedeutet das, daß mehrere Eigenwerte  $E_\ell$  gleich sind; zum Beispiel, wenn  $|\varphi_1\rangle$  und  $|\varphi_2\rangle$  Eigenkets von  $H$  mit den gleichen Eigenwerten  $E$  sind, so schreibt man ebenfalls

$$\begin{aligned} H |\varphi_1\rangle &= E_1 |\varphi_1\rangle, \\ H |\varphi_2\rangle &= E_2 |\varphi_2\rangle, \end{aligned} \quad (\text{III.4})$$

weiß aber zugleich, daß

$$E_1 = E_2 = E. \quad (\text{III.5})$$

Für ein makroskopisches System, bezeichnet man den « *mikroskopischen Zustand* » (manchmal sagt man auch « *Mikrozustand* » ) als den reinen Quantenzustand, in dem es sich zu einem bestimmten Zeitpunkt befindet. Wie wir bereits im Abschnitt I.B.2 im Rahmen der Newtonschen Mechanik darlegten, ist es unmöglich und im Übrigen ohne Interesse, den mikroskopischen Zustand eines Systems von makroskopischer Größe zu kennen : Die Anzahl der Observablen, die man gleichzeitig messen müßte, wäre kaum weniger groß als die der Parameter, die den Zustand des Systems in der klassischen Mechanik definieren (vgl. Formel (I.2)). Man wird sich also mit einer (sehr) unvollständigen Information über den Zustand des Systems zufrieden geben, was sich in der Verwendung von Wahrscheinlichkeiten in der Beschreibung dieser nur unvollständig bekannten Zustände ausdrückt (siehe Abschnitt I.B.3 hier oben).

## 2. Der makroskopische Zustand : Eine « statistische Mischung »

### a. Definition einer statistischen Mischung von reinen Zuständen.

Um Situationen behandeln zu können, in denen der (reine) Zustand eines Systems unbekannt ist, man aber eine Information von probabilistischer Art über ihn besitzt, wird man dazu gebracht, den Begriff Quantenzustand zu verallgemeinern : Man sagt, daß der Zustand des Systems eine *statistische Mischung* der reinen Zustände  $|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_m\rangle, \dots$  mit den Wahrscheinlichkeiten  $P_1, P_2, \dots, P_m, \dots$  ist, wenn bekannt ist, daß das System sich mit der Wahrscheinlichkeit  $P_1$  im Zustand  $|\psi_1\rangle$ , mit der Wahrscheinlichkeit  $P_2$  im

Zustand  $|\psi_2\rangle, \dots$  befindet. Wohlgermerkt, diese Wahrscheinlichkeiten müssen die gewöhnlichen Bedingungen erfüllen :

$$0 \leq P_m \leq 1 \quad \forall m \quad \text{und} \quad \sum_m P_m = 1. \quad (\text{III.6})$$

Wir wollen nun genau angeben, wie man die physikalischen Vorhersagen bezüglich der Eigenschaften des Systems konkret berechnet, wenn sein Zustand eine solche statistische Mischung ist. Nehmen wir zum Beispiel an, daß man beabsichtigt, eine physikalische Größe  $A$  zu messen, und daß man die Wahrscheinlichkeit wissen will, bei dieser Messung den Wert  $a$  dieser Größe zu erhalten; wir nehmen zur Vereinfachung den Eigenwert  $a$  als nicht entartet zu  $A$  an und schreiben den entsprechenden Eigenvektor mit  $|u\rangle$ . Unsere Überlegungen sind nun die folgenden : Wenn das System im reinen Zustand  $|\psi\rangle$  ist, dann wäre die gesuchte Wahrscheinlichkeit (wenn wir dafür sorgen,  $|\psi\rangle$  und  $|u\rangle$  auf 1 zu normieren)

$$P_m(a) = |\langle u | \psi_m \rangle|^2; \quad (\text{III.7})$$

da dieses mögliche Ereignis (das System ist im reinen Zustand  $|\psi_m\rangle$ ) die Wahrscheinlichkeit  $P_m$  hat, muß man (III.7) mit  $P_m$  multiplizieren, um die Wahrscheinlichkeit dafür zu erhalten, daß sich das System im Zustand  $|\psi_m\rangle$  befindet *und* daß man  $a$  in der Messung findet; schließlich erhält man die gesamte Wahrscheinlichkeit dafür, daß man  $a$  findet (das System sei in einem beliebigen reinen Zustand), durch Summation über alle möglichen Zustände  $|\psi_m\rangle$  :

$$P(a) = \sum_m P_m P_m(a) = \sum_m P_m |\langle u | \psi_m \rangle|^2. \quad (\text{III.8})$$

Man erhält auf analoge Weise den Mittelwert  $\bar{A}$  einer physikalischen Größe  $A$  in der Situation, wo der Zustand des Systems eine statistische Mischung ist : In dem Zustand  $|\psi_m\rangle$  wäre der Mittelwert von  $A$  (den wir zur Vereinfachung mit  $A_m$  bezeichnen)

$$A_m = \langle \psi_m | A | \psi_m \rangle; \quad (\text{III.9})$$

der globale Mittelwert  $\bar{A}$  wird wie oben erhalten, indem man dieses Ergebnis mit der Wahrscheinlichkeit  $P_m$  der entsprechenden möglichen Ereignisse gewichtet und über alle diese möglichen Ereignisse summiert :

$$\bar{A} = \sum_m P_m A_m. \quad (\text{III.10})$$

Wenn der Zustand eines Systems ein statistisches Gemisch ist, braucht man also einfach nur die *Mittelwerte über die statistische Verteilung*  $\{P_m\}$

der reinen Zustände zu bilden, gemäß den im Anhang I erläuterten Techniken, also Mittelwerte der verschiedenen Eigenschaften, Größen und Quantitäten, die man für jeden reinen Zustand berechnen kann.

**b. Charakterisierung des makroskopischen Zustandes.** Die statistischen Mischungen von reinen Zuständen, die wir soeben im Abschnitt 1 eingeführt haben, sind sehr nützlich, sogar für Systeme von mikroskopischer Größe. Aber in Folge der Gründe, die wir am Ende des Abschnitts 1 oben angegeben haben, sind sie zur Quantenbeschreibung von Systemen makroskopischer Größe unverzichtbar: Der Quantenzustand eines solchen Systems ist notwendigerweise eine statistische Mischung von reinen Zuständen (das heißt von mikroskopischen Zuständen, siehe Abschnitt 1), die man seinen « *makroskopischen Zustand* » oder « *Makrozustand* » nennt.

Der am besten an die Charakterisierung und an die Handhabung der statistischen Mischungen angepaßte Formalismus ist der der Dichteoperatoren, der eingehend in der Ergänzung I.H beschrieben wird. Im Rahmen dieses Formalismus läßt sich am Natürlichsten die statistische Mechanik entwickeln. Wir werden sie allerdings aus didaktischen Gründen hier nicht gebrauchen: Wir wählen eine weniger allgemeine Methode, die aber viel einfacher und intuitiver ist, und behalten uns für einige Ergänzungen die abstrakteren Betrachtungen vor, die auf den Dichteoperatoren beruhen.

Im Folgenden charakterisieren wir daher den makroskopischen Zustand eines Systems als eine *statistische Mischung von reinen stationären Zuständen*. Um noch genauer zu sein, setzen wir die Eigenzustände  $|\varphi_\ell\rangle$  des Hamiltonoperators  $H$  des Systems und die entsprechenden Energien  $E_\ell$  (Gleichung (III.2)) als bekannt voraus. Zur Vereinfachung *bezeichnen wir die stationären Zustände  $|\varphi_\ell\rangle$  einfach mit  $(\ell)$* <sup>13</sup>. Wir betrachten für diese Zustände  $(\ell)$  nur mögliche mikroskopische Zustände des Systems. Von nun an wird ein makroskopischer Zustand durch die *Angabe des Ensembles der Wahrscheinlichkeiten  $P_\ell$*  charakterisiert sein, da das System sich in jedem der möglichen stationären Zustände  $(\ell)$  befinden kann<sup>14</sup>. Der Grund, warum diese Beschreibung in der Tat ausreicht, zumindest auf dem elementaren Ni-

<sup>13</sup>Das Symbol  $(\ell)$  repräsentiert im allgemeinen einen Satz von mehreren « Quantenzahlen », die zur vollständigen Spezifizierung eines Eigenzustandes von  $H$  notwendig sind. Für ein makroskopisches System ist ihre Anzahl riesig (vgl. §B.2.b.α).

<sup>14</sup>Wir werden später feststellen (vgl. §IV.A.2 hier unten und Teil IV, §I.A), daß die Zustände  $(\ell)$ , die wir hier benutzen, als Eigenzustände des Hamiltonoperators  $H$  strenggenommen nicht stationär sind. Dies ist darauf zurückzuführen, daß es für ein System von makroskopischer Größe unmöglich ist, seinen exakten Hamiltonoperator anzugeben; der Operator  $H$ , der uns dazu dient die Zustände  $(\ell)$  zu definieren, ist gezwungenermaßen ein approximativer Hamiltonoperator, und seine Eigenzustände  $(\ell)$  sind nur näherungsweise stationär.

veau, auf dem wir uns hier befinden, folgt aus der entscheidenden Rolle, die die Energie als physikalische Größe in der statistischen Mechanik ebenso wie in der mikroskopischen Physik spielt<sup>15</sup>.

## B. Dichte von (mikroskopischen) Zuständen

Für ein System von makroskopischer Größe genügt es oftmals, die Anzahl der mikroskopischen Zustände ( $\ell$ ) zu einer bestimmten Energie zu kennen. Diese Anzahl wird von der « Zustandsdichte » des Systems geliefert. Wir werden jetzt die Zustandsdichte definieren und dann an einfachen Beispielen untersuchen (diese Untersuchung wird in den Ergänzungen I.E und I.F weiter geführt).

### 1. Definition und Anwendung der Zustandsdichte

**a. Summation über die Zustände und Summation über die Energien.** Wir haben den Abschnitt A abgeschlossen, indem wir zeigten, daß die Eigenschaften eines Systemes makroskopischer Größe durch Mittelwertbildung über die möglichen mikroskopischen Zustände erhalten werden. Im konkreten Fall müssen die *Summen über die stationären Zustände* ( $\ell$ ) in der Form

$$\bar{A} = \sum_{(\ell)} P_{\ell} A_{\ell}, \quad (\text{III.11})$$

berechnet werden, wobei  $P_{\ell}$  die Wahrscheinlichkeit dafür darstellt, daß das System sich in dem mikroskopischen Zustand ( $\ell$ ) befindet und  $A_{\ell}$  der Wert ist, den in diesem Zustand die Größe, die uns interessiert, annimmt. Wir weisen noch einmal auf die Notwendigkeit hin, *nicht Energieniveaus und stationäre Zustände zu verwechseln*: Es gibt im allgemeinen mehrere unterscheidbare mikroskopische Zustände ( $\ell$ ) (beschrieben durch linear unabhängige Kets), die an die gleiche Energie geknüpft sind. Ursprünglich geht die Summation in einem Ausdruck der Form (III.11) über die Zustände ( $\ell$ ).

Ziemlich oft jedoch hängt der aufzusummierende Term tatsächlich nur von der Energie  $E_{\ell}$  des Zustandes ( $\ell$ ) ab. Wir nehmen an, daß dies der Fall ist, und setzen

$$P_{\ell} A_{\ell} = f(E_{\ell}). \quad (\text{III.12})$$

<sup>15</sup>Die Eigenzustände ( $\ell$ ) des Hamiltonoperators einzeln zu betrachten, wie wir es machen, läuft darauf hinaus, sie als Basis im Zustandsraum des Systems zu wählen. Die Bedeutung der Rolle der Energie in der Physik zeigt sich daran, daß die Dichteoperatoren in der statistischen Mechanik in einer solchen Basis diagonal sind, das heißt, daß die stationären Zustände ( $\ell$ ) auf alle Fälle bevorzugt werden.

Wir betrachten jetzt einen bestimmten Wert der Energie  $E_\ell$ . Wenn  $g(E_\ell)$  die Anzahl der unterscheidbaren Zustände ( $\ell$ ) ist, die mit ihr verknüpft sind (Grad der Entartung), dann enthält die Summe (III.11)  $g(E_\ell)$  unterscheidbare Terme, die gleich  $f(E_\ell)$  sind. Man kann also *die Summe über die Zustände ( $\ell$ ) durch eine Summe über die Energieniveaus ersetzen*, unter der Bedingung, den Faktor  $g(E_\ell)$ , den wir soeben erklärten, mit zu berücksichtigen; wir schreiben

$$\bar{A} = \sum_{(\ell)} P_\ell A_\ell = \sum_{E_\ell} g(E_\ell) f(E_\ell). \quad (\text{III.13})$$

Da die vereinfachte Schreibweise, die wir hier gebrauchen (ein einziger Index  $\ell$  für die stationären Zustände), eine gewisse Mehrdeutigkeit mit sich bringt, präzisieren wir, daß die Summation über  $E_\ell$  nur einen Term für jedes Niveau, das heißt für jeden möglichen Energiewert berücksichtigt : Wenn, wie in (III.5), zwei Energien  $E_\ell$  gleich sind, so entsprechen ihnen zwei Terme in der Summe über die Zustände ( $\ell$ ), aber ein einziger in der Summe über die Niveaus  $E_\ell$ .

**b. Übergang zu einem Integral über die Energie.** Obgleich groß, sind die von der statistischen Mechanik untersuchten Systeme *endlich*, das bedeutet, daß es immer einen « Kasten » von makroskopischem aber endlichem Volumen gibt, der das System vollständig umhüllt. Dieses hat zur Folge, wie man in der Quantenmechanik leicht zeigen kann, daß *die Energieniveaus* eines solchen Systems *ein diskretes Spektrum bilden*<sup>16</sup>. Deswegen ist in der Formel (III.13) die Summe über die Niveaus eine diskrete Summe  $\sum$ , ebenso wie über die Zustände.

Dennoch ist für ein System makroskopischer Dimension *der Unterschied zwischen zwei aufeinanderfolgenden Niveaus sehr gering* : Tatsächlich ist die Quantelung der Energie im makroskopischen Maßstab nicht direkt wahrnehmbar.

Betrachten wir nun ein Energieintervall  $\delta E$ , das die zwei folgenden Eigenschaften besitzt :

- (i) Es ist genügend klein, um *auf makroskopischem Niveau zu verschwinden*, das heißt

$$\delta E \ll \delta_0 E, \quad (\text{III.14})$$

<sup>16</sup>Diese Eigenschaft folgt direkt aus der Bedingung, daß die Aufenthaltswahrscheinlichkeit der Bestandteile des Systems (Moleküle eines Gases zum Beispiel) außerhalb des Kastens, der das System umhüllt, verschwinden muß : Wie im einfachen Fall des eindimensionalen Potentialtopfes hat die Eigenwertgleichung des Hamiltonoperators nur für solche diskrete Energiewerte Lösungen, die die Randbedingungen erfüllen (vgl. C.D.L., S. 350).

dabei ist  $\delta_0 E$  die Unsicherheit, mit der die Messungen der makroskopischen Energie behaftet sind;

- (ii) es ist genügend groß, um *eine große Anzahl mikroskopischer Niveaus*  $E_\ell$  zu enthalten, das heißt, daß es sehr groß ist im Vergleich zu dem Unterschied zwischen zwei aufeinanderfolgenden Niveaus :

$$\delta E \gg |E_{\ell+1} - E_\ell|. \quad (\text{III.15})$$

Diese zwei Forderungen widersprechen einander nicht und können gleichzeitig an einem System von makroskopischer Dimension verifiziert werden. Es sei  $\delta n(E)$  die *Anzahl der mikroskopischen Zustände* ( $\ell$ ), deren Energie  $E_\ell$  zwischen  $E$  und  $E + \delta E$  liegen. Per Definition ist die *Zustandsdichte*  $\rho(E)$  des betrachteten Systems so beschaffen, daß gilt

$$\delta n(E) := \rho(E) \delta E. \quad (\text{III.16.a})$$

Wohlverstanden, wir benutzen hier wie immer in vergleichbaren Fällen, die differentielle Schreibweise :

$$dn(E) := \rho(E) dE \quad (\text{III.16.b})$$

ist die Anzahl der Zustände ( $\ell$ ), deren Energie zwischen  $E$  und  $E + dE$  liegen.

Die Kenntnis der Zustandsdichte  $\rho(E)$  erlaubt, die Summen von der Form (III.13) zu ermitteln. Man erwartet in der Tat, daß für ein System von makroskopischer Größe die aufzusummierende Funktion  $f(E_\ell)$  (Formel (III.12)) verschwindend wenig zwischen zwei aufeinanderfolgenden Niveaus variiert, das heißt, daß sie praktisch den gleichen Wert  $f(E)$  für alle Niveaus, deren Energien zwischen  $E$  und  $E + dE$  liegen, annimmt. Der Beitrag zur Summe (III.13) des Ensembles dieser Niveaus ist daher  $f(E)$ , multipliziert mit der Anzahl der entsprechenden Zustände  $\rho(E)d(E)$ . Um über alle möglichen Niveaus zu summieren, muß man schließlich die Anteile aller Intervalle  $(E, E + dE)$  an  $E$  von seiner untersten Grenze  $E_0$  ( $\ll$  Grundniveau  $\gg$ ) bis  $+\infty$  aufsummieren, was per Definition darauf hinausläuft, ein *Integral über die Energie* zu berechnen :

$$\bar{A} := \sum_{(\ell)} P_\ell A_\ell = \int_{E_0}^{+\infty} f(E) \rho(E) dE. \quad (\text{III.17})$$

*Anmerkungen :*

- (i) Unterstreichen wir noch einmal, daß  $\rho(E)dE$  die Anzahl von *Zuständen* ist, deren Energie zwischen  $E$  und  $E + dE$  liegen, nicht die Anzahl der Niveaus.

- (ii) Es ist manchmal bequemer, zuerst die Anzahl  $\Phi(E)$  von Zuständen, deren Energie kleiner als  $E$  ist, zu berechnen. Wenn man die Funktion  $\Phi(E)$  kennt, dann ist daraus sofort die Zustandsdichte  $\rho(E)$  herleitbar, denn die Anzahl von Zuständen  $dn(E)$ , deren Energie zwischen  $E$  und  $E + dE$  liegen, beträgt

$$dn(E) = \Phi(E + dE) - \Phi(E). \quad (\text{III.18})$$

Die Zustandsdichte ist damit die Ableitung der Funktion  $\Phi(E)$  :

$$\rho(E) = \frac{d\Phi(E)}{dE}. \quad (\text{III.19})$$

- (iii) Es sollte nach dem Voranstehenden klar sein, daß die Formel (III.17) einen *Näherungswert* an die ursprüngliche Summe (III.11) (ergänzt durch die Bedingung (III.12)) darstellt. Für ein System von makroskopischer Dimension ist diese Näherung *in der Mehrzahl der Fälle ausgezeichnet*. Man stellt sich dennoch vor, daß sie ungeeignet sein könnte, wenn die Funktion  $f(E_\ell)$  nicht langsam genug in bestimmten Energiebereichen variiert, zum Beispiel für die allerersten Niveaus; in allen solchen Fällen wird die diskrete Summe (III.13) nicht gut durch das Integral (III.17) wiedergegeben. Wenn es bezüglich der Gültigkeit einer Näherungsformel, wie (III.17) Zweifel gibt, muß man auf den ursprünglichen Ausdruck (III.13) zurückgreifen.

## 2. Beispiel : Freie Teilchen in einem makroskopischen Kasten

Hier ist ein sehr einfaches Beispiel. Es erlaubt einige Punkte zu präzisieren, die klarer in einer konkreten Situation erfaßt werden, sowie eine Idee davon zu geben, wie die Zustandsdichte  $\rho(E)$  mit der Energie  $E$  variiert.

**a. Einteilchensystem.** Wir beginnen mit einem einfachen nicht makroskopischen System, das aus einem einzigen freien Teilchen (ohne Spin) besteht. Aber wir nehmen an, daß dieses Teilchen in einem Kasten von makroskopischer Dimension eingeschlossen ist, was uns erlaubt, eine Zustandsdichte zu definieren. Wir betrachten einen spatförmigen Kasten (Parallelepipet, auch Parallelfach) mit den Kantenlängen  $L_x$ ,  $L_y$ ,  $L_z$  und dem Volumen  $V = L_x \cdot L_y \cdot L_z$ .

Die Quantenuntersuchung dieses Systems ist im Anhang II (§II.4) zusammengefaßt. Wenn man den Wänden des Kastens periodische Randbedingungen auferlegt, so kann man als Funktionen von stationären Wellen des

betrachteten Teilchens die ebenen Wellen nehmen

$$\varphi_{\vec{K}}(\vec{r}) = \frac{1}{\sqrt{V}} e^{i\vec{K}\cdot\vec{r}}; \quad (\text{III.20})$$

$\hbar\vec{K}$  ist der *Impuls* des Teilchens im stationären Zustand  $\varphi_{\vec{K}}$ ; die zu diesem Zustand zugehörige Energie beträgt

$$E_{\vec{K}} = \frac{\hbar^2 \vec{K}^2}{2m}. \quad (\text{III.21})$$

*Die möglichen Impulse sind gequantelt*: Die allein unter diesen periodischen Bedingungen erlaubten Vektoren  $\vec{K}$  sind von der Form

$$\vec{K} = l_x \frac{2\pi}{L_x} \vec{e}_x + l_y \frac{2\pi}{L_y} \vec{e}_y + l_z \frac{2\pi}{L_z} \vec{e}_z, \quad (\text{III.22})$$

wobei die Parameter  $l_x, l_y, l_z$  alle ganzzahligen Werte annehmen. Der Index, den wir im allgemeinen Fall ( $\ell$ ) nennen, ist hier nun ein Satz von drei ganzzahligen Indizes:

$$(\ell) \equiv (l_x, l_y, l_z). \quad (\text{III.23})$$

Die Energien (III.21) bilden ein *diskretes Spektrum*<sup>17</sup>. Aber man stellt auch leicht fest, daß die Unterschiede zwischen aufeinanderfolgenden Niveaus umso geringer sind, je größer die Abmessungen  $L_x, L_y, L_z$  des Kastens sind; für makroskopische Maße *verschwinden diese Unterschiede vollständig*<sup>18</sup>. Dieses erlaubt, die Zustandsdichte zu definieren und zu berechnen. Nach

<sup>17</sup>Dies ist auf die Einschränkungen über die möglichen Werte der Parameter  $l_x, l_y, l_z$  zurückzuführen, und diese Einschränkungen folgen ihrerseits aus den Randbedingungen: In dem Abschnitt II.4 des Anhangs II verlangen die Gleichungen (26) nur, daß  $u, v, w$  imaginäre Exponentialfunktionen sind, und erst die Bedingungen (29) und (35) legen die den Wellenvektoren erlaubten Werte fest. Man verifiziere daher konkret an diesem Beispiel die allgemeine Behauptung, die am Anfang von Abschnitt I.b aufgestellt wurde.

<sup>18</sup>Für ein Elektron ( $m \simeq 10^{-30}$  kg), eingeschlossen in einem Würfel mit der Kantenlänge  $L = 1$  cm, beträgt der Unterschied zwischen zwei aufeinanderfolgenden Niveaus

$$E_{l_x+1, l_y, l_z} - E_{l_x, l_y, l_z} = (2l_x + 1) \frac{2\pi^2 \hbar^2}{mL^2} \quad (1)$$

ungefähr  $10^{-33}$  J für ein  $l_x$  der Größenordnung 1. Wir werden in (III, §V.C) sehen, daß bei hohen Temperaturen  $T$  die Größenordnung der ganzzahligen  $\ell$  ( $l_x, l_y$  oder  $l_z$ ) durch

$$\frac{2\pi^2 \hbar^2}{mL^2} \ell^2 \lesssim \frac{1}{2} kT. \quad (2)$$

gegeben ist. Für solche Werte von  $l_x$  sieht man, daß gilt

$$E_{l_x+1, l_y, l_z} - E_{l_x, l_y, l_z} \lesssim \left[ kT \frac{4\pi^2 \hbar^2}{mL^2} \right]^{1/2} \quad (3)$$

(III.21) bedeutet eine feste Energie auch einen festen Betrag (Modul) des Vektors  $\vec{K}$ . Die Zustände  $(l_x, l_y, l_z)$ , deren Energie zwischen  $E$  und  $E + dE$  liegt, zu berechnen, bedeutet daher, die erlaubten Vektoren  $\vec{K}$  zu berechnen, deren Betrag zwischen  $K$  und  $K + dK$  liegt, mit

$$K = \sqrt{\frac{2mE}{\hbar^2}}, \quad (\text{III.24.a})$$

$$dK = \frac{1}{2} \sqrt{\frac{2m}{\hbar^2}} E^{-1/2} dE. \quad (\text{III.24.b})$$

Dieses Problem läßt sich leicht<sup>19</sup> durch geometrische Betrachtungen im

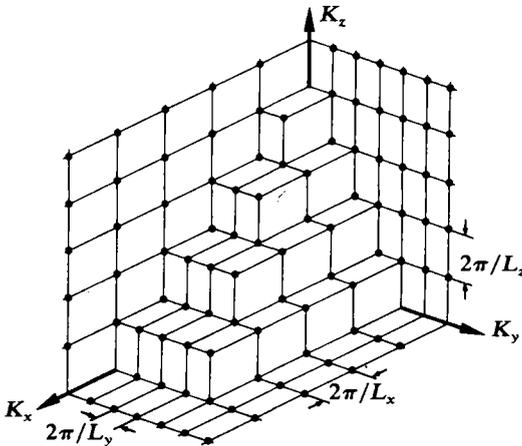


ABBILDUNG 1

Geometrische Konstruktion der « erlaubten » Endpunkte der Vektoren  $\vec{K}$ , die die stationären Zustände kennzeichnen.

dreidimensionalen Raum der Vektoren  $\vec{K}$  lösen. Man findet, daß die gesuchte

Selbst bei Temperaturen von mehreren tausend Kelvin übersteigt diese Energiedifferenz nicht  $10^{-16}$  eV.

Man beachte, daß die Werte für Teilchen von größerer Masse (Atome, Moleküle, ...) noch viel kleiner wären.

<sup>19</sup>Die Endpunkte der Vektoren  $\vec{K}$ , die uns interessieren, liegen in einem der Knoten eines dreidimensionalen Netzes (Abb. 1), das aus Elementarzellen aufgebaut ist, deren Seiten parallel zu den Achsen sind mit den Maßzahlen  $2\pi/L_x$ ,  $2\pi/L_y$  und  $2\pi/L_z$ . Wir wollen die Anzahl der Vektoren bestimmen, für die diese Endpunkte in dem Intervall zwischen zwei Kugeln liegen mit dem Ursprung als Mittelpunkt und den Radien  $K$  und  $K + dK$ , und die gemäß (III.24) gegeben sind. Für die makroskopischen Werte von  $L_x$ ,  $L_y$ ,  $L_z$  sind die Abmessungen der Elementarzellen sehr klein gegenüber der betrachteten Zunahme  $dK$ ,

Zahl, die per Definition gleich  $\rho(E)dE$  ist, den Wert

$$\rho(E)dE = \frac{V}{2\pi^2} K^2 dK, \quad (\text{III.25})$$

hat, was unter Berücksichtigung der Relation (III.24)

$$\rho(E) = \frac{V\sqrt{2}}{2\pi^2\hbar^3} m^{3/2} E^{1/2}. \quad (\text{III.26})$$

ergibt.

Wir erklären im Anhang II (§II.4.c.α), warum es bequemer ist, periodische Randwertbedingungen gegenüber « konstanten » (oder starren) Randwerten zu gebrauchen. Es läßt sich ohne Schwierigkeiten zeigen<sup>20</sup>, daß das

und zwar derart, daß die gesuchte Zahl sehr groß ist. Man kann sie nun auf folgende Art erhalten (auf einige Stellen genau, was angesichts der Größe der Zahl selbst ohne Bedeutung ist).

Die kugelförmige Schale zwischen den Kugeln mit den Radien  $K$  und  $dK$  besitzt im Raum der  $\vec{K}$  das Volumen

$$4\pi K^2 dK. \quad (1)$$

Die Anzahl der Elementarzellen, die in diesem Gebiet des Raumes der  $\vec{K}$  enthalten sind, ist nahezu gleich der Anzahl wie oft sein Volumen das der Elementarzelle enthält; diese Anzahl beträgt

$$\frac{2\pi}{L_x} \cdot \frac{2\pi}{L_y} \cdot \frac{2\pi}{L_z} = \frac{8\pi^3}{V}. \quad (2)$$

Man erhält so

$$\frac{4\pi K^2 dK}{8\pi^3/V}. \quad (3)$$

Dies sind die Knoten des Netzes, das heißt die Ecken der Elementarzellen, die wir zählen wollen. Nun besitzt zwar jede Zelle 8 Ecken, aber jede Ecke gehört zu 8 verschiedenen Zellen. Es gibt daher im Mittel auf eine große Anzahl von Zellen eine Ecke pro Zelle. (3) ist demzufolge in sehr guter Näherung die Zahl, die wir suchten.

<sup>20</sup>Im Fall konstanter Randbedingungen (Anhang II, §II.4.b) kann man auch die drei Zahlen  $(\ell_x, \ell_y, \ell_z)$  durch den Vektor  $\vec{K}$  ersetzen, der definiert wird durch

$$\vec{K} := \ell_x \frac{\pi}{L_x} \vec{e}_x + \ell_y \frac{\pi}{L_y} \vec{e}_y + \ell_z \frac{\pi}{L_z} \vec{e}_z; \quad (1)$$

der Ausdruck für die Energie ist dann

$$E_{\vec{K}} = \frac{\hbar^2 \vec{K}^2}{2m} \quad (2)$$

wie in (III.21), aber die Werte, die  $\ell_x$ ,  $\ell_y$  und  $\ell_z$  annehmen können, sind nur *ganze positive Zahlen* (Null ausgeschlossen).

Die Berechnung der Zustandsdichte ist dann analog zu der in Fußnote 19 mit den folgenden zwei Unterschieden: Einerseits hat die Elementarzelle im Vektorraum  $\vec{K}$  das Volumen  $(\pi/L_x) \cdot (\pi/L_y) \cdot (\pi/L_z)$ , das 8-mal kleiner ist als im vorhergehenden Fall, andererseits müssen die drei Komponenten der zulässigen Vektoren  $\vec{K}$  positiv sein; das Volumen

Ergebnis bezüglich der Zustandsdichte unverändert bleibt, wenn man in seinen Überlegungen von Bedingungen mit « konstanten » Randwerten ausgeht. Solange wir es mit freien, in einem Kasten eingeschlossenen Teilchen zu tun haben, verwenden wir daher *systematisch periodische Randbedingungen*.

**b. System von  $N$  freien und unabhängigen Teilchen.** Wir betrachten jetzt den gegenüber Abschnitt a realistischen Fall von  $N$  freien Teilchen (ohne Spin), die in einem Kasten von makroskopischer Größe eingeschlossen sind. Die  $N$  Teilchen sind untereinander identisch, wir nehmen den Kasten als spatförmig an (mit Kantenlängen  $L_x, L_y, L_z$ , Volumen  $V = L_x \cdot L_y \cdot L_z$ ).

*α. Zustände des Systems und Zustände eines jeden Teilchens.* Die  $N$  Teilchen sind *unabhängig*, wenn der Hamiltonoperator  $H$  des Systems einfach die Summe von  $N$  Termen ist, von denen jeder sich auf ein einziges Teilchen bezieht :

$$H = \sum_{i=1}^N H_i, \quad (\text{III.27})$$

darin hängt  $H_i$  nur von Observablen ab, die an das mit  $i$  nummerierte Teilchen geknüpft sind; *die Wechselwirkungsterme* zwischen den verschiedenen Teilchen würden von mehreren untereinander assoziierten Observablen abhängen und *sind hier Null*. In diesem Fall läßt sich in der Quantenmechanik zeigen (Anhang II, §III.2.a), daß *man die Eigenzustände von  $H$  durch ein (Tensor-)produkt aus den Eigenzuständen der  $N$  individuellen Hamiltonoperatoren  $H_i$  konstruieren kann* (man muß natürlich  $N$  individuelle Zustände betrachten, einen für jedes Teilchen), *da die entsprechenden Energien  $N$  individuelle Energien sind*.

Hier nehmen wir zusätzlich an, daß jedes Teilchen *frei* ist, was bedeutet, daß der Hamiltonoperator  $H_i$  sich einfach schreibt zu

$$H_i = \frac{\vec{P}_i^2}{2m}, \quad (\text{III.28})$$

worin  $\vec{P}_i$  der Impulsoperator des Teilchens ( $i$ ) ist. Wir können daher für jedes Teilchen die Ergebnisse vom obigen Abschnitt benutzen.

Aber es tritt noch eine *Schwierigkeit* auf : Für identische Teilchen muß das *Symmetrisierungspostulat* (Anhang II, §IV.2) auf die Zustände des Systems angewendet werden; dieses kompliziert beträchtlich das Problem, so

---

der Kugelschale zwischen den Kugeln mit den Radien  $K$  und  $dK$  muß daher durch 8 geteilt werden. Diese beiden Faktoren kompensieren sich und lassen daher das Ergebnis unverändert.

daß schließlich seine exakte Lösung unmöglich wird, zumindest in der Form, wie wir es dargestellt haben. Wir tun hier so, *als ob die  $N$  Teilchen unterschiedlicher Natur wären* und werden später auf die Schlußfolgerungen aus der Ununterscheidbarkeit der Teilchen zurückkommen<sup>21</sup>.

Ein stationärer Zustand des Systems ist nun durch die Größe der  $N$  gewöhnlichen Vektoren

$$\{\vec{K}_1, \vec{K}_2, \dots, \vec{K}_N\}, \quad (\text{III.29})$$

charakterisiert, dabei ist  $\hbar\vec{K}_i$  der Impulsoperator des Teilchens ( $i$ ), wenn wir periodische Randbedingungen wählen. Die Energie dieses Zustandes hat den Wert

$$E = \varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_N, \quad (\text{III.30})$$

mit

$$\varepsilon_i = \frac{\hbar^2 \vec{K}_i^2}{2m}, \quad (\text{III.31})$$

der Energie des Teilchens ( $i$ ). Die Komponenten des Vektors  $\vec{K}_i$  müssen ganzzahlige Vielfache von  $2\pi/L_x$ ,  $2\pi/L_y$  bzw.  $2\pi/L_z$  sein. Hierbei repräsentiert das Symbol ( $\ell$ ), das wir benutzen, um die mikroskopischen Zustände des Systems zu charakterisieren,  $3N$  Indizes, denn man benötigt 3 ganze Zahlen  $\ell_x, \ell_y, \ell_z$  um den individuellen Zustand eines jeden der  $N$  Teilchen zu kennzeichnen.

*$\beta$ . Zustandsdichte des Systems.* Wir wollen die stationären Zustände des Systems berechnen, deren Energie zwischen  $E$  und  $E + dE$  liegen. Die geometrische Methode, die uns im Abschnitt a gedient hat, kann auf einen  $3N$ -dimensionalen Vektorraum verallgemeinert werden (III.29); diese Rechnung wird in der Ergänzung I.E durchgeführt. Wir gehen hier jetzt anders vor und leiten die Zustandsdichte des Systems, die wir mit  $\rho_N(E)$  bezeichnen, aus der Zustandsdichte  $\rho(\varepsilon)$  eines einzigen Teilchens ab.

Beginnen wir damit, die Anzahl  $\Phi_N(E)$  von Zuständen des Systems zu berechnen, deren Energie kleiner als  $E$  ist. Eine einfache Überlegung<sup>22</sup> liefert

<sup>21</sup>Der Fall von Systemen mit identischen und nicht unterscheidbaren Teilchen wird zuerst näherungsweise in der Ergänzung II.C und im Abschnitt IV.A.3 von Teil III (Maxwell-Boltzmann-Approximation) behandelt, anschließend vollständiger im Teil V, Abschnitt IV (Quantenstatistik).

<sup>22</sup>Wenn man die Auswirkungen des Symmetrisierungspostulates außer acht läßt, so ist für ein System von  $N = 2$  unabhängigen Teilchen die Anzahl von Zuständen, in denen die Energie des Teilchens (1) zwischen  $\varepsilon_1$  und  $\varepsilon_1 + d\varepsilon_1$  und die des Teilchens (2) zwischen  $\varepsilon_2$  und  $\varepsilon_2 + d\varepsilon_2$  liegt, das Produkt

$$\rho(\varepsilon_1)d\varepsilon_1 \cdot \rho(\varepsilon_2)d\varepsilon_2 \quad (1)$$

der Anzahl der Zustände von jedem Teilchen, die die entsprechende Bedingung erfüllen :

den Ausdruck

$$\Phi_N(E) = \int_0^{+\infty} d\varepsilon_1 d\varepsilon_2 \dots d\varepsilon_N \rho(\varepsilon_1) \rho(\varepsilon_2) \dots \rho(\varepsilon_N) \theta(E - \varepsilon_1 - \varepsilon_2 - \dots - \varepsilon_N), \quad (\text{III.32})$$

worin  $\theta(x)$  die « Heaviside-Funktion » ist :

$$\theta(x) := \begin{cases} 0 & \text{für } x < 0; \\ 1 & \text{für } x > 0. \end{cases} \quad (\text{III.33})$$

Man weiß (Formel (III.19)), daß die gesuchte Zustandsdichte  $\rho_N(E)$  die Ableitung von  $\Phi_N(E)$  ist :

$$\rho_N(E) = \frac{d}{dE} \Phi_N(E). \quad (\text{III.34})$$

Nun tritt  $E$  in (III.32) aber nur in der Funktion  $\theta$  auf, deren Ableitung bekannt ist<sup>23</sup> :

$$\frac{d}{dx} \theta(x) = \delta(x), \quad (\text{III.35})$$

darin ist  $\delta$  die « Dirac-Funktion » . Man erhält schließlich

$$\rho_N(E) = \int_0^{+\infty} d\varepsilon_1 d\varepsilon_2 \dots d\varepsilon_N \rho(\varepsilon_1) \rho(\varepsilon_2) \dots \rho(\varepsilon_N) \delta(E - \varepsilon_1 - \varepsilon_2 - \dots - \varepsilon_N). \quad (\text{III.36})$$

Tatsächlich kann man wegen der Unabhängigkeit der beiden Teilchen jeden Zustand des Teilchens (1), dessen Energie zum Intervall  $(\varepsilon_1, \varepsilon_1 + d\varepsilon_1)$  gehört, mit einem beliebigen Zustand des Teilchens (2), dessen Energie zum Intervall  $(\varepsilon_2, \varepsilon_2 + d\varepsilon_2)$  gehört, verknüpfen, um somit einen Zustand des Systems zu konstruieren. In der Formel (1) erscheint die gleiche Zustandsdichte  $\rho(\varepsilon)$  für die beiden Teilchen, weil diese identisch sind. In der Formel (III.32) ist das Produkt

$$\rho(\varepsilon_1) d\varepsilon_1 \cdot \rho(\varepsilon_2) d\varepsilon_2 \dots \cdot \rho(\varepsilon_N) d\varepsilon_N \quad (2)$$

unter dem Integral ebenfalls die Anzahl der Zustände des Systems von  $N$  unabhängigen Teilchen (die so behandelt werden, als wenn sie nicht identisch wären), für die die Energie des Teilchens ( $j$ ) zwischen  $\varepsilon_j$  und  $\varepsilon_j + d\varepsilon_j$  liegt mit  $j = 1, 2, \dots, N$ . Um  $\Phi_N(E)$  zu erhalten, muß man alle Zahlen in (2) aufsummieren, die den Werten  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N$  entsprechen, so daß gilt

$$\varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_N < E, \quad (3)$$

denn die Energie eines Zustandes des Systems ist die Summe der einzelnen Energien der Teilchen (wenn letztere unabhängig sind). In der Formel (III.32) wird dies durchgeführt, indem über alle möglichen Werte für jedes  $\varepsilon_j$  (das von 0 für  $\vec{K}_i = \vec{0}$  bis Unendlich variiert) summiert wird, nachdem die Beiträge, die die Bedingung (3) erfüllen, mit 1, die Beiträge, die (3) verletzen, mit 0 multipliziert wurden.

<sup>23</sup>Siehe zum Beispiel C.D.L., Anhang II, Abschnitt 3.a .

Was uns im Prinzip interessiert, ist das Verhalten von  $\rho_N(E)$  als Funktion der Energie  $E$  des Systems. Dieses Verhalten kann durch folgende Überlegung erhalten werden.

In der Formel (III.36) stellt die Funktion  $\delta$  sicher, daß die Summe der  $\varepsilon_i$  gleich  $E$  ist. Man erwartet schließlich, daß jedes  $\varepsilon_i$  von der Größe  $E/N$  ist. Da  $\varepsilon_i$  die Dimension einer Energie hat, kann man auch ansetzen

$$\varepsilon_i = \frac{E}{N} \eta_i, \quad (\text{III.37})$$

worin die neuen Variablen  $\eta_i$  dimensionslos sind. Im Integral in (III.32) ist  $E$  so festgelegt, daß gilt

$$d\varepsilon_i = \frac{E}{N} d\eta_i. \quad (\text{III.38})$$

Andererseits weiß man durch Formel (III.26), daß gilt

$$\rho(\varepsilon) = AV\varepsilon^{1/2}, \quad (\text{III.39.a})$$

mit

$$A = \frac{\sqrt{2}}{2\pi^2\hbar^3} m^{3/2}. \quad (\text{III.39.b})$$

Man erhält schließlich

$$\begin{aligned} \Phi_N(E) &= \left(\frac{E}{N}\right)^N A^N V^N \left(\frac{E}{N}\right)^{N/2} \int_0^{+\infty} d\eta_1 d\eta_2 \dots d\eta_N [\eta_1 \eta_2 \dots \eta_N]^{1/2} \\ &\quad \cdot \theta\left[\frac{E}{N}(N - \eta_1 - \eta_2 - \dots - \eta_N)\right] \\ &= A^N V^N \left(\frac{E}{N}\right)^{3N/2} \int_0^{+\infty} \left[\prod_{i=1}^N d\eta_i \eta_i^{1/2}\right] \theta\left[N - \sum_{i=1}^N \eta_i\right] \end{aligned} \quad (\text{III.40})$$

denn

$$\theta(ax) = \theta(x) \quad \text{für} \quad a > 0. \quad (\text{III.41})$$

Das Integral muß numerisch berechnet werden. Es genügt also nach  $E$  abzuleiten, um die Zustandsdichte in der Form

$$\rho_N(E) = C(N) V^N \left(\frac{E}{N}\right)^{\frac{3N}{2}-1}, \quad (\text{III.42})$$

zu erhalten, worin  $C(N)$  ein numerischer Koeffizient ist, der von dem Volumen  $V$  und der Energie  $E$  unabhängig ist. Man kann diesen Koeffizienten berechnen (vgl. Ergänzung I.E), und man erhält

$$C(N) = \frac{3}{2} \left(\frac{mN}{2\pi\hbar^2}\right)^{3N/2} \frac{1}{\Gamma\left(\frac{3N}{2} + 1\right)} \simeq \frac{1}{2} \sqrt{\frac{3}{N\pi}} \left(\frac{me}{3\pi\hbar^2}\right)^{3N/2} \quad (\text{III.43})$$

Der wichtigste Aspekt der vorangehenden Überlegung ist *der Gebrauch der Dimensionsanalyse*: Man braucht nur zu verstehen, daß jedes  $\varepsilon_i$ , das die Dimension einer Energie hat, zu  $E$  proportional sein muß; man erhält dann sofort aus (III.32) und (III.39) die Eigenschaft  $\Phi_N(E)$ , indem man einfach die Differentiale und die  $\rho(\varepsilon)$  berechnet.

Die Formel (III.42) zeigt, daß für große  $N$   $\rho_N(E)$  eine *sehr schnell wachsende Funktion* von  $E$  ist. Es liegt dann nahe, davon den Logarithmus zu nehmen, der groß ist (linear in  $N$ ), aber langsam mit  $E$  variiert (wie  $\ln E$ ).

*Achtung*: die Formel (III.43) unterstellt, daß die  $N$  Teilchen unterscheidbar sind. Für identische und daher nicht unterscheidbare Teilchen läßt sich ein angenähertes Resultat erhalten, indem man  $C(N)$  durch  $N!$  dividiert (vgl. Fußnote 21).

### 3. Verallgemeinerung zu einem System von $N$ schwach gekoppelten Elementen

Die Überlegung des voranstehenden Abschnittes stützt sich prinzipiell auf die Unabhängigkeit der  $N$  Teilchen des Systems: Sie gestattet, die Ausgangsformel (III.32) zu schreiben, in der die Energie des Systems die Summe der einzelnen Energien der  $N$  Teilchen, und die Anzahl der Zustände des Systems das Produkt der Anzahl von Zuständen eines jeden Teilchens ist. Die Tatsache, daß es sich um freie Teilchen handelt, ist in Wirklichkeit sekundär, wenn man dabei die Größenordnungen beachtet. Dies werden wir in diesem Abschnitt zeigen.

Wir betrachten ein System aus  $N$  schwach gekoppelten Elementen. Dies bedeutet, daß der Hamiltonoperator  $H$  des Systems zum Beispiel von der Form

$$H = \sum_{i=1}^N H_i + \sum_{\substack{i,j=1 \\ i \neq j}}^N H_{ij}, \quad (\text{III.44})$$

ist, wobei in  $H_{ij}$  die Bestandteile ( $i$ ) und ( $j$ ) gleichzeitig auftreten, jedoch so, daß *die Wechselwirkungsterme  $H_{ij}$  klein sind* gegenüber den Einzeltermen  $H_i$ <sup>24</sup>. Man kann daher in erster Näherung die Wechselwirkungsterme vernachlässigen und erhält somit ein System von unabhängigen Elementen<sup>25</sup>.

<sup>24</sup>In zahlreichen Systemen sind die Wechselwirkungen nicht vernachlässigbar. Dennoch gelingt es manchmal sogar für solche Systeme, das Problem auf den hier betrachteten Fall zurückzuführen: Es gibt Methoden, die es gestatten, den Hamiltonoperator  $H$  durch einen approximativen Hamiltonoperator zu ersetzen, wobei diesmal die Einzeltermen groß gegenüber den Kopplungstermen zwischen den Bestandteilen sind; wir werden dazu Beispiele in den Ergänzungen III.E bis III.K betrachten.

<sup>25</sup>Man kann in gewissen Fällen die Approximation verbessern, indem man anschließend die Wechselwirkungsterme mit der Störungstheorie behandelt (siehe zum Beispiel die

Beginnen wir mit der Untersuchung eines dieser Elemente, das als isoliert angenommen und durch den Hamiltonoperator  $H$  beschrieben wird. Wenn die korrespondierenden Energieniveaus äquidistant (mit einem Zwischenabstand  $\delta$ ) und von gleichem Entartungsgrad wären, wie im Fall des eindimensionalen harmonischen Oszillator, so wäre die Anzahl  $\Phi(\varepsilon)$  von unterscheidbaren Zuständen, deren Energie kleiner als  $\varepsilon$  ist, von der Form

$$\Phi(\varepsilon) \propto \varepsilon/\delta, \quad (\text{III.45})$$

das heißt linear in  $\varepsilon$ . Im allgemeinen Fall sind die Eigenwerte von  $H$  nicht äquidistant, und in der Formel (III.45) tritt dann der mittlere Zwischenabstand  $\delta$  der Eigenwerte auf, die kleiner als  $\varepsilon$  sind; im übrigen muß dieser mittlere Zwischenabstand der Tatsache Rechnung tragen, daß für diese Niveaus auch der Entartungsgrad nicht der gleiche ist. Kurz gesagt,  $\delta$  hängt im allgemeinen von  $\varepsilon$  ab. Dennoch ist diese Abhängigkeit nicht so stark, daß sie die Abschätzung (III.45) vollständig ändern könnte. Im großen und ganzen hat man daher immer

$$\Phi(\varepsilon) \propto \varepsilon^\alpha, \quad (\text{III.46})$$

mit

$$\alpha \sim 1. \quad (\text{III.47})$$

Die Zustandsdichte  $\rho(\varepsilon)$  folgt durch Ableitung daraus :

$$\rho(\varepsilon) \propto \varepsilon^{\alpha-1}. \quad (\text{III.48})$$

Wir gehen nun zu einem System von  $N$  Elementen über. In dem Maße, wie ihre Wechselwirkungen vernachlässigbar sind, kann man auf die Formel (III.32) und den Ausdruck (III.36), der sich daraus für die Zustandsdichte  $\rho_N(E)$  des Systems ableitet, zurückgreifen. Wie wir am Ende des Abschnittes 2.b.β unterstrichen haben trägt jedes Differential  $d\varepsilon_i$  in (III.32) einen Faktor  $E/N$  bei und jedes  $\rho(\varepsilon_i)$  einen Faktor  $(E/N)^{\alpha-1}$  (nach (III.48)) bei. Man erhält somit

$$\Phi_N(E) \propto \left(\frac{E}{N}\right)^{N\alpha}. \quad (\text{III.49})$$

Die Energieabhängigkeit von  $\rho_N(E)$  ist daher derjenigen vergleichbar, die wir weiter oben für freie und unabhängige Teilchen gefunden haben. Diese (des öfteren angenäherte) Abhängigkeit der verschiedenen Elemente ist die wesentliche Eigenschaft, um zu einem Resultat dieses Typs zu gelangen; für den Fall, daß es sich nicht mehr um freie Teilchen handelt, wird einfach im Exponenten  $3N/2$  durch  $N\alpha$  ersetzt, wobei  $\alpha$  kaum von 1 abweicht.

## C. Beschreibung eines klassischen Systems

Im allgemeinen erlaubt die Quantenmechanik die Untersuchung der Eigenschaften eines beliebigen Systems, ob groß oder klein. Doch gibt es Fälle, in denen die klassische Newtonschen Mechanik gültig ist; unter gewissen Bedingungen beschreibt sie in guter Näherung das Verhalten von Systemen (siehe Abschnitt I.4 der Ergänzung I.A). Wir werden hier die Grundlagen der klassischen statistischen Mechanik aufzeigen<sup>26</sup>.

### 1. Klassischer mikroskopischer Zustand

In der klassischen Mechanik (siehe Anhang III) ist der Zustand eines punktförmigen Teilchens zu einem gegebenen Zeitpunkt durch seine Position  $\vec{r}$  und seine Geschwindigkeit  $\vec{v}$  gegeben, oder im Hamiltonschen Formalismus durch die Lage  $\vec{r}$  und den Impuls  $\vec{p}$ , den zur Lage konjugierte Impuls. Wenn man es mit einem System aus  $N$  Teilchen zu tun hat, so ist sein augenblicklicher klassischer Zustand durch die Angabe von  $6N$  Parametern definiert :

$$\{\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N; \vec{p}_1, \vec{p}_2, \dots, \vec{p}_N\}, \quad (\text{III.50})$$

$\vec{r}_j$  und  $\vec{p}_j$  sind die Position und der Impuls des mit  $j$  nummerierten Teilchens.

Dies läßt sich für ein beliebiges System verallgemeinern. Um zu einem bestimmten Zeitpunkt den Zustand des Systems zu kennen (das heißt die Position und die Geschwindigkeit eines jeden Punktes), benötigt man eine gewisse Anzahl  $n$  von Parametern  $q_i$  ( $i = 1, 2, \dots, n$ ), die *generalisierte Koordinaten* heißen, sowie (im Hamiltonschen Formalismus) die *n konjugierten Impulse*  $p_i$  der Parameter  $q_i$ . Der momentane Zustand des Systems ist dann durch die  $2n$  Größen

$$\{q_1, q_2, \dots, q_n; p_1, p_2, \dots, p_n\} \quad (\text{III.51})$$

charakterisiert. Die Bewegung des Systems wird durch die zeitliche Entwicklung seines Zustandes beschrieben : durch die  $2n$  Funktionen der Zeit<sup>27</sup>

$$\{q_i(t), p_i(t); i = 1, 2, \dots, n\}. \quad (\text{III.52})$$

Jeder *physikalischen Größe*, die man an einem System messen kann, ist eine Funktion von Zustandsvariablen  $q_i$  und  $p_i$  zugeordnet (diese Funktion

<sup>26</sup>Historisch gesehen ist die statistische Mechanik vor der Quantenmechanik entwickelt worden und basiert daher auf der klassischen Mechanik.

<sup>27</sup>Diese Funktionen sind natürlich nicht beliebig : Sie müssen die *Bewegungsgleichungen* erfüllen, Differentialgleichungen also, die das Fundamentalgesetz der Newtonschen Mechanik ausdrücken (Anhang III, §I.2 und §II.2).

kann unter anderem explizit von der Zeit abhängen), die den Wert dieser Größe angibt, wenn man den Zustand (III.51) kennt. Somit ist der Gesamtenergie des Systems die *Hamilton-Funktion*  $\mathcal{H}(q_i, p_i; t)$  zugeordnet. Beispielsweise schreibt sich für ein System von  $N$  punktförmigen Teilchen, in dem Kräfte von einer potentiellen Energie  $U(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n)$  ausgehen, die Hamilton-Funktion einfach zu

$$\mathcal{H}(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N; \vec{p}_1, \vec{p}_2, \dots, \vec{p}_N) = \sum_{j=1}^N \frac{\vec{p}_j^2}{2m_j} + U(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n) \quad (\text{III.53})$$

( $m_j$  ist die Masse des Teilchens ( $j$ )) : Sie ist die Summe der kinetischen Energie der  $N$  Teilchen und der gesamten potentiellen Energie.

Es ist praktisch, sich einen  $2n$ -dimensionalen Raum vorzustellen ( $6N$  für ein System von  $N$  punktförmigen Teilchen), genannt « *Phasenraum* », in dem jeder mögliche Zustand des Systems durch einen Punkt mit den Koordinaten  $q_1, q_2, \dots, q_n, p_1, p_2, \dots, p_n$  dargestellt wird. Die Bewegung des Systems wird durch eine Verschiebung des darstellenden Punktes<sup>28</sup> im Phasenraum beschrieben; er durchläuft eine Bewegungsbahn, die durch die vorgegebenen Funktionen (III.52) ausgedrückt wird.

In den Fällen, in denen die klassische Mechanik korrekt die Eigenschaften des untersuchten Systems wiedergibt, wird somit sein mikroskopischer Zustand durch die Vorgabe von  $2n$  Variablen charakterisiert, die man als die Koordinaten eines Punktes in dem dem betrachteten System zugeordneten Phasenraum betrachten kann.

## 2. Makroskopischer Zustand

Wenn die Anzahl von  $2n$  Parametern zur vollständigen Beschreibung des Systems groß ist (das heißt zum Beispiel, daß die Anzahl  $N$  der Teilchen, die das System bilden, groß ist), ist es unmöglich, seinen mikroskopischen Zustand exakt zu kennen. Zur Definition eines makroskopischen Zustandes, der groß genug ist, um die tatsächlich noch zugänglichen makroskopischen Eigenschaften des Systems zu berechnen, bedient man sich dann der Wahrscheinlichkeiten.

In der klassischen Mechanik bilden die möglichen mikroskopischen Zustände ein Kontinuum, praktisch dargestellt durch den Phasenraum (§1) : Der Index ( $\ell$ ), der die Quantenzustände kennzeichnet, wird hier durch die Gesamtheit der  $2n$  Variablen (III.51) ersetzt. Der makroskopische Zustand

<sup>28</sup> *Achtung* : Die  $p_i$  sind wie die  $q_i$  hier *Positions*-Koordinaten des darstellenden Punktes im Phasenraum; die Geschwindigkeit dieses Punktes auf seiner Trajektorie ist der  $2n$ -dimensionale Vektor mit Komponenten  $\{\dot{q}_i, \dot{p}_i; i = 1, 2, \dots, n\}$ .

ist daher durch eine *Wahrscheinlichkeitsdichte im Phasenraum* charakterisiert, das heißt durch eine Funktion der generalisierten Koordinaten  $q_i$  und ihrer konjugierten Impulse  $p_i$ . Diese Funktion hängt außerdem im allgemeinen Fall explizit von der Zeit ab :

$$w(q_1, q_2, \dots, q_n; p_1, p_2, \dots, p_n; t). \quad (\text{III.54})$$

Diese Funktion liefert die Wahrscheinlichkeit

$$dP(q_1, q_2, \dots, q_n; p_1, p_2, \dots, p_n; t) = w(q_1, q_2, \dots, q_n; p_1, p_2, \dots, p_n; t) dq_1 dq_2 \dots dq_n dp_1 dp_2 \dots dp_n, \quad (\text{III.55})$$

dafür, daß zum Zeitpunkt  $t$  der mikroskopische Zustand des Systems seinen repräsentativen Punkt in dem infinitesimalen Volumen

$$dq_1 dq_2 \dots dq_n dp_1 dp_2 \dots dp_n$$

hat, das den Punkt  $\{q_1, q_2, \dots, q_n; p_1, p_2, \dots, p_n\}$  des Phasenraumes enthält.

Diese Wahrscheinlichkeit ist normiert : Zu jedem Zeitpunkt  $t$  gilt<sup>29</sup>

$$\int w(q_i, p_i; t) \prod_{i=1}^n dq_i dp_i = 1, \quad (\text{III.56})$$

dabei wird über den Phasenraum integriert. Der Mittelwert einer beliebigen physikalischen Größe  $f(q_i, p_i; t)$  berechnet sich nach der üblichen Formel :

$$\bar{f} = \int w(q_i, p_i; t) f(q_i, p_i; t) \prod_{i=1}^n dq_i dp_i. \quad (\text{III.57})$$

Dieser Mittelwert hängt nur noch von der Zeit  $t$  ab.

Im Anhang III (§II.3.b) wird eine wichtige Eigenschaft der Wahrscheinlichkeitsdichte im Phasenraum dargelegt und bewiesen (Liouville-Gleichung).

## IV. Ergodenhypothese und statistisches Ensemble

Wir werden in diesem Abschnitt IV die Hypothese behandeln, auf der die wahrscheinlichkeitstheoretische Beschreibung von Systemen in der statistischen Mechanik beruht. Wir werden zunächst diese Hypothese in einem einfachen Beispiel darstellen, anschließend werden wir zeigen, wie sie sich verallgemeinern läßt und wie sie durch Einführung der Wahrscheinlichkeiten wiedergegeben werden kann. Wir enden mit einer kurzen Diskussion dieser fundamentalen Hypothese.

<sup>29</sup>Der Integrand der Formel (III.56) ist nichts anderes als der Ausdruck (III.55) in vereinfachter Schreibweise.

## A. Zeitliche Entwicklung eines makroskopischen Systems

### 1. Untersuchung eines einfachen Beispiels

Wir betrachten zunächst ein Gas, das aus einer großen feststehenden Anzahl  $N$  von identischen Molekülen besteht. Die Bewegung dieser Gasmoleküle auf mikroskopischem Niveau kann mit der klassischen Mechanik behandelt werden (siehe Ergänzung I.A, §III.2); das erlaubt uns intuitiv vorzugehen.

Wenn man die Entwicklung eines Gases im Verlauf der Zeit untersucht, so kann man zwei Arten von Bewegung unterscheiden.

**a. Entwicklung auf mikroskopischem Niveau : « thermische Bewegung » und Fluktuationen.** Nehmen wir in erster Linie an, daß sich das Gas aus makroskopischer Sicht im Gleichgewicht befindet. Dies wird der Fall sein, wenn sich die ihm auferlegten äußeren Bedingungen (Volumen des Behälters, Temperatur, ...) nicht ändern und wenn man genügend lange nach der Füllung des Behälters gewartet hat.

Wir wissen, daß dieses makroskopische Gleichgewicht tatsächlich eine ständige Bewegung auf mikroskopischem Niveau verbirgt : Die Moleküle sind mit ziemlich beträchtlichen Geschwindigkeiten ( $\simeq 500$  m/s) unaufhörlich in Bewegung; sie stoßen sich gegenseitig gegen die Wände des Behälters, so daß sich oft ihre Geschwindigkeit in Größe und Richtung ändert. Das scheinbare Gleichgewicht ist darauf zurückzuführen, daß die Moleküle extrem zahlreich sind, und daß sie zahlreiche Zusammenstöße erleiden : Es gibt somit in jedem Augenblick praktisch ebenso viele Moleküle, die sich in der rechten Hälfte wie in der linken Hälfte des Behälters befinden, und praktisch ebenso viele, die sich mit der Geschwindigkeit  $v$  bzw. mit der Geschwindigkeit  $-v$  bewegen. Diese Bewegung auf mikroskopischem Niveau wird häufig « *thermische Bewegung* » genannt, denn die Geschwindigkeiten sind umso größer je höher die Temperatur ist.

Die thermische Bewegung ist von *Fluktuationen* begleitet. Zum Beispiel ist die Anzahl der Moleküle, die sich in der rechten Hälfte des Behälters befinden, nicht exakt und konstant gleich  $N/2$  : Sie fluktuiert im Lauf der Zeit um diesen Mittelwert, die Amplitude dieser Fluktuation ist dabei klein im Relativwert ( $\sim \sqrt{N}/N$ ), aber nicht Null. Ebenso ist die Anzahl der Moleküle mit der Geschwindigkeit  $v$  einmal größer, ein anderes mal kleiner als die von Molekülen, die in diesem Augenblick die Geschwindigkeit  $-v$  besitzen, denn diese Zahlen fluktuieren ebenfalls im Lauf der Zeit. Wir erinnern daran, daß diese durch Bewegungen auf mikroskopischem Niveau verursachten Fluktuationen experimentell gezeigt werden können, was die Realität der Moleküle und ihrer Bewegung beweist.

Konkret stellen wir uns vor, wir könnten zu jedem Zeitpunkt  $t$  die Moleküle zählen, die sich in einem bestimmten und festen Gebiet des Behälters, der das Gas enthält, befinden (Abb. 2.a)<sup>30</sup>. Die erhaltenen Zahlen  $n(t)$  als Funktion der Zeit  $t$  sind in der Abbildung 2.b dargestellt : Fluktuationen von kleiner Amplitude relativ zum Zentralwert  $n_0$ .

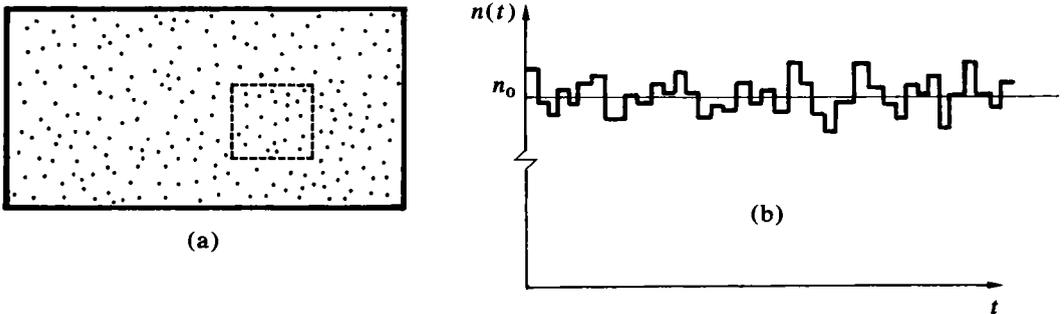


ABBILDUNG 2

In einem Gas, das sich im makroskopischen Gleichgewicht befindet, mißt man die Anzahl von Molekülen  $n(t)$ , die sich zu einem bestimmten Zeitpunkt  $t$  in einem festen Gebiet des Behälters befinden, begrenzt durch eine gepunktete Linie in der Abbildung a. Der Verlauf der Änderungen von  $n(t)$  ist in der Abbildung b dargestellt.

**b. Entwicklung auf makroskopischem Niveau.** Das im Abschnitt 1 betrachtete Gas befand sich aus makroskopischer Sicht im Gleichgewicht, aber das ist nicht immer der Fall. Stellen wir uns zum Beispiel vor, daß der Behälter, der das Gas enthält, durch eine bewegliche Zwischenwand in zwei Kammern aufgeteilt wird (Abb. 3.a). Man präpariert das System derart, daß eine der Kammern leer ist und sich das Gas in der anderen befindet. Wenn man zu einem bestimmten Zeitpunkt die Trennwand zurückzieht, breitet sich das Gas auch in die ursprünglich leere Kammer aus und erfüllt schließlich den gesamten zur Verfügung stehenden Raum. Die Änderung der Anzahl  $n(t)$

<sup>30</sup>Um die Moleküle zu « zählen », müßte man die *Dichte* des Gases in dem betrachteten Gebiet mit einer genügend großen Genauigkeit messen, um Fluktuationen wahrnehmen zu können. Die Dichte ist hier aus praktischen Gründen der Darstellung ausgewählt worden; jede andere makroskopische Größe zeigte im wesentlichen das gleiche Verhalten.

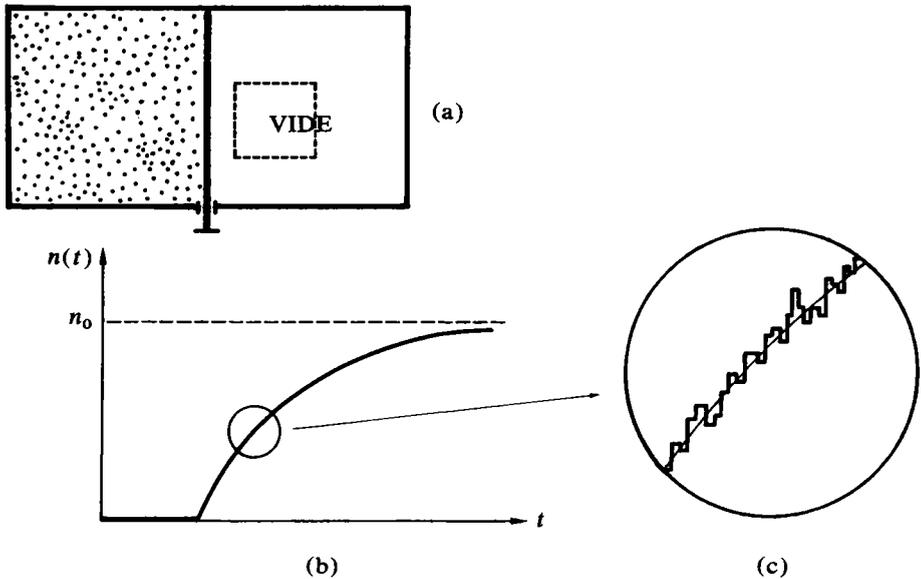


ABBILDUNG 3

Man mißt die Anzahl  $n(t)$  der Moleküle, die sich in dem Gebiet befinden, das durch die gestrichelte Linie der Abbildung a gekennzeichnet ist; diesmal entwickelt sich das Gas auf makroskopischem Niveau: Zum Zeitpunkt  $t = 0$  zieht man die in Abbildung a dargestellte bewegliche Zwischenwand weg. Die Abbildung b zeigt den Kurvenverlauf der Änderung von  $n(t)$ , von dem ein Ausschnitt in Abbildung c vergrößert dargestellt ist.

von Molekülen, die sich im gleichen Gebiet des Behälters aufhalten, wird offensichtlich im Laufe der Zeit qualitativ verschieden sein von derjenigen, die im makroskopischen Gleichgewicht beobachtet wird: Die Ausdehnung des Gases drückt sich in einer makroskopischen Änderung von  $n$  aus, die dabei von 0 bis  $n_0$  geht (Abb. 3.b). Betrachtet man diese Änderung in einem vergrößerten Maßstab (Abb. 3.c), so stellt man fest, daß die makroskopische Änderung ebenfalls unregelmäßig erfolgt, denn die Bewegung auf mikroskopischem Niveau ist natürlich noch immer vorhanden.

## 2. Verallgemeinerung für beliebiges makroskopisches System

In dem soeben untersuchten Beispiel kann die mikroskopische Bewegung des Gases durch die klassische Mechanik beschrieben werden. Aber im allgemeinen erfordert die Beschreibung eines Systems auf mikroskopischem Niveau

den Einsatz der Quantenmechanik. Wir betrachten in diesem Fall (siehe Abschnitt III.A.2.b) den Eigenzustand ( $\ell$ ) des Hamiltonoperators  $H$  eines Systems mit den Energien  $E_\ell$ . Wenn der Hamiltonoperator eines Systems zeitunabhängig ist, dann sind seine Eigenzustände stationär : Befindet sich das System zu einem bestimmten Zeitpunkt in dem Eigenzustand des Hamiltonoperators, so bleibt es darin unendlich lange, und seine Eigenschaften ändern sich nicht im Laufe der Zeit. Man muß dabei die grundlegende Tatsache berücksichtigen, daß *für ein System von makroskopischer Größe der Hamiltonoperator nur unvollständig definiert ist*. Es gibt dafür sowohl theoretische als auch praktische Gründe. Der Hauptgrund ist, daß das zu untersuchende System nicht rigoros vom Rest des Universums isoliert werden kann : Äußere Einflüsse, die auf makroskopischem Niveau vernachlässigbar sind, müßten im mikroskopischen Hamiltonoperator berücksichtigt werden, und dies ist natürlich nicht möglich. Zum Beispiel diskutieren wir im Anhang II (§II.4.c.α) die Randbedingungen dafür, daß das System in einem Behälter von makroskopischem Volumen eingeschlossen ist : Wenn die Ausmaße des Behälters im makroskopischem Maßstab wohl definiert sind, sind sie noch weit davon entfernt, mit ausreichender Genauigkeit als genau bestimmt in einer mikroskopischen Berechnung betrachtet zu werden. Ebenso wird die mikroskopische Bewegung der Bestandteile des Systems durch Kräfte beeinflusst, die auf sie von Körpern außerhalb des Systems ausgeübt werden. Diese Kräfte können unmöglich im Hamiltonoperator erfaßt werden und sind auf makroskopischem Niveau vernachlässigbar. Wir kommen auf diese Effekte im Teil IV zurück und bestimmen ihre Größenordnung in der Ergänzung IV.B.

Die Folge davon ist, daß der Hamiltonoperator  $H$  notwendigerweise angenähert ist, und seine Eigenzustände ( $\ell$ ) in Wirklichkeit nicht stationär sind : *Das System geht ständig im Laufe der Zeit von einem Zustand ( $\ell$ ) in einen anderen über*<sup>31</sup>, und zwar auf unkontrollierbare Art und Weise. Daraus resultieren, wie im Beispiel des Gases (§1), zeitliche Fluktuationen der verschiedenen physikalischen Größen. Wie im Abschnitt 1 sind diese Fluktuationen sowohl vorhanden, wenn das System sich makroskopisch ändert, als auch wenn es makroskopisch im Gleichgewicht ist.

---

<sup>31</sup>In der Quantenmechanik ist der (reine) Zustand eines Systems im allgemeinen eine Linearkombination von Basiszuständen (hier den Zuständen ( $\ell$ )). Einfach aus praktischen Gründen tun wir so, als ob die Zustände ( $\ell$ ) die einzigen möglichen mikroskopischen Zustände wären (vgl. §III.A.2).

## B. Mittelwert über die mikroskopischen Fluktuationen

Die statistische Mechanik setzt sich zum Ziel, so genau wie möglich die makroskopische Entwicklung von solchen Systemen zu beschreiben, wie wir sie im Abschnitt A betrachtet haben. Aber wir wissen, daß das unmöglich ist, und überdies unnötig ihre mikroskopische Bewegung in alle Einzelheiten zu « verfolgen ». Letztere wird daher mit wahrscheinlichkeitstheoretischen Techniken behandelt : Man bildet Mittelwerte über die durch die mikroskopische Bewegung erzeugten Fluktuationen.

### 1. Zeitlicher Mittelwert

**a. Definition.** Untersuchen wir zuerst den Fall, in dem sich das betrachtete System (zum Beispiel das im Abschnitt A.1.a behandelte Gas) aus makroskopischer Sicht im Gleichgewicht befindet. Die Änderung irgendeiner Größe  $f(t)$  mit der Zeit zeigt auch den Kurvenverlauf, wie er für  $n(t)$  in der Abbildung 2.b dargestellt ist.

Auf makroskopischem Niveau interessieren wir uns hauptsächlich für den Zentralwert  $f_0$ , um den  $f(t)$  schwankt, denn die Amplitude der Fluktuationen besitzt einen relativ kleinen Wert. Um diesen Zentralwert zu bestimmen, genügt es, das System während eines ausreichend langen Zeitintervalles  $\tau$ , damit sich eine große Anzahl von Fluktuationen ereignen, zu untersuchen. Man berechnet dann auf gewöhnliche Art und Weise den Mittelwert von der im Verlauf des Zeitintervalls  $\tau$  gemessenen Größe : Mißt man zwischen  $t$  und  $t + dt$  den Wert  $f(t)$ , so muß dieser noch mit dem « Gewicht »  $dt/\tau$  versehen werden, und man muß anschließend über alle Werte von  $t$  zwischen dem Anfang  $t_0$  und dem Ende  $t_0 + \tau$  der Beobachtungszeit summieren. Man gelangt somit zur Definition des *zeitlichen Mittelwertes*  $\hat{f}$  der Größe  $f$  :<sup>32</sup>

$$\hat{f} := \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_{t_0}^{t_0 + \tau} f(t) dt. \quad (\text{IV.1})$$

Befindet sich das System auf makroskopischem Niveau im Gleichgewicht, so ist der gefundene Wert  $\hat{f}$  von dem gewählten Anfangszustand  $t_0$  unabhängig.

---

<sup>32</sup>Der zeitliche Mittelwert  $\hat{f}$  ist *a priori* verschieden von dem probalistischen Mittelwert, der im Anhang I (§I.4.b) definiert wird. Tatsächlich wiederholt man hier die Messung von  $f$  an dem gleichen System, statt jedesmal ein neues System zu nehmen (immer auf gleiche Art präpariert), damit jedes Experiment unabhängig von den vorhergehenden ist. Man kann jedoch die Definition (IV.1) benutzen, um durch eine Formel, analog zur Definition (29) des Anhanges I, den quadratischen Mittelwert einer Größe, die die Amplitude der zeitlichen Fluktuationen charakterisiert, ebenso zu berechnen.

**b. Beschreibung der makroskopischen Entwicklung.** Läßt sich die voranstehende Definition so verallgemeinern, daß das betrachtete System sich wie in dem Beispiel des Abschnittes A.1.b makroskopisch entwickelt?

Diese Verallgemeinerung stößt sofort auf eine grundlegende Schwierigkeit, die sich leicht an Hand der Abbildung 3.b verstehen läßt. Die makroskopische Entwicklung dauert eine endliche Zeit, an deren Ende sich das System in einem neuen Gleichgewichtszustand stabilisiert (in der Abbildung 3.b erreicht  $n(t)$  den Wert  $n_0$ , um den es anschließend schwankt wie in Abbildung 2.b). Der Anteil dieser Übergangsperiode an der Summe (IV.1) sinkt, wenn  $\tau$  wächst und wird schließlich vernachlässigbar. Der mit (IV.1) gefundene Mittelwert entspricht dem neuen Gleichgewichtszustand; er sagt nichts über die makroskopische Entwicklung des Systems aus, die wir beschreiben wollten.

Wenn die mikroskopischen Fluktuationen, wie die Abbildung 3.c suggeriert, jedoch sehr viel schneller sind als die makroskopische Entwicklung, so kann man die Formel (IV.1) derart modifizieren, daß sie brauchbar wird. Um es genauer zu beschreiben, sei  $\tau_m$  eine charakteristische Zeit der mikroskopischen Bewegung:  $\tau_m$  ist ein Zeitintervall definierter Größenordnung, in dem das System geringfügig fluktuiert. Ebenso sei  $\tau_M$  die charakteristische Zeit der makroskopischen Entwicklung. Wir nehmen an, daß es möglich ist, Zeitintervalle  $\tau$  so zu wählen, daß die zwei Bedingungen

$$\tau_m \ll \tau \ll \tau_M \quad (\text{IV.2})$$

erfüllt sind. Dann hängen die über einem solchen Intervall  $\tau$  gemäß (IV.1) berechneten Mittelwerte (ohne den Grenzwert  $\tau \rightarrow \infty$ ) von der Zeit  $t_0$  ab, was der makroskopischen Entwicklung des Systems entspricht. Damit die Bedingungen (IV.2) erfüllt werden können, muß  $\tau_M$  einige Größenordnungen größer als  $\tau_m$  sein.

## 2. Mittelwert über ein Ensemble

Seit Gibbs (1839 - 1903) ersetzt man in der statistischen Mechanik die zeitlichen Mittelwerte über die Fluktuationen eines bestimmten Systems durch Mittelwerte von Ensembles, die eine große Anzahl von identischen Systemen umfassen.

**a. Statistisches Ensemble.** Anstatt ein einziges makroskopisches System zu untersuchen, konstruieren wir davon Duplikate: Wir nehmen zum Beispiel mehrere identische Behälter, füllen sie mit der gleichen Anzahl von Molekülen des gleichen Gases und unterwerfen sie den gleichen Bedingun-

gen. Wird die Anzahl dieser Behälter sehr groß, stellen sie ein *statistisches Ensemble* dar.

Betrachten wir nun dieses Ensemble von Systemen zu *einem bestimmten Zeitpunkt*. Findet in ihnen eine makroskopische Entwicklung statt (wie im Abschnitt A.1.b), so erwartet man, daß sich zu einem bestimmten Zeitpunkt alle Systeme *im gleichen Stadium dieser makroskopischen Entwicklung* befinden : Wenn man in einer großen Anzahl von Fällen, die mit denen in Abbildung 3.a identisch sind, im selben Augenblick und auf die gleiche Art die Zwischentrennwand entfernt, so äußert sich die Expansion des Gases in jedem Behälter in der gleichen makroskopischen Änderung von  $n(t)$  gemäß Abbildung 3.b. Im Gegensatz dazu erwartet man, daß sich die verschiedenen Systeme des Ensembles zum gleichen Zeitpunkt in *verschiedenen Stadien der mikroskopisch fluktuierenden Entwicklung* befinden. Tatsächlich sind die Systeme des statistischen Ensembles im makroskopischen Maßstab nicht identisch, denn in diesem Maßstab ist ihre Präparation kontrollierbar : Es ist zum Beispiel unmöglich zu gewährleisten, daß die verschiedenen Behälter an der gleichen Stelle genau die gleiche Anzahl von Molekülen enthalten; ebenso ist ihr Volumen nur mit der Genauigkeit makroskopischer Messungen bestimmbar. Konsequenterweise befinden sich die Systeme schon bei ihrer Präparation in verschiedenen mikroskopischen Zuständen, und ihre fluktuierende Entwicklung auf mikroskopischem Niveau folgt verschiedenen, jedoch analogen Kurven.

**b. Charakterisierung eines statistischen Ensembles.** Bedienen wir uns wieder der Quantenbeschreibung der mikroskopischen Zustände und nehmen an, daß jedes der identischen Systeme (auf makroskopischem Niveau) des statistischen Ensembles sich in einem Eigenzustand ( $\ell$ ) des Hamiltonoperators  $H$  befindet. Nach dem bereits gesagtem befinden sich die verschiedenen Systeme in unterschiedlichen mikroskopischen Zuständen ( $\ell$ )<sup>33</sup>. Es sei  $\mathcal{N}$  die Gesamtzahl der Systeme des Ensembles und  $\mathcal{N}_\ell$  die Anzahl derjenigen, die sich zu einem bestimmten Zeitpunkt im mikroskopischen Zustand ( $\ell$ ) befinden.

Nehmen wir an, daß man eine bestimmte Größe  $f$  auswählt und ihren Wert in jedem der Systeme des Ensembles mißt und anschließend daraus den Mittelwert bildet (Summe der  $\mathcal{N}$  erhaltenen Werte durch  $\mathcal{N}$ ). Man erhält so den Ensemble-Mittelwert der Größe  $f$  zu einem betrachteten Zeitpunkt. Es ist klar, daß dieser Ensemble-Mittelwert, falls  $f$  den Wert  $f_\ell$  im Zustand

<sup>33</sup>Die äußeren Bedingungen, die dem System (im makroskopischen Maßstab) auferlegt sind, schränken die möglichen mikroskopischen Zustände ein. Aber die Zahl der Zustände ( $\ell$ ), die mit diesen Bedingungen kompatibel sind, ist immer noch enorm groß.

( $\ell$ ) annimmt, zu

$$\bar{f}_N = \frac{1}{N} \sum_{(\ell)} \mathcal{N}_\ell f_\ell, \quad (\text{IV.3})$$

wird, worin die Summe sich über die verschiedenen in dem Ensemble vorhandenen Zustände ( $\ell$ ) erstreckt. Im Grenzfall, wo  $N$  gegen Unendlich geht, wird das Verhältnis  $\mathcal{N}_\ell/N$  eine Wahrscheinlichkeit :

$$P_\ell := \lim_{N \rightarrow \infty} \frac{\mathcal{N}_\ell}{N}, \quad (\text{IV.4})$$

und der Ensemble-Mittelwert  $\bar{f}_N$  strebt gegen den Mittelwert, der mit diesen Wahrscheinlichkeiten berechnet wurde :

$$\lim_{N \rightarrow \infty} \bar{f}_N = \bar{f} = \sum_{(\ell)} P_\ell f_\ell. \quad (\text{IV.5})$$

Genau genommen ist  $P_\ell$  die Wahrscheinlichkeit dafür, daß sich ein System bei zufälliger Auswahl aus dem Ensemble im mikroskopischen Zustand ( $\ell$ ) befindet. In der Praxis verfügt man jedoch niemals über eine unendliche Anzahl (noch über eine große Anzahl) von identischen Systemen, sondern nur über ein System. Man interpretiert daher die Wahrscheinlichkeit  $P_\ell$ , definiert durch (IV.4), als die *Wahrscheinlichkeit dafür, daß sich das untersuchte System im mikroskopischen Zustand ( $\ell$ ) zu einem bestimmten Zeitpunkt befindet*. Somit erlaubt der Begriff des statistischen Ensembles auf präzise Art die Wahrscheinlichkeiten  $P_\ell$ , die wir in den vorhergehenden Abschnitten eingeführt haben, zu definieren. Umgekehrt ist das statistische Ensemble eines determinierten Systems für einen bestimmten Zeitpunkt vollständig charakterisiert durch die statistische Verteilung der mikroskopischen Zustände, das heißt durch das Ensemble der Wahrscheinlichkeiten  $P_\ell$  für alle möglichen Zustände ( $\ell$ )<sup>34</sup>.  $P_\ell$  wird oft « *Besetzungswahrscheinlichkeit* » des Zustandes ( $\ell$ ) genannt<sup>35</sup>.

**c. Entwicklung des Ensembles mit der Zeit.** In der Definition (IV.4) ist  $\mathcal{N}_\ell$  die Anzahl der Systeme des Ensembles, die sich zu *einem gegebenen*

<sup>34</sup>Die Zustände ( $\ell$ ), die mit den äußeren Bedingungen inkompatibel sind, haben eine Wahrscheinlichkeit Null. Gemäß praktischer Erwägungen kann man diese Zustände in die « möglichen Zustände » ein- oder auch nicht einbeziehen.

<sup>35</sup>Wenn die Bewegung des betrachteten mikroskopischen Systems durch die klassische Mechanik beschrieben werden kann, wird das zugeordnete statistische Ensemble durch eine Wahrscheinlichkeitsdichte  $w(q_1, q_2, \dots, q_n; p_1, p_2, \dots, p_n; t)$  in dem Phasenraum charakterisiert (siehe Abschnitt III.C).

Zeitpunkt im Zustand  $(\ell)$  befinden; die Wahrscheinlichkeiten  $P_\ell$  hängen daher *a priori* von der Zeit ab :  $P_\ell(t)$ . Dennoch ist anzumerken, daß sie nach (IV.4) stets die folgende Normierungsbedingung erfüllen :

$$\sum_{(\ell)} P_\ell(t) = 1. \quad (\text{IV.6})$$

Wir werden im Teil IV die Entwicklung der Besetzungswahrscheinlichkeiten  $P_\ell(t)$  mit der Zeit untersuchen; wir begnügen uns hier mit allgemeinen Betrachtungen.

Nehmen wir zunächst an, daß die Wahrscheinlichkeiten  $P_\ell$  zeitlich konstant sind :

$$\frac{dP_\ell}{dt} = 0 \quad \forall(\ell). \quad (\text{IV.7})$$

Der Ensemble-Mittelwert  $\bar{f}$  irgendeiner Größe  $f$  (selbst unabhängig von  $t$ ) ist dann von der Zeit unabhängig. Wenn die Ensemble-Mittelwerte korrekt die makroskopischen Eigenschaften des Systems beschreiben (siehe Abschnitt 3 unten), so werden diese Eigenschaften im gegenwärtigen Fall von der Zeit unabhängig, das bedeutet, daß das System sich *im makroskopischen Gleichgewicht* befindet. Genauer gesagt, ist die Bedingung (IV.7) jetzt *Definition* des makroskopischen Gleichgewichts, das wir bis jetzt im wesentlichen anschaulich behandelt haben<sup>36</sup>.

Beachten wir, daß jedes der Systeme des Ensembles sich mit der Zeit entwickelt und dabei unaufhörlich die mikroskopischen Zustände  $(\ell)$  verändert. Die Bedingung (IV.7) impliziert nur, daß sich das Verhältnis  $N_\ell/\mathcal{N}$  von Systemen des Ensembles, die sich im Zustand  $(\ell)$  befinden, nicht im Laufe der Zeit ändert (im Grenzfall  $\mathcal{N} \rightarrow \infty$ ); das heißt, daß es im Mittel ebenso viele Systeme des Ensembles gibt, die den Zustand  $(\ell)$  während eines Zeitintervalls  $\delta t$  verlassen, wie solche Systeme, die im Verlaufe des gleichen Zeitintervalls in dem gleichen Zustand  $(\ell)$  ankommen. Es handelt sich daher um ein *statistisches Gleichgewicht* : Es sind die Eigenschaften des Ensembles, nicht diejenigen der einzelnen Systeme (auf mikroskopischem Niveau), die von der Zeit unabhängig sind.

Wenn sich im Gegensatz dazu die Wahrscheinlichkeiten  $P_\ell(t)$  mit der Zeit ändern, so wird das gleiche mit den Ensemble-Mittelwerten erfolgen :

$$\bar{f}(t) = \sum_{(\ell)} f_\ell P_\ell(t). \quad (\text{IV.8})$$

<sup>36</sup>Die Eigenschaften eines Systems können unabhängig von der Zeit sein, ohne daß es im Gleichgewicht ist : Dies ist zum Beispiel der Fall für einen Strom durchflossenen Leiter, wenn der Strom stationär ist. Der Unterschied zwischen einem Gleichgewichtszustand und einem stationären Zustand außerhalb des Gleichgewichts besteht in dem Fließen (von Ladung, Energie oder von Teilchen ... ); diese Größen sind im zweiten Fall nicht alle Null.

Nach den obigen Ausführungen im Abschnitt a erwartet man, daß sich dies in einer *makroskopischen Entwicklung* ausdrückt.

### 3. Diskussion : Die Ergodenhypothese

Die Ensemble-Mittelwerte vom Abschnitt 2 stellen einen beträchtlichen Vorteil gegenüber den zeitlichen Mittelwerten des Abschnittes 1 dar : Nimmt man *alle Systeme des Ensembles zum gleichen Zeitpunkt  $t$* , so eliminiert man die grundsätzlichen Schwierigkeiten der zeitlichen Mittelwertbildung (§1.b); die makroskopische Entwicklung als zeitliche Änderung der Mittelwerte  $\bar{f}(t)$  verhindert nicht mehr die Berechnung der Mittelwerte von mikroskopischen Fluktuationen.

Doch ist der Begriff des Ensemble-Mittelwertes in anderer Hinsicht problematisch. Tatsächlich untersucht man ja wegen der *zeitlichen Fluktuationen* immer *ein einziges* System, so daß die mikroskopische Bewegung durch Mittelwerte ausgedrückt werden kann. Es stellt sich daher die Frage nach der *Äquivalenz von Ensemble-Mittelwerten und zeitlichen Mittelwerten*, wobei man nur letztere messen kann.

Im Abschnitt 2.a haben wir plausible Argumente für eine solche Äquivalenz geliefert. Tatsächlich « erwartet man », daß sich die verschiedenen Systeme eines statistischen Ensembles auf makroskopischem Niveau identisch entwickeln, wobei der Ensemble-Mittelwert diese makroskopische Entwicklung gewissermaßen erhält. Aber es wird nicht deutlich, daß sich die verschiedenen mikroskopischen Zustände ( $\ell$ ), die zu einem gegebenen Zeitpunkt in dem statistischen Ensemble vorhanden sind, ebenso mit den Häufigkeiten  $\mathcal{N}_\ell/N$  entwickeln, die genau denen entsprechen, die während der zeitlichen Fluktuationen eines einzigen Systems auftreten. Wir werden diese Frage in Einzelheiten in der Ergänzung IV.A diskutieren und sehen, daß sie gegenwärtig noch nicht vollständig gelöst ist.

Die statistische Mechanik basiert auf der *Ergodenhypothese*, die die Äquivalenz zwischen Ensemble-Mittelwert und zeitlichem Mittelwert postuliert. Wir akzeptieren diese Hypothese unter der Annahme, daß ihre Gültigkeit auf der experimentellen Verifizierung beruht : Künftig werden wir die im Abschnitt 2 eingeführten Begriffe als Ausgangspunkt verwenden.

## ERGÄNZUNGEN ZU TEIL I

---

<b>I.A. Größenordnungen</b>	Sehr physikalisch. Leicht lesbar, schwer verständlich. Könnte nach dem Studium von II und III nochmals aufgegriffen werden.
<b>I.B. Statistische Verteilung einer Funktion von Zufallsvariablen</b>	Ergänzungen über die Wahrscheinlichkeiten. Erfordert die Kenntnis der Begriffe, die im Anhang I wiederholt werden.
<b>I.C. Die Binomialverteilung im Grenzfall großer Zahlen</b>	I.B und I.C : ein wenig technisch; wird zum Studium empfohlen, wenn sich das Problem stellt; der in I.C entwickelte Gedankengang dient als Grundlage für viele andere Situationen.
<b>I.D. Zufallsbewegung und Brownsche Bewegung</b>	I.D : leicht, interessant wegen der zahlreichen Anwendungsmöglichkeiten der Zufallsbewegung.
<b>I.E. Zustandsdichte eines idealen Gases</b>	Leicht. Beweis eines Ergebnisses aus dem Hauptteil.
<b>I.F. Elektronenzustände in Festkörpern</b>	Behandlung eines wichtigen physikalischen Problems. Einige technische Schwierigkeiten, aber nicht unüberwindbar.
<b>I.G. Information und Entropie</b>	Entwickelt den wichtigen Begriff der statistischen Entropie, der im Hauptteil eingeführt wurde, und diskutiert im besonderen seine Anwendung auf kontinuierliche Wahrscheinlichkeitsverteilungen. Ziemlich leicht.
<b>I.H. Dichteoperator</b>	Führt das zur Formulierung der statistischen Quantenmechanik geeignete Werkzeug ein, viel allgemeiner und aussagekräftiger, als das im Hauptteil behandelte. Erfordert gute Kenntnis in den Grundtechniken der Quantenmechanik.

---

---

## ERGÄNZUNG I.A

# Größenordnungen

---

### I. Mikroskopische Größenordnungen

1. Die physikalischen Fundamentalkonstanten
2. Das isolierte Atom
  - a. Natürliche Längen- und Energieeinheiten
  - b. Ionisationsenergie und Atomradius
3. Moleküle und Festkörper
  - a. Bindungsenergien und interatomare Abstände
  - b. Schwingungs- und Rotationsenergien
  - c. Elektrische und magnetische Momente
4. Klassische Mechanik oder Quantenmechanik ?

### II. Der Übergang zum mikroskopischen Bereich

1. Makroskopische Größenordnungen
2. Temperaturskala
3. Elektrische und magnetische Felder

### III. Beispiel : Klassisches ideales Gas

1. Definitionen
2. Gültigkeit der klassischen Beschreibung
  - a. Gültigkeitsbedingung
  - b. Molekülgas
  - c. Elektronengas in einem Metall
3. Vergleich zwischen kinetischer und potentieller Energie

---

Diese Ergänzung ist dazu bestimmt, den Leser mit der Größenordnung von physikalischen Größen, die in diesem Buch behandelt werden, vertraut zu machen. Statt einen Katalog anzugeben, zeigen wir vielmehr wie man schnell die meisten Werte durch eine Anwendung der Dimensionsanalyse findet. Wir versuchen somit, das Interesse an qualitativen Methoden zur Abschätzung der Größenordnungen des bei der Lösung eines Problems zu erwartenden Resultates zu wecken. Diese Methoden erlauben nicht nur, die notwendigen Approximationen zu rechtfertigen, sondern sie liefern ein noch besseres Verständnis des physikalischen Inhaltes; sie helfen bei der quantitativen Lösung und Analyse des erhaltenen Ergebnisses<sup>1</sup>.

---

<sup>1</sup>Im Verlauf der Diskussion kommt es vor, daß wir auf späteren Stoff vorgreifen. Wir bitten den Leser später begründete Ergebnisse zu akzeptieren und raten ihm, nach Lektüre der entsprechenden Kapitel auf diese Ergänzung zurückzukommen.

# I. Mikroskopische Größenordnungen

## 1. Die physikalischen Fundamentalkonstanten

Auf den Vorsatzblättern dieses Buches befinden sich die physikalischen Fundamentalkonstanten. Vernachlässigt man die relativistischen Korrekturen, so lassen sich die Eigenschaften der Materie im atomarem Bereich (den wir als *mikroskopisch* bezeichnen), durch nur vier Konstanten beschreiben<sup>2</sup> :

- der Absolutbetrag der Elektronladung (Elementarladung)

$$e = 1,602 \cdot 10^{-19} \text{ C} \simeq 1,6 \cdot 10^{-19} \text{ C}, \quad (\text{A.1})$$

- die Masse des Elektrons

$$m_e = 9,109 \cdot 10^{-31} \text{ kg} \simeq 10^{-30} \text{ kg}, \quad (\text{A.2})$$

- die Masse des Nukleons (Proton oder Neutron)

$$M_N = 1,67 \cdot 10^{-27} \text{ kg} \simeq 1,7 \cdot 10^{-27} \text{ kg}, \quad (\text{A.3})$$

- die Plancksche Konstante

$$\hbar = \frac{h}{2\pi} = 1,055 \cdot 10^{-34} \text{ J} \cdot \text{s} \simeq 10^{-34} \text{ J} \cdot \text{s}. \quad (\text{A.4})$$

Zu diesen Konstanten muß im MKSA-System\* die Dielektrizitätskonstante des Vakuums (elektrische Feldkonstante im Vakuum)

$$\epsilon_0 = \frac{10^{-19}}{36\pi} \quad (\text{A.5})$$

hinzugefügt werden. Tatsächlich können die elektrischen Ladungen nur paarweise *über* die Coulombsche Wechselwirkung wirksam werden; eher als die Ladung  $e$  ist daher oft

$$q^2 := \frac{e^2}{4\pi\epsilon_0} = 2,310 \cdot 10^{-28} \text{ J} \cdot \text{m} \simeq 2,3 \cdot 10^{-28} \text{ J} \cdot \text{m} \quad (\text{A.6})$$

als physikalische Größe zu betrachten.

---

<sup>2</sup>Da es hier nur um Größenordnungen geht, vernachlässigen wir die Kernbindungsenergie : Die Masse eines Kernes mit  $A$  Nukleonen ist  $A \cdot M_N$ .

\*Anmerkung des Übersetzers : Heute ist das SI-System verbindlich :  $\epsilon_0 = 8,854 \cdot 10^{-12} \text{ Fm}^{-1}$ .

## 2. Das isolierte Atom

**a. Natürliche Längen- und Energieeinheiten.** Da die Masse des Kerns beträchtlich größer als die des Elektrons ist ( $M_N/m_e = 1836$ ), kann man voraussetzen, daß die Bewegung des Kernes keine bedeutende Rolle bei den atomaren Eigenschaften spielt : Man kann den Kern als unendlich schwer und als fest im Raum fixiert annehmen. Der « Radius » und die Energie des Atoms dürfen daher nur die drei Konstanten<sup>3</sup>  $m_e$ ,  $\hbar$ , und  $q^2$  enthalten. Man kann sich leicht davon überzeugen<sup>4</sup>, daß die Verknüpfungen dieser Konstanten die Dimensionen einer Länge bzw. einer Energie besitzen :

$$a_0 = \frac{\hbar^2}{m_e q^2} = 0,53 \cdot 10^{-10} \text{ m} = 0,53 \text{ \AA} \quad (\text{A.7})$$

bzw.

$$E_0 = \frac{m_e q^4}{\hbar^2} = 4.37 \cdot 10^{-18} \text{ J} = 27,3 \text{ eV}; \quad (\text{A.8})$$

$a_0$  und  $E_0$  erscheinen hierbei wie *natürliche Einheiten*<sup>5</sup> der Atomphysik von Länge und Energie : Man erwartet, daß jede Länge und jede Energie im atomaren Bereich proportional zu  $a_0$  bzw.  $E_0$  ist, wobei sich die Proportionalitätskonstante in vernünftigen Grenzen um 1 (sagen wir zwischen  $10^{-3}$  und  $10^3$ ) hält.

In der Praxis bevorzugt man gegenüber  $E_0$  die Einheit

$$R_y := \frac{E_0}{2} = \frac{m_e q^4}{2\hbar^2} \simeq 13,6 \text{ eV}, \quad (\text{A.9})$$

die als Rydberg-Konstante (das Rydberg) bezeichnet wird.  $R_y$  ist gleich der Ionisationsenergie des Wasserstoffatoms aus seinem Grundzustand :  $a_0$  ist der diesem Zustand entsprechende « Radius ». Es ist daher praktisch,  $a_0$  und  $R_y$  als Bezugsgrößen zu wählen.

<sup>3</sup>Unter diesem Gesichtspunkt erscheint  $\hbar$  wie die « fehlende Konstante » in der klassischen Mechanik. Es ist in der Tat unmöglich, eine *charakteristische Länge* nur mit den Konstanten  $m_e$  und  $q^2$  zu erhalten. Natürlich gibt es in der klassischen Mechanik Längen, aber diese Längen haben keinen « natürlichen Maßstab » : Zum Beispiel kann die Bahn eines Satelliten, der die Erde umkreist, gemäß seiner Geschwindigkeit irgend einen Radius zwischen dem der Erde und Unendlich besitzen.

<sup>4</sup>Es genügt, die drei Konstanten  $m_e$ ,  $\hbar$  und  $q^2$  mit Hilfe der Fundamentalgrößen Länge, Zeit und Masse auszudrücken und die Werte von  $\alpha$ ,  $\beta$  und  $\gamma$  so zu bestimmen, daß die Beziehung  $[m_e^\alpha \hbar^\beta q^{2\gamma}]$  einer Länge oder einer Energie entspricht. In beiden Fällen gibt es nur eine Lösung.

<sup>5</sup>Diese natürlichen Einheiten können übrigens « auf natürliche Weise » miteinander verknüpft werden :  $E_0$  ist die Wechselwirkungsenergie  $q^2/a_0$  von zwei Elementarladungen  $e$  im Abstand  $a_0$ .

**b. Ionisationsenergie und Atomradius.** Das Wasserstoffatom wird, wie wir soeben gesehen haben, in seinem Grundzustand durch einen Radius  $a_0$  und eine Ionisationsenergie  $R_y$  charakterisiert. Wenn man schwerere Atome mit Atomnummern  $Z$  (auch Kernladungszahl bzw. Ordnungszahl im Periodensystem der Elemente) betrachtet, so kann man durch einfache Überlegungen die Abhängigkeit von  $Z$  von Dimensionen und Energien herleiten.

Wir nehmen einfach an, daß man ein Atom aufbauen kann, indem man nacheinander einzelne Elektronen zum Kern hinzufügt. Für das erste Elektron erhält man das gleiche Ergebnis wie für das Wasserstoffatom unter Berücksichtigung, daß die Ladung des Kernes jetzt  $Ze$  ist; es genügt also  $q^2$  durch  $Zq^2$  in den Ausdrücken (A.9) und (A.7) zu ersetzen, um die Bindungsenergie des ersten Elektrons zu erhalten :

$$E_1 \simeq Z^2 R_y, \quad (\text{A.10})$$

und seine mittlere Entfernung vom Kern :

$$r_1 \simeq \frac{a_0}{Z}. \quad (\text{A.11})$$

Für große  $Z$  ist die Bindungsenergie daher beträchtlich größer und der Radius kleiner als beim Wasserstoffatom. Fügt man das nächste Elektron hinzu, wird es auch stark an den Kern gebunden sein; tatsächlich ist die Abstoßungsenergie zwischen den beiden Elektronen (die sich im Mittel in einer Entfernung der Größe  $r_1$  voneinander befinden)  $Z$ -mal schwächer als die Bindungsenergie eines jeden Elektrons an den Kern. Wenn man somit das  $(n + 1)$ -te Elektron hinzufügt, verhält sich das aus dem Kern und den  $n$  ersten Elektronen gebildete Ion wie ein « Kern » der Ladung  $(Z - n)e$ ; die Bindung der Elektronen an den Kern wird immer schwächer und ihr Abstand vom Kern immer größer. Wird das letzte Elektron addiert, so hat das Ion die Ladung  $e$ ; man erwartet, daß der « Radius » des Atoms in der Größenordnung von  $a_0$  und die Ionisationsenergie<sup>6</sup> in der Größenordnung von  $R_y$  ist.

Diese qualitativen Betrachtungen können nur eine Abschätzung der Phänomene geben; sie ersetzen auf keinen Fall eine korrekte quantitative Theorie. Nur durch Lösen der Schrödinger-Gleichung läßt sich zeigen, daß  $a_0$  genau der Radius des Wasserstoffatoms und  $R_y$  seine Ionisationsenergie ist. Wenn die Schrödinger-Gleichung nur die Größen  $m_e$ ,  $q^2$  und  $\hbar$  benötigt, so braucht

<sup>6</sup>Die Ionisationsenergie ist die kleinste Energie, die man aufwenden muß, um ein Elektron aus dem Atom zu entfernen; das am schwächsten gebundene Elektron wird demnach sich als erstes lösen.

der Atomradius nur noch in Einheiten von  $a_0$  ausgedrückt zu werden; allerdings kann die Proportionalitätskonstante von 1 ziemlich verschieden sein: Wenn wir zum Beispiel  $a_0$  mit  $h$  statt mit  $\hbar$  definiert hätten, würde die Schrödinger-Gleichung den Radius des Wasserstoffatoms nur auf den Faktor  $4\pi^2$  ( $\approx 40$ ) genau ergeben. Man kann also nie sicher sein, daß man sich bei einem Zahlenfaktor « täuscht ». Dennoch gibt eine durch die Praxis geschickt getroffene Wahl (zum Beispiel  $\hbar$  statt  $h$ )<sup>7</sup>, im allgemeinen die richtige Größenordnung wieder, das heißt bis auf einen Faktor 10 - 1/10 ein korrektes Resultat.

Dies läßt sich an der Tabelle I überprüfen, in der der Radius und die Ionisationsenergie von einigen Atomen angegeben ist: Sie sind in allen Fällen mit  $a_0$  und  $R_y$  vergleichbar. Dennoch bemerkt man signifikante Unterschiede zwischen den Atomen: Die Edelgase (He und Ne) sind besonders stabil; die Elemente, die in der Chemie als « elektropositiv » bezeichnet werden, geben leicht ein Elektron ab (Li, Na), die « elektronegativen » Elemente (O, Cl) weniger leicht.

TABELLE I.

**Ionisationsenergie und Radius von einigen Atomen.**

Atom	Ionisationsenergie (eV)	Atomradius (Å)
H	13,6	0,53
He	24,5	0,30
Li	5,4	1,67
C	11,2	0,65
O	13,6	0,46
Ne	21,5	0,37
Na	5,1	1,80
Cl	13,0	0,75

### 3. Moleküle und Festkörper

**a. Bindungsenergien und interatomare Abstände.** Es sei ein zweiatomiges Molekül gegeben; welche Größen legen seine Bindungsenergie und den Abstand zwischen den beiden Kernen fest?

In erster Näherung kann man die Kerne als unbeweglich, mit einem festen Abstand  $d$  untereinander, betrachten. Diese Kerne sind von « Wolken » der

<sup>7</sup>Diese Wahl ist offensichtlich um so leichter, wenn man das Ergebnis schon kennt.

äußeren Elektronen der beiden Atome umgeben; die Energie dieser Elektronen ist eine Funktion von  $d$ . Der stabile Zustand des Moleküls wird erhalten, indem man den Wert von  $d$  sucht, der die Gesamtenergie zum Minimum macht. Auch hier sind die einzigen beteiligten Fundamentalkonstanten  $m_e$ ,  $\hbar$  und  $q^2$ . Daraus folgt, daß die « natürlichen Einheiten » von Länge und Energie noch immer  $a_0$  und  $R_y$  sind, so daß die Längen und charakteristischen Energien des Moleküls in Größenordnungen von  $a_0$  und  $R_y$  angegeben werden können. Das bestätigt die Tabelle II. Man bemerkt jedoch, daß die

TABELLE II.

Dissoziationsenergie und interatomare Abstände von einigen zweiatomigen Molekülen.

Molekül	Dissoziationsenergie (eV)	Kernabstand (Å)
H <sub>2</sub>	4,5	0,75
N <sub>2</sub>	9,8	1,09
O <sub>2</sub>	5,1	1,20
Cl <sub>2</sub>	2,5	1,99
HCl	4,4	1,27
CO	8,4	1,13
NaCl	3,6	2,51

Dissoziationsenergien der Moleküle wesentlich schwächer als die Ionisationsenergien der Atome sind : Es ist im allgemeinen leichter ein Molekül zu dissoziieren als ein Atom zu ionisieren, andererseits sind die Abstände zwischen den Kernen systematisch größer als  $a_0$ <sup>8</sup>.

Die gleichen Überlegungen gelten offensichtlich für kompliziertere Moleküle und für Festkörper : Die Größenordnung der internuklearen Abstände ist das Angström, die der Kohäsionsenergien das Elektronenvolt (für einen

<sup>8</sup>Dieser Effekt kann der elektrostatischen Abstoßung zwischen zwei Kernen zugeschrieben werden. Wenn ihre Entfernung  $d$  um einige Å größer ist, wechselwirken sie praktisch nicht, da ihre Ladung durch die den Kern umhüllenden Elektronen « abgeschirmt » wird. Aber wenn  $d$  genügend klein ist, so daß die beiden « Elektronenwolken » sich durchdringen können, besetzen die äußeren Elektronen « Orbitale », die beide Kerne umhüllen und sich gegenseitig « sehen ». Für die Elektronen sind gewisse molekulare « Orbitale » um beide Kerne herum energetisch günstiger als atomare « Orbitale » um einen einzigen Kern. Dieses erlaubt schließlich die Bildung eines stabilen Moleküls. Gegen diese Tendenz wirkt aber die elektrostatische Abstoßung der Kerne, und zwar umso stärker, je kleiner  $d$  ist. Dadurch kommen die « großen » Werte von  $d$  (in Einheiten von  $a_0$ ) und die « kleinen » Werte der Dissoziationsenergie (in Einheiten von  $R_y$ ) zustande.

Festkörper ist die Kohäsionsenergie die Bindungsenergie durch Molekül; siehe Tabelle III).

TABELLE III

**Kohäsionsenergie und Abstand zwischen nächsten Nachbarn für einige Elemente im festen Zustand.**

Festkörper	Kohäsionsenergie (eV)	Abstand zwischen nächsten Nachbarn (Å)
C (Diamant)	7,36	1,5
Na	1,13	3,7
Al	3,34	2,9
Si	4,64	2,3
Ar	0,08	3,8
Fe	4,29	2,5
Te	2,0	2,9

**b. Schwingungs- und Rotationsenergien.** Neben den Elektronenanregungsenergien, von denen wir schon gesprochen haben, gibt es im Molekül Schwingungs- und Rotationsanregungen. Diese bringen die Bewegung des Moleküls in Abhängigkeit von seinem Massenmittelpunkt ins Spiel und hängen damit direkt von der Masse der Atomkerne ab.

Für ein zweiatomiges Molekül, das aus zwei Kernen der Massen  $M_1$  und  $M_2$  besteht, läßt die Separierung von Bewegung und Lage des Massenmittelpunktes einen einzigen Parameter für die Relativbewegung zu, die *reduzierte Masse*

$$M_R = \frac{M_1 M_2}{M_1 + M_2}. \quad (\text{A.12})$$

Wir drücken sie mit Hilfe ihrer natürlichen Einheit, der Masse  $M_N$  des Nukleons, aus :

$$M_R = \alpha M_N, \quad (\text{A.13})$$

wobei  $\alpha$  ein Parameter ist, der explizit von den Nukleonenzahlen  $A_1$  und  $A_2$  der beiden Kerne abhängt :

$$\alpha = \frac{A_1 A_2}{A_1 + A_2}. \quad (\text{A.14})$$

$\alpha$ . *Schwingungsenergien.* Die Schwingungen des Moleküls von geringer Amplitude werden durch die zweiten Ableitungen des Wechselwirkungspotentials zwischen den beiden Atomen in der Umgebung des Gleichgewichtsabstandes bestimmt<sup>9</sup>. Wir haben nun aber gesehen, daß der Maßstab dieses Potentials die Rydbergkonstante  $R_y$  ist, und daß es in den charakteristischen Längen in der Größenordnung von  $a_0$  variiert. Man kann daher in erster Näherung die Schwingungsbewegungen des Moleküls mit einem harmonischen Potential beschreiben, dessen Rückstellkonstante  $K$  die Größenordnung

$$K \simeq \frac{R_y}{a_0^2} \quad (\text{A.15})$$

hat<sup>10</sup>.

Die entsprechende Schwingungsfrequenz ist

$$\omega_v = \sqrt{\frac{K}{M_R}}. \quad (\text{A.16})$$

Man kann dieses Resultat aus praktischen Gründen in einer Energieeinheit ausdrücken, indem man  $\omega_v$  mit  $\hbar$  multipliziert<sup>11</sup>. Man erhält schließlich unter Benutzung der Ausdrücke (A.7) und (A.9)

$$\hbar\omega_v \simeq \hbar \sqrt{\frac{R_y}{a_0^2 M_R}} = R_y \sqrt{\frac{m_e}{M_N}} \cdot \frac{1}{\sqrt{\alpha}} \quad (\text{A.17})$$

<sup>9</sup>Es sei  $V(d)$  das Wechselwirkungspotential zwischen den beiden Atomen mit dem Abstand  $d$ . Entwickelt man  $V(d)$  bis zur 2. Ordnung in der Umgebung seines Gleichgewichtswertes  $d_0$  (was die erste Ableitung von  $V(d)$  Null werden läßt), so erhält man

$$V(d) = V(d_0) + \frac{1}{2}(d - d_0)^2 V''(d_0) + \dots \quad (1)$$

Für kleine Oszillationen um  $d_0$  verhält sich  $V(d)$  wie ein harmonisches Potential von konstanter Rückstellkraft

$$K = V''(d_0). \quad (2)$$

<sup>10</sup>Die einzigen charakteristischen Größen des Potentials  $V(d)$  sind  $R_y$  und  $a_0$ . Seine erste Ableitung  $V'(d)$ , dessen Dimension der Quotient aus einer Energie und einer Länge ist, wird in Einheiten von  $R_y/a_0$  gemessen; ebenso erhält man  $R_y/a_0^2$  für seine zweite Ableitung  $V''(d)$ . Wie immer wählen wir die Proportionalitätskonstante in der Größenordnung von 1 (wir merken an, daß in gewissen Fällen diese Konstante Null ist; dies ist der Fall für  $V'(d)$  an der Stelle  $d = d_0$ ).

<sup>11</sup>Wir erinnern daran, daß in der Quantenmechanik die Energieniveaus dieses harmonischen Oszillators durch

$$\epsilon_n^v = \left(n + \frac{1}{2}\right) \hbar\omega_v,$$

gegeben sind, wobei  $n$  alle positiven ganzzahligen Werte einschließlich Null annimmt.

(wir vernachlässigen einen Faktor 2, der ohne Bedeutung in einer Abschätzung dieser Größenordnung ist) mit

$$R_y \sqrt{\frac{m_e}{M_N}} \simeq 0,32 \text{ eV}. \quad (\text{A.18})$$

Die Tabelle IV gibt die experimentellen Werte von  $\hbar\omega_v$  für einige zweiatomige Moleküle an. In allen Fällen liegen diese Werte in der Nähe der durch (A.17) gegebenen Abschätzung<sup>12</sup>.

Für Moleküle, die mehr als zwei Atome enthalten, sind die Ergebnisse

TABELLE IV

**Schwingungs- und Rotationsquanten für einige zweiatomige Moleküle.**  $\alpha$  ist die reduzierte Masse (in der Masseinheit des Nukleons). In Klammern sind die durch die Abschätzungen (A.17) und (A.20) « vorausgesagten » Werte angegeben.

Molekül	reduzierte Masse $\alpha = M_R/M_N$	Schwingungsquant $\hbar\omega_v(\text{eV})$	Rotationsquant $\hbar\omega_r(\text{eV})$
H <sub>2</sub>	0,5	0,53 (0,45)	$7,5 \cdot 10^{-3}$ ( $15 \cdot 10^{-3}$ )
D <sub>2</sub>	1	0,39 (0,32)	$3,8 \cdot 10^{-3}$ ( $7 \cdot 10^{-3}$ )
N <sub>2</sub>	7	0,29 (0,12)	$2,5 \cdot 10^{-4}$ ( $10^{-3}$ )
O <sub>2</sub>	8	0,19 (0,11)	$1,8 \cdot 10^{-4}$ ( $9 \cdot 10^{-4}$ )
Cl <sub>2</sub>	17,8	0,07 (0,08)	$3 \cdot 10^{-5}$ ( $4 \cdot 10^{-4}$ )
HCl	0,97	0,36 (0,32)	$1,3 \cdot 10^{-3}$ ( $7 \cdot 10^{-3}$ )
NO	7,5	0,23 (0,12)	$2 \cdot 10^{-4}$ ( $10^{-3}$ )

komplizierter; es gibt dann mehrere Typen von Eigenschwingungen, aber die Größenordnungen bleiben die gleichen.

In einem Festkörper führt die Kopplung zwischen den Atomen zu einer sehr großen Anzahl von Eigenschwingungen (sogenannten Phononen), deren Frequenzen sehr dicht beieinander liegen. Sie bilden « Bänder » von praktisch kontinuierlichen Werten, deren Größe durch (A.17) gegeben ist. Die Tabelle V gibt für einige Festkörper eine der charakteristischen Schwingungsfrequenzen im Kristall an: die Debye-Frequenz  $\omega_D$  ( $\pi/\omega_D$  ist die Größenordnung der Zeit, die eine Schallwelle benötigt, um einen interatomaren Abstand zu durchlaufen).

<sup>12</sup>Man könnte für jedes Molekül die Abschätzung noch verfeinern, indem man in (A.15) die in Tabelle II angegebenen experimentellen Werte mit einbezieht.

TABELLE V

Debye-Frequenz  $\omega_D$  einiger Festkörper. In Klammern Abschätzung nach der Formel (A.17), in der  $\alpha$  durch  $A$  ersetzt wurde.

Festkörper	Atommasse $A$	Debye-Frequenz $\hbar\omega_D$ (eV)
Be	9	$8,6 \cdot 10^{-2}$ ( $11 \cdot 10^{-2}$ )
Na	23	$1,3 \cdot 10^{-2}$ ( $7 \cdot 10^{-2}$ )
Al	27	$3,4 \cdot 10^{-2}$ ( $6 \cdot 10^{-2}$ )
Si	28	$5,4 \cdot 10^{-2}$ ( $6 \cdot 10^{-2}$ )
Ar	40	$0,7 \cdot 10^{-2}$ ( $5 \cdot 10^{-2}$ )
Cu	63,5	$2,7 \cdot 10^{-2}$ ( $4 \cdot 10^{-2}$ )
NaCl	58,5	$2,8 \cdot 10^{-2}$ ( $4 \cdot 10^{-2}$ )

$\beta$ . *Rotationsenergien.* Die Rotation eines zweiatomigen Moleküls um seinen Massenmittelpunkt hängt von der reduzierten Masse  $M_R$  und dem interatomaren Abstand  $d$  der Größenordnung  $a_0$  ab. Nimmt man das Molekül als *starr* an, so sind  $M_R$  und  $d$  die einzigen Größen, die die Rotations-eigenschaften des Moleküls festlegen (in Form des Ausdruckes  $I = M_R d^2$ , dem Drehmoment in bezug auf den Massenmittelpunkt). Man verfügt daher, um eine charakteristische Frequenz dieser Rotation zu bestimmen, nur über die Konstanten  $M_R$ ,  $d$  und  $\hbar$ . Sie sind durch die folgende Beziehung in der Einheit einer Frequenz verknüpft<sup>13</sup>

$$\omega_r = \frac{\hbar}{M_R d^2}. \quad (\text{A.19})$$

Auch hier wird man eher das Energiequant  $\hbar\omega_r$  betrachten; es ist daher praktisch,  $\hbar\omega_r$  in Termen von  $R_y$  auszudrücken, indem man  $d$  durch  $a_0$  in (A.19) ersetzt. Es folgt sofort (auf einen Faktor 2 genau)

$$\hbar\omega_r \simeq R_y \frac{m_e}{M_N} \cdot \frac{1}{\alpha}, \quad (\text{A.20})$$

mit

$$R_y \frac{m_e}{M_N} \simeq 7,4 \cdot 10^{-3} \text{ eV}. \quad (\text{A.21})$$

<sup>13</sup>Man erhält die gleiche Abschätzung, indem man die Rotationsenergie  $I\omega_r^2/2$  näherungsweise gleich  $\hbar\omega_r$  schreibt.

Die Rotationsenergien<sup>14</sup> sind daher beträchtlich kleiner als die Schwingungsenergien. Dies läßt sich an Tabelle IV<sup>15</sup> verifizieren.

### c. Elektrische und magnetische Momente

*α. Elektrische Momente.* In gewissen Molekülen fällt der Schwerpunkt der positiven Ladungen nicht mit dem der negativen Ladungen zusammen; das Molekül verfügt daher über ein elektrisches Dipolmoment  $\vec{p}$ . Die entsprechenden Ladungen sind in der Größenordnungen von  $e$ , die Abstände in der Größenordnung von  $a_0$ ; die natürliche Einheit des mikroskopischen Dipolmomentes ist daher

$$p_0 = ea_0 \simeq 10^{-29} \text{ C} \cdot \text{m}. \quad (\text{A.22})$$

Offensichtlich ist für einige Moleküle das Dipolmoment in Abwesenheit eines elektrischen Feldes Null : Dies ist zum Beispiel für Moleküle der Fall, die aus zwei identischen Atomen bestehen ( $\text{N}_2, \text{O}_2, \text{H}_2, \dots$ ). Aber wenn ein Molekül ein permanentes Dipolmoment zeigt, wird dieses gleich einige  $10^{-29} \text{ C} \cdot \text{m}$  groß sein; der Proportionalitätsfaktor liegt im allgemeinen zwischen 10 und  $1/10$ . Die Tabelle VI gibt für einige Moleküle die Werte von  $p$  an<sup>16</sup>.

*β. Magnetische Momente.* Besitzt ein Atom oder ein Ion einen von Null verschiedenen Gesamtdrehimpuls, so verursacht die entsprechende Bewegung der geladenen Teilchen einen elektrischen Strom, und das Atom (oder Ion) zeigt ein permanentes magnetische Moment  $\vec{\mu}$ . Die natürliche Einheit

<sup>14</sup>Die Rotationsenergien eines Moleküls sind in der Quantenmechanik durch

$$\epsilon_j^r = j(j+1) \frac{\hbar^2}{2I}$$

gegeben, worin  $j$  alle positiven ganzen Zahlen einschließlich Null annimmt.

<sup>15</sup>(A.20) gibt die Größenordnung des experimentellen Wertes von  $\omega_r$  gut wieder. Die Abschätzung ist jedoch systematisch zu groß. Dies ist darauf zurückzuführen, daß der Abstand  $d$  zwischen den Kernen stets größer als  $a_0$  ist (siehe Tabelle II und Fußnote 8). Auch da könnte man die Abschätzung verfeinern, indem man die Angaben der Tabelle II benutzt.

<sup>16</sup>Man stellt fest, daß das Dipolmoment groß ist (im Verhältnis zur natürlichen Einheit), wenn eins der Atome stark < elektronegativ > (Cl) und das andere stark < elektropositiv > (Na oder Cs) ist. Das Dipolmoment des Moleküls  $\text{CO}_2$  ist aus Symmetriegründen Null : Die drei Atome sind linear angeordnet, die O's umgeben das C. Im  $\text{H}_2\text{O}$  bilden die Atome im Gegensatz dazu ein gleichschenkliges Dreieck und das < elektronegative > O-Atom zieht die Elektronen-< Wolke > an sich, was zu positiven Teilladungen auf den H-Atomen führt.

TABELLE VI

Elektrisches Dipolmoment einiger Moleküle.

Molekül	Dipolmoment ( $10^{-29} \text{C} \cdot \text{m}$ )
H <sub>2</sub> O	0,62
HCl	0,36
NaCl	3,00
CsCl	3,48
NH <sub>3</sub>	0,49
CO	0,04
CO <sub>2</sub>	0,0

des atomaren magnetischen Momentes ist das *Bohrsche Magneton*<sup>17</sup>

$$\mu_B = \frac{e\hbar}{2m_e} = 0,927 \cdot 10^{-23} \text{ J/T} \simeq 10^{-23} \text{ J/T}. \quad (\text{A.23})$$

<sup>17</sup>Man findet dieses Ergebnis mit der gewohnten Methode: Ein magnetisches Moment ist das Produkt aus der Stromstärke (gleich einer elektrischen Ladung durch Zeit) und der Fläche; hier tritt jetzt  $e$  und nicht mehr  $q^2$  auf; man muß also durch Verknüpfung von  $e$ ,  $m_e$  und  $\hbar$  die Einheit des magnetischen Momentes bilden. Auch da ist die Lösung eindeutig.

Man kann auch nach « physikalischen » Überlegungen so vorgehen, daß man die Stromstärke  $i$  aus der Bewegung einer Ladung  $e$  auf einer Kreisbahn vom Radius  $a_0$  ermittelt (das heißt die Ladung  $e$  multipliziert mit der Anzahl der Umdrehungen durch Sekunde):

$$i \simeq e \frac{v}{a_0}, \quad (1)$$

wobei  $v$  die « Geschwindigkeit » des Elektrons auf seiner Bahn von der Größenordnung

$$\frac{1}{2} m_e v^2 \simeq R_y \quad (2)$$

ist, woraus

$$\mu \simeq i a_0^2 \simeq a_0 e \sqrt{\frac{R_y}{m_e}} \quad (3)$$

folgt. Indem man  $R_y$  und  $a_0$  durch ihre Werte (A.7) und (A.9) ersetzt, findet man

$$\mu \simeq \frac{e\hbar}{m_e}, \quad (4)$$

das auf einen Faktor 2 gleich dem Bohrschen Magneton  $\mu_B$  ist. Wir weisen nochmals darauf hin: Das Voranstehende ist *keine Berechnung* des Wertes des Bohrschen Magnetons, sondern nur eine Methode, um die Größenordnung der atomaren magnetischen Momente zu ermitteln.

Für ein freies Atom (oder Ion) schreibt sich das magnetische Moment

$$\vec{\mu} = -\mu_B(\vec{L} + 2\vec{S}), \quad (\text{A.24})$$

wobei  $\vec{L}$  und  $\vec{S}$  der Gesamtbahndrehimpuls respektive der Gesamtspin des Atoms in Einheiten von  $\hbar$  sind. Die Projektion von  $\vec{\mu}$  auf eine Quantisierungsachse hat schließlich den Wert von einigen  $\mu_B$ . Die Situation ist für ein Atom in einem Kristall komplexer, aber die Größenordnung ist die gleiche.

Auch der Kern besitzt im allgemeinen einen Drehimpuls und daher ein magnetisches Moment. Man erhält seine Größenordnung durch Ersetzen von  $m_e$  durch  $M_N$  in (A.23). Die magnetischen Kernmomente sind daher ungefähr 1000 mal kleiner als die atomaren magnetischen Momente (sie betragen einige  $10^{-26}$  J/T).

#### 4. Klassische Mechanik oder Quantenmechanik?

Die Eigenschaften der Materie werden auf mikroskopischem Niveau durch die Quantenmechanik beschrieben. Es gibt jedoch Fälle, wo die klassische Mechanik ausreicht, eine korrekte Beschreibung der beobachteten Phänomene zu liefern. Diese klassische Beschreibung ist im allgemeinen leichter durchzuführen als die Quantenbeschreibung; es ist daher wichtig, ein einfaches Kriterium zu kennen, wann ihre Benutzung gerechtfertigt ist.

Die charakteristische Konstante der Quantenmechanik ist die *Plancksche Konstante*  $\hbar$ . Diese Konstante hat die Dimension einer *Wirkung*, das heißt eines Produktes von Zeit mal Energie, oder eines Produktes von Länge mal Bewegungsgröße oder Drehimpuls. Das Kriterium lautet folgendermaßen :

*Wenn alle natürlichen Variablen eines physikalischen Systems, die die Dimensionen einer Wirkung haben, sehr viel größer als die Plancksche Konstante  $\hbar$  sind, dann wird das System korrekt durch den Formalismus der klassischen Mechanik beschrieben.*

Dieses Kriterium wird nahezu immer von den Objekten makroskopischer Größe, aber auch von Fällen auf mikroskopischem Niveau erfüllt<sup>18</sup>. Man muß übrigens anmerken, daß die klassische Mechanik (verbessert durch einige einfache quantenmechanische Begriffe) oft mit Erfolg benutzt werden kann, um Abschätzungen von Größenordnungen zu erhalten, sogar in Fällen, wo es fraglich wäre, sie für quantitative Berechnungen einzusetzen; dies haben wir bereits mehrfach in dieser Ergänzung getan.

<sup>18</sup>Gewisse makroskopische Ausnahmesysteme wie supraflüssiges (suprafluides) Helium, erfüllen nicht das Kriterium und zeigen im makroskopischem Maßstab typische Quanteneigenschaften. Umgekehrt zeigen wir im Abschnitt III.2, daß die meisten Gase auch auf mikroskopischem Niveau mit der klassischen Mechanik behandelt werden können.

## II. Der Übergang zum makroskopischen Bereich

### 1. Makroskopische Größenordnungen

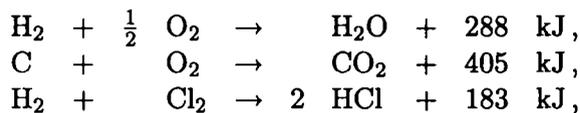
Die Größenordnung der meisten makroskopischen physikalischen Größen wird durch Multiplikation der Werte der mikroskopischen Größen mit der Avogadro-Zahl \*

$$N_A = 6,02 \cdot 10^{23} \quad (\text{A.25})$$

erhalten.

Wir geben einige Beispiele :

- (i) Bei einer chemischen Reaktion dissoziieren die Moleküle, um neue Moleküle zu bilden. Die mit dieser Neuordnung von Atomen verbundenen Energien müssen in der Größenordnung der Bindungsenergie der Moleküle liegen, das heißt von 1 bis 10 eV pro Elementarprozess. Die chemischen Reaktionswärmen sind somit in der Größenordnung von (1 bis 10)  $\cdot N_A$  eV durch Mol, das heißt  $10^5$  bis  $10^6$  Joules durch Mol. Man kennt zum Beispiel



etc.

- (ii) Ebenso ist die erforderliche Energie, um ein Mol eines Kristalles bei konstanter Temperatur und konstantem Druck zu verdampfen, die als *latente Sublimationswärme* bezeichnet wird, von der gleichen Größenordnung wie die Kohäsionsenergie des Kristalles multipliziert mit der Avogadro-Zahl. Nach der Tabelle III erwartet man (0,1 bis 10)  $\cdot N_A$  eV durch Mol, das sind  $10^4$  bis  $10^6$  J durch Mol.<sup>19</sup> Dies kann man gut in der Tabelle VII.a nachlesen.

Die *latenten Verdampfungswärmen* (Übergang vom flüssigen Zustand zum gasförmigen Zustand) sind geringfügig kleiner, aber die Größenordnung ist ebenfalls nach obiger Überlegung gegeben (Tabelle VII.b).

---

\* Anmerkung des Übersetzers : Die *Avogadro-Konstante* entspricht dieser Zahl mal der Einheit  $\text{mol}^{-1}$  (s. I, (I.1)).

<sup>19</sup> *Achtung* : Die latente Wärme ist nicht so einfach mit der Kohäsionsenergie verbunden. Tatsächlich besitzen die Atome eines Kristalls eine Schwingungsenergie thermischen Ursprungs, und erst die Differenz zwischen der Kohäsionsenergie und dieser Schwingungsenergie ergibt die latente Wärme. Letztere hängt daher von der Temperatur ab; dennoch erhält man eine korrekte Größenordnung, wenn man nur von der Kohäsionsenergie ausgeht.

TABELLE VII

Latente Wärmen. VII.a : Latente Sublimationswärmen für einige Metalle (die Temperatur ist nicht angegeben, denn zwischen 0 K und dem Schmelzpunkt ändert sich  $L_s$  nur um einige Prozent). VII.b : Latente Schmelz- ( $L_f$ ) und Verdampfungswärme ( $L_v$ ) für einige Stoffe am Schmelzpunkt ( $T_f$ ) bzw. am Siedepunkt ( $T_v$ ) bei einem Druck von einer Atmosphäre.

Metall	$L_s$ (kJ/mol)	Festkörper oder Flüssigkeit	$T_f$ (K)	$L_f$ (kJ/mol)	$T_v$ (K)	$L_v$ (kJ/mol)
K	82	N <sub>2</sub>	63	0,7	77	5,6
Fe	436	O <sub>2</sub>	55	0,5	90	6,8
Cu	338	Cl <sub>2</sub>	172	6,4	238	18,4
Ag	284	Ar	84	1,2	87,3	6,3
W	878	Hg	234	2,3	630	54,5
Au	367	H <sub>2</sub> O	273	6,0	373	40,6
Pb	197	NH <sub>3</sub>	198	7,7	240	23,2

VII.a

VII.b

Die *latenten Schmelzwärmen* (Übergang von fest zu flüssig) sind systematisch kleiner. Sie betragen zwischen 0,5 und 50 kJ durch Mol (Tabelle VII.b). Dies erklärt sich aus der Tatsache, daß in einer Flüssigkeit die Bindungsenergie zwischen Molekülen nicht vernachlässigt werden kann wie in einem Gas : Sie ist sicherlich kleiner als die Kohäsionsenergie in einem Festkörper, aber nur geringfügig.

- (iii) Gewisse Stoffe, die als *ferroelektrisch* bezeichnet werden, zeigen eine permanente elektrische Polarisation (elektrisches Dipolmoment durch Volumen) im makroskopischen Maßstab. Andere, die *ferromagnetisch* sind, verfügen über eine von Null verschiedene Magnetisierung (magnetisches Moment durch Volumen). Diese Eigenschaften werden dadurch verursacht, daß die mikroskopischen Momente sich auf makroskopische Entfernungen hin parallel zueinander ausrichten statt sich zufällig zu verteilen. Somit ist das Moment durch Volumen gleich dem Produkt aus einem mikroskopischen Moment und der Anzahl dieser Momente durch Volumen. Für einen interatomaren Abstand  $d_0$  der Länge 2 Å beträgt die Größenordnung der Polarisation eines ferroelektrischen Stoffes

$$\mathcal{P}_0 = p_0 \cdot \frac{1}{d_0^3} \simeq 1 \text{ C/m}^2, \quad (\text{A.26})$$

und die der Magnetisierung eines ferromagnetischen Stoffes

$$\mathcal{M}_0 = \mu_B \cdot \frac{1}{d_0^3} \simeq 10^6 \text{ A/m}. \quad (\text{A.27})$$

Zum Beispiel zeigen die Kristalle von Bariumtitanat ( $\text{BaTiO}_3$ ) eine Polarisation von  $0,26 \text{ C/m}^2$  bei Zimmertemperatur; die Magnetisierung des Eisens bei der gleichen Temperatur beträgt  $1,7 \cdot 10^6 \text{ A/m}$ .

Natürlich wird durch solche Betrachtungen nicht erklärt, warum das Eisen zum Beispiel ferromagnetisch ist, noch viel weniger, warum es oberhalb von  $1043 \text{ K}$  paramagnetisch wird. Doch lassen sie hoffen, in dem Maße wie die Größenordnungen vergleichbar sind, die makroskopischen Phänomene, wie den Ferromagnetismus, mittels der magnetischen Momente der atomaren Bestandteile zu verstehen.

## 2. Temperaturskala

Der Übergang zum makroskopischen Bereich läßt sich noch subtiler mit Hilfe des Temperaturbegriffes beschreiben. Dieser Begriff wird im Teil II definiert und in den nachfolgenden Kapiteln ausführlich untersucht; wir teilen hier nur mit, daß die *absolute Temperatur*  $T$  (Einheit Kelvin, abgekürzt K) definitionsgemäß gleich der um  $273,15$  vergrößerten Temperatur  $t$  in Grad Celsius  $^\circ\text{C}$  ist :

$$T(\text{K}) = t(^\circ\text{C}) + 273,15. \quad (\text{A.28})$$

Mit jedem Wert der Temperatur kann man eine mikroskopische *charakteristische Energie* verknüpfen, die durch das Produkt  $kT$  aus der absoluten Temperatur  $T$  und der Boltzmann-Konstanten  $k$  gegeben ist

$$k = 1,38 \cdot 10^{-23} \text{ J/K}. \quad (\text{A.29})$$

Wir werden sehen, daß die Energie  $kT$  für ein System bei der Temperatur  $T$  eine *Bezugsgrößenordnung* liefert, die mit der Energie des Systems in Zusammenhang gebracht werden kann<sup>20</sup>. Es ist daher wichtig, schnell eine Temperatur in eine Energie oder umgekehrt « übersetzen » zu können, und dafür sich die folgenden Ziffern zu merken :

<sup>20</sup>Die Tatsache, daß sich eine Äquivalenz zwischen Temperatur und Energie herstellen läßt, bedeutet *weder*, daß die Temperatur eine Energie ist, noch daß die Energie eines makroskopischen Systems bei der Temperatur  $T$  das Produkt von  $kT$  mit der Anzahl von Atomen ist, die es enthält (wir werden sehen, daß die Realität komplexer ist). Dennoch stimmt es manchmal (aber nicht immer!), daß die mittlere Energie pro Atom eines Systems bei der Temperatur  $T$  in der *Größenordnung* von  $kT$  ist. Dies kann zur Abschätzung von Größenordnungen benutzt werden. Diese Eigenschaften und seine Gültigkeit werden im einzelnen im Kapitel III (§V.C) untersucht.

- 1 K ist äquivalent zu  $\simeq 10^4$  eV ( $0,8617 \cdot 10^4$  eV)
- 1 eV ist äquivalent zu  $\simeq 12\,000$  K (11 605 K)
- die Zimmertemperatur ( $20^\circ\text{C} \simeq 293$  K) ist äquivalent zu  $\simeq 1/40$  eV.

Zur Abschätzung der Skala der absoluten Temperaturen verweisen wir schließlich auf Tabelle VIII, in der einige Bezugstemperaturen (in K und eV) angegeben sind.

### 3. Elektrische und magnetische Felder

Wir beenden diesen Abschnitt, indem wir einige Größenordnungen der im Labor zugänglichen elektrischen und magnetischen Felder angeben.

Auf *makroskopischem Niveau* sind die stärksten gegenwärtig realisierbaren elektrischen Felder in der Größenordnung von  $10^7$  V/m; die stärksten kontinuierlichen magnetischen Felder betragen mehr als 30 Tesla (das heißt 300 000 Gauß), können aber für sehr kurze Zeitintervalle Werte bis zu 100 T erreichen.

Diese Werte bleiben weit unter denen der entsprechenden Felder auf *mikroskopischem Niveau*. Zum Beispiel beträgt das elektrische Feld, das das Elektron im Wasserstoffatom festhält,

$$E_{at} \simeq \frac{e}{4\pi\epsilon_0 a_0^2} \simeq 5 \cdot 10^{11} \text{ V/m}. \quad (\text{A.30})$$

Dieser Wert gestattet uns, einen oberen Grenzwert für die im Labor erreichbaren makroskopischen Felder zu bestimmen: Wenn das elektrische Feld eines Leiters den Wert (A.30) erreicht, wird die Kraft des elektrostatischen Druckes auf den Leiter von der gleichen Größenordnung sein wie die Kohäsionskraft, die die Atome zusammenhält, und der Leiter wird zerstört. Für den Betrag des maximal erreichbaren magnetischen Feldes  $B_{\max}$  gilt, daß die entsprechende magnetische Kraft von der gleichen Größenordnung ist wie die elektrische Kraft, die das Elektron festhält:

$$evB_{\max} \simeq eE_{at}; \quad (\text{A.31})$$

wegen

$$\frac{1}{2} m_e v^2 \simeq R_y, \quad (\text{A.32})$$

erhält man

$$B_{\max} \simeq E_{at} \sqrt{\frac{m_e}{R_y}} \simeq 3 \cdot 10^5 \text{ T}. \quad (\text{A.33})$$

TABELLE VIII

Einige Temperaturen in K und in eV.

	$T(K)$	$kT(eV)$
Zentrum der Sonne	$> 5 \cdot 10^6$	$> 400$
Oberfläche der Sonne	$\sim 5500$	0,47
Schmelzen von Wolfram	3650	0,31
Schmelzen von Eisen	1808	0,16
Schmelzen von Blei	600	$5,2 \cdot 10^{-2}$
Sieden von Wasser	373	$3,2 \cdot 10^{-2}$
Schmelzen von Eis	273	$2,4 \cdot 10^{-2}$
Sieden von Stickstoff	77	$6,6 \cdot 10^{-3}$
Schmelzen von Stickstoff	63,3	$5,5 \cdot 10^{-3}$
Sieden von Wasserstoff	20,3	$1,7 \cdot 10^{-3}$
Schmelzen von Wasserstoff	14	$1,2 \cdot 10^{-3}$
Sieden von Helium 4	4,2	$3,6 \cdot 10^{-4}$
Übergang zur Supraflüssigkeit von Helium-4	2,17	$1,9 \cdot 10^{-4}$
Übergang zur Supraflüssigkeit von Helium-3	$2,6 \cdot 10^{-3}$	$2,2 \cdot 10^{-7}$

Die Materie, die die zur Realisierung eines solchen magnetischen Feldes erforderlichen Ströme transportiert, ist nicht mehr stabil.

Tatsächlich treten bereits die Probleme für die elektrischen wie für die magnetischen Felder unterhalb dieser theoretischen Grenzen auf; deshalb sind die im Labor erhaltenen maximalen Werte deutlich kleiner.

### III. Beispiel: Klassisches ideales Gas

#### 1. Definitionen

Ein Ensemble von  $N$  identischen Teilchen der Masse  $m$ , eingeschlossen in einem Rezipienten vom Volumen  $V$ , stellt ein Gas dar, wenn die kinetische Energie der Teilchen gegenüber ihrer potentiellen Wechselwirkungsenergie groß ist<sup>21</sup>.

Dieses Gas wird

- (i) *klassisch* genannt, wenn man die Bewegung der enthaltenen Teilchen mit dem Formalismus der klassischen Mechanik beschreiben kann;

<sup>21</sup>Überragt umgekehrt die potentielle Energie bei weitem die kinetische Energie, so handelt es sich um einen Festkörper. In einer Flüssigkeit sind beide Energien von gleicher Größenordnung (was die theoretische Untersuchung beträchtlich kompliziert).

(ii) *ideal* genannt, wenn die potentielle Energie vernachlässigbar ist.

Wenn die zweite Bedingung ohne die erste erfüllt wird, so hat man es mit einem *idealen Quantengas* zu tun.

Es ist klar, daß die beiden voranstehenden Bedingungen um so besser erfüllt sind, je *verdünnter* das Gas ist; doch muß man diese Aussage für jede der beiden Bedingungen genauer untersuchen.

## 2. Gültigkeit der klassischen Beschreibung

**a. Gültigkeitsbedingung.** Wie wir im Abschnitt I.4 gesehen haben, müssen wir  $\hbar$  mit den physikalischen Größen des Systems, die die Dimensionen einer Wirkung haben, vergleichen. Für die Bewegung der Moleküle eines Gases gibt es eine einzige Größe, die man durch das Produkt aus dem Impuls  $mv$  eines Teilchens und dem Abstand  $d$  zwischen zwei Teilchen erhält (das Produkt aus der kinetischen Energie eines Teilchens und der Zeit, die zwischen zwei Stößen verfließen ist, ist systematisch größer). Die Bedingung (i) läßt sich dann durch die Beziehung

$$mvd \gg \hbar. \quad (\text{A.34})$$

ausdrücken.

Offensichtlich ist der Variationsbereich von  $d$  und  $v$  *a priori* sehr groß :  $d$  kann zwischen dem Durchmesser eines Teilchens und der größten Ausdehnung des Rezipienten liegen,  $v$  zwischen Null und Unendlich. Aber die Geschwindigkeiten und Abstände zwischen den Teilchen liegen praktisch immer in der Nähe der Mittelwerte  $\bar{v}$  und  $\bar{d}$  von  $v$  und  $d$ . Es genügt daher, daß die Bedingung

$$m\bar{v}\bar{d} \gg \hbar \quad (\text{A.35})$$

erfüllt wird. Indem wir die de Broglie-Wellenlänge einführen, die mit der mittleren Geschwindigkeit verknüpft ist :

$$\lambda^* := \frac{\hbar}{m\bar{v}}, \quad (\text{A.36})$$

kann man diese Bedingung in die Form

$$\lambda^* \ll \bar{d}. \quad (\text{A.37})$$

umschreiben.

Der mittlere Abstand  $\bar{d}$  wird leicht erhalten :

$$\bar{d} = \left( \frac{V}{N} \right)^{1/3}. \quad (\text{A.38})$$

Im Teil III (§V.C) werden wir zeigen, daß die mittlere Geschwindigkeit  $\bar{v}$  von der Temperatur des Gases abhängt; die Größenordnung ist dabei<sup>22</sup>

$$\frac{1}{2}m\bar{v}^2 \simeq kT. \quad (\text{A.39})$$

Die Bedingung (A.37) ist daher gleichbedeutend mit

$$kT \gg \frac{\hbar^2}{m} \left( \frac{N}{V} \right)^{2/3}. \quad (\text{A.40})$$

**b. Ein Molekülgas.** Betrachten wir zum Beispiel Wasserstoff bei Zimmertemperatur und gewöhnlichem Druck. Die Masse eines Moleküls beträgt zweimal die des Nukleons und die damit verknüpfte de Broglie-Wellenlänge beträgt bei Raumtemperatur (300 K) ungefähr

$$\lambda^* \simeq \frac{\hbar}{\sqrt{4M_N kT}} \simeq 0,20 \text{ \AA}. \quad (\text{A.41})$$

Andererseits gibt die Zustandsgleichung für ein Gas, das als ideal und klassisch betrachtet wird (vgl. Fußnote 22)<sup>23</sup>,

$$pv = NkT, \quad (\text{A.42})$$

bei einem Druck  $p$  von 1 Atmosphäre ( $10^5$  Pa) ist

$$\bar{d} = \left( \frac{V}{N} \right)^{1/3} = \left( \frac{kT}{p} \right)^{1/3} \simeq 34 \text{ \AA}. \quad (\text{A.43})$$

Die Bedingung (A.37) ist damit gut verifiziert, und die klassische Näherung ausgezeichnet.

Die Moleküle anderer Gase verfügen über viel größere Massen; die de Broglie-Wellenlänge wird dabei kleiner und die Näherung besser. *Alle gewöhnlichen Gase werden gut durch die klassische Mechanik beschrieben*<sup>24</sup>.

<sup>22</sup>Tatsächlich ist die Beziehung (A.39) nur gültig, wenn das Gas sich wie ein klassisches Gas verhält, das heißt, wenn die Bedingung (A.37) erfüllt ist. Solche Überlegungen sind in der Physik geläufig, und wir werden sie im folgenden noch oft benutzen. Sie bestehen darin, die Eigenschaften eines Systems mittels einer Näherung, die man anschließend mit Hilfe des erhaltenen Resultates rechtfertigt (oder wieder aufhebt), zu analysieren.

<sup>23</sup>Dieses Resultat gilt für ein beliebiges klassisches ideales Gas. Es kann daher nützlich sein, sich in Erinnerung zu rufen, daß unter gewöhnlichen Bedingungen der mittlere Abstand zwischen zwei Molekülen eines Gases bei 30 Å liegt.

<sup>24</sup>Die klassische Näherung könnte sich bei tiefen Temperaturen oder bei hohen Drücken als schlecht erweisen (vgl. (A.41) und (A.43)). Tatsächlich stellt sich dieses Problem nur bei Helium, das bei tiefer Temperatur (oder hohem Druck) gasförmig ist.

**c. Elektronengas in einem Metall.** Ein kristalliner Festkörper besteht aus im Raum (in den Knoten eines Gitternetzes) regelmäßig angeordneten Atomen. Wenn dieser Festkörper ein Metall ist, so halten sich die äußeren Elektronen der verschiedenen Atome nicht an einer Stelle auf, sondern bewegen sich quasi frei im gesamten Kristall; dieser besteht daher aus einem Gitter von *positiven Ionen*, die in ein « *Gas* » von Leitungselektronen eingebettet sind.

Die Masse des Elektrons ist sehr viel kleiner (um einen Faktor  $2 \cdot 1836$ ) als die Masse des im Abschnitt *b* betrachteten Wasserstoffmoleküls; seine de Broglie-Wellenlänge ist daher sehr viel größer<sup>25</sup> als (A.41) :

$$\lambda^* \simeq 0,20\sqrt{2 \cdot 1836} = 12 \text{ \AA} . \quad (\text{A.44})$$

Im Gegensatz dazu ist der Abstand zwischen den Teilchen sehr viel kleiner : wenn es ein Leitungselektron pro Ion gibt (das ist der Fall zum Beispiel im Natrium oder Kupfer), dann ist  $\bar{d}$  der Abstand zwischen benachbarten Ionen in dem Metall,

$$\bar{d} \simeq 2 \text{ \AA} . \quad (\text{A.45})$$

Die Bedingung (A.37) ist daher nicht erfüllt. Die Elektronen eines Metalles bilden ein zu dichtes Gas, als daß eine klassische Beschreibung adäquat erscheint. Eine quantenmechanische Behandlung ist unerlässlich, indem man explizit das Pauli-Ausschließungsprinzip berücksichtigt. Man sagt, daß die Leitungselektronen eines Metalles ein *entartetes Quantengas* darstellen.

### 3. Vergleich zwischen kinetischer und potentieller Energie

Wir beschränken die Untersuchung der zweiten Bedingung auf Molekülgase. Wie wir soeben gesehen haben, läßt sich diese Untersuchung im klassischen Rahmen durchführen.

Die kinetische Energie der Moleküle eines Gases hängt von der Temperatur ab. Ihre Größenordnung haben wir bereits angegeben (Formel (A.39)), und sie ist

$$E_{\text{kin}} \simeq NkT . \quad (\text{A.46})$$

Schwieriger ist es nun, die potentielle Energie der Wechselwirkung zu berechnen. Die Wechselwirkung zwischen zwei Molekülen eines Gases hängt von ihrem relativen Abstand ab. Wird dieser Abstand kleiner als ihre eigene Größenordnung, so üben die beiden Moleküle aufeinander äußerst große

<sup>25</sup>Wie im Fall des Wasserstoffs berechnen wir  $\lambda^*$ , indem wir die Beziehung (A.39) als erfüllt annehmen. Tatsächlich ist sie es für die Elektronen eines Metalles nicht; sie wäre es, wenn die Elektronen sich wie ein klassisches ideales Gas verhalten würden. (vgl. Fußnote 22).

Abstoßungskräfte aus (es ist im wesentlichen das Pauli-Prinzip, dem die Elektronen gehorchen; es macht die Moleküle füreinander undurchdringbar). Bei großen Abständen ziehen sie sich hingegen schwach an (van der Waals-Kräfte). Das Wechselwirkungspotential  $u(r)$  zwischen zwei Molekülen als Funktion ihres relativen Abstandes  $r$  hat daher das in der Abbildung 1 dargestellte Aussehen.

Für einfache Moleküle ist es möglich,  $u(r)$  mit Hilfe der Quantenme-

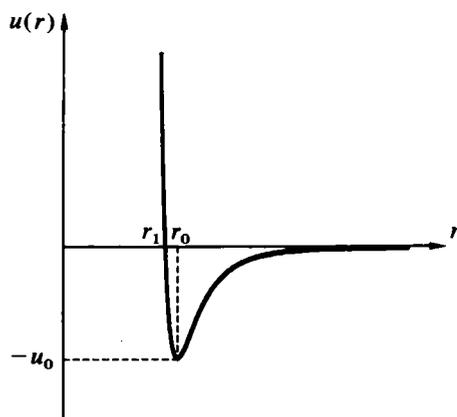


ABBILDUNG 1

Typisches Wechselwirkungspotential zwischen zwei Molekülen als Funktion ihres relativen Abstandes  $r$ .

chanik zu berechnen. Aber man verwendet im allgemeinen semi-empirische Beziehungen der Form

$$u(r) = -u_0 f(r/r_0); \quad (\text{A.47})$$

$u(r)$  ist im wesentlichen durch zwei mikroskopische Parameter bestimmt, die das Minimum festlegen:  $r_0$  und  $u_0$ ;  $f(x)$  ist eine Funktion ohne Dimension, die die Gestalt der Kurve in Abbildung 1 besitzt.

Wir nennen  $r_1$  den Abstand, bei dem  $u(r)$  Null wird. In erster Näherung<sup>26</sup> können wir sagen, daß sich ein gegebenes Molekül mit der Wahrscheinlichkeit

<sup>26</sup>Diese Approximation ist um so besser gerechtfertigt, je größer die kinetische Energie der Moleküle gegenüber der potentiellen Energie ist (genau diese Bedingung versuchen wir hier anzugeben). Tatsächlich haben zwei Moleküle wegen ihrer Anziehungskräfte großer Reichweite eine höhere Wahrscheinlichkeit, einander nahe zu sein als voneinander entfernt zu sein; aber dieser Effekt wird vernachlässigbar, wenn die kinetische Energie groß wird.

Null einem Molekül auf einen Abstand kleiner als  $r_1$  nähert und sich stets mit der gleichen Wahrscheinlichkeit in einer Entfernung größer als  $r_1$  an irgendeinem Punkt des Rezipienten befindet. Die Wahrscheinlichkeit  $dP(r)$  zwei Moleküle in einer Entfernung  $r$  voneinander (mit ungefähr  $dr$ ) anzutreffen, ist somit

$$dP(r) \simeq \begin{cases} 0 & \text{für } r < r_1 \\ \frac{4\pi r^2 dr}{V} & \text{für } r > r_1. \end{cases} \quad (\text{A.48})$$

Man leitet daraus die mittlere Wechselwirkungsenergie zwischen zwei Molekülen ab<sup>27</sup>:

$$\bar{u} \simeq - \int_{r_1}^{\infty} u_0 f(r/r_0) \frac{4\pi r^2}{V} dr, \quad (\text{A.49})$$

was man auch mit Substitution der Variablen  $x = r/r_0$  schreiben kann

$$\bar{u} \simeq - \frac{4\pi r_0^3}{V} u_0 \int_{x_1}^{\infty} f(x) x^2 dx; \quad (\text{A.50})$$

$x_1 = r_1/r_0$  ist eine Zahl, die geringfügig kleiner als 1 ist. Die Funktion  $f(x)$  hat einen Träger von einigen Einheiten; sie variiert derart zwischen 0 und 1 im Integrationsintervall (siehe Abbildung 1), daß das Integral in (A.50) eine Zahl wird, die geringfügig kleiner als 1 ist. In Größenordnungen erhält man schließlich

$$\bar{u} \simeq - \frac{u_0 r_0^3}{V}. \quad (\text{A.51})$$

Wenn  $N$  die Anzahl der Moleküle des Gases ist, dann gibt es

$$\frac{1}{2} N(N-1) \simeq \frac{N^2}{2} \quad (\text{A.52})$$

Paare von Molekülen, und die mittlere Wechselwirkungsenergie der Gesamtheit des Gases beträgt

$$U \simeq - \frac{N^2}{2} u_0 \frac{r_0^3}{V}. \quad (\text{A.53})$$

Die potentielle Wechselwirkungsenergie wird daher gegenüber der kinetischen Energie zu vernachlässigen sein, und das Gas kann als ideal betrachtet werden, wenn

$$\frac{kT}{u_0} \cdot \frac{V}{Nr_0^3} \gg 1. \quad (\text{A.54})$$

<sup>27</sup>Die Reichweite der Funktion  $u(r)$  beträgt nur einige  $r_0$ , das heißt sie ist von mikroskopischer Größenordnung. Man kann daher, ohne den Wert von  $\bar{u}$  zu ändern, die Integration in (A.49) bis ins Unendliche ausdehnen.

Somit erscheint  $u_0/k$  als natürliche Temperaturskala und  $r_0^3$  als der natürliche Maßstab von Volumen durch Teilchen  $V/N$ .

Mit festgehaltener Dichte  $N/V$  verhält sich ein Gas bei hoher Temperatur wie ein ideales Gas. Bei festgehaltener Temperatur ist die Approximation um so besser, je verdünnter das Gas ist.

Es bleiben noch die Parameter  $u_0$  und  $r_0$  genauer zu beschreiben. Da sie Parameter aus dem atomaren Bereich sind, erwartet man für  $u_0$  die Größenordnung von eV, für  $r_0$  die Größenordnung von Å. Tatsächlich ist  $u_0$ , das die Wechselwirkungsenergie zwischen zwei *chemisch stabilen* Molekülen darstellt, für gewöhnliche Gase sehr viel kleiner, und  $r_0$  systematisch größer:  $u_0$  ist in der Größenordnung von nur einigen Hundertstel eV und  $r_0$  liegt zwischen 3 und 10 Å gemäß der Größe des Moleküls. Zum Beispiel beträgt  $r_0$  3,9 Å für Argon, 4,2 Å für Stickstoff und 4,8 Å für Kohlenstoffgas; die entsprechenden Werte von  $u_0$  sind  $10^{-2}$  eV (118 K),  $8 \cdot 10^{-3}$  eV (95 K) und  $1,7 \cdot 10^{-2}$  eV (198 K)<sup>28</sup>.

Bei Zimmertemperatur (300 K) und unter Atmosphärendruck erhält man mit  $10^{-2}$  eV für  $u_0$  und 4 Å für  $r_0$  (vgl. (A.43))

$$\left(\frac{kT}{u_0}\right)\left(\frac{V}{Nr_0^3}\right) \simeq 1500. \quad (\text{A.55})$$

Die Näherung des idealen Gases ist daher in diesem Fall gerechtfertigt. Sie kann jedoch für tiefere Temperaturen oder höhere Drücke schlecht werden.

Zusammenfassend kann man sagen, daß sich ein Gas wie ein klassisches ideales Gas verhält, wenn es genügend verdünnt ist. Dann sind die beiden Bedingungen (A.40) und (A.54) gleichzeitig erfüllt. Anzumerken ist, daß praktisch für alle gewöhnlichen Gase die Bedingung (A.54) vor der Bedingung (A.40) verletzt wird: Von außergewöhnlichen Umständen abgesehen wird ein Gas durch die klassische Mechanik korrekt beschrieben; nur bei nicht zu tiefen Temperaturen und bei genügend geringer Dichte kann man die Wechselwirkungen vernachlässigen, und die Approximation des idealen Gases ist dann gerechtfertigt.

---

<sup>28</sup>Die Werte von  $u_0$  und  $r_0$  hängen noch von der für die Darstellung von  $u(r)$  gewählten Funktion  $f(x)$  ab. Die hier angegebenen Werte entsprechen dem Modell von Lennard-Jones, für das gilt

$$f(x) = \frac{2}{x^6} - \frac{1}{x^{12}}.$$

# Statistische Verteilung einer Funktion von Zufallsvariablen

---

I. Statistische Verteilung  
einer Größe Funktion einer  
einzigsten Zufallsvariablen

1. Wahrscheinlichkeitsdichte  
einer streng monotonen  
Funktion

- a. Streng wachsende  
Funktion
- b. Streng fallende  
Funktion

2. Allgemeiner Fall einer  
beliebigen Funktion

II. Statistische Verteilung  
einer Größe Funktion von  
mehreren Zufallsvariablen

1. Verteilungsfunktion

2. Wahrscheinlichkeitsdichte

3. Sonderfall :  $A(\vec{r})$  als  
Summe von unabhängigen  
Zufallsvariablen

III. Beispiele

- 1. Projektion eines Zufallsvek-  
tors mit konstantem Betrag
- 2. Kinetische Energie eines  
Teilchens mit zufälliger  
Geschwindigkeit

---

Es sei  $x$  eine kontinuierliche Zufallsvariable, von der die Wahrscheinlichkeitsdichte  $w(x)$  bekannt ist. Wir betrachten eine physikalische Größe  $A$  als Funktion dieser Variablen :  $A = A(x)$ . Wie lautet die entsprechende Wahrscheinlichkeitsdichte  $W(A)$ ? Anders gefragt, wie lautet die Wahrscheinlichkeit

$$dP(A) = W(A)dA \tag{B.1}$$

dafür, daß die Größe  $A$  dem infinitesimalen kleinen Intervall  $[A, A + dA]$  angehört?

Die Formeln (24) und (25) im Anhang I zeigen, daß sich diese Wahrscheinlichkeit angeben läßt mit

$$W(A)dA = \int_{dD(A)} w(x)dx, \tag{B.2}$$

wobei  $dD(A)$  das infinitesimale Gebiet der  $x$ -Werte ist, so daß gilt

$$A \leq A(x) \leq A + dA. \tag{B.3}$$

Wir zeigen in dieser Ergänzung, wie man die Wahrscheinlichkeitsdichte  $W(A)$  einer Größe  $A$  praktisch berechnen kann, die zunächst Funktion einer einzigen, dann von mehreren Zufallsvariablen ist<sup>1</sup>. Wir enden, indem wir explizit ein Beispiel für jeden Fall behandeln.

## I. Statistische Verteilung einer Größe Funktion einer einzigen Zufallsvariablen

Um den Gedankengang klar herauszustellen, untersuchen wir zunächst den Fall, in dem  $A$  eine monotone Funktion der Zufallsvariablen  $x$  ist, bevor wir den allgemeinen Fall behandeln.

### 1. Wahrscheinlichkeitsdichte einer streng monotonen Funktion

**a. Streng wachsende Funktion.** Wenn  $A(x)$  eine streng wachsende Funktion von  $x$  ist, so ist die Relation zwischen  $A$  und  $x$  eine eindeutig umkehrbare Abbildung (eindeutige Abbildung), und  $x$  ist selbst eine eindeutig definierte und wachsende Funktion von  $A$  :

$$A = A(x) \text{ wachsend} \iff x = x(A) \text{ wachsend.} \tag{B.4}$$

Die Abbildung 1.a definiert nun eindeutig den Bereich  $dD(A)$  der Formel (B.2), der sich als

$$W(A)dA = \int_{x(A)}^{x(A+dA)} w(x')dx' \tag{B.5}$$

schreibt. Die Variablensubstitution  $x' = x(A')$  ergibt

$$W(A)dA = \int_A^{A+dA} w[x(A')] \frac{dx}{dA'} dA'. \tag{B.6}$$

Aber das Integrationsintervall  $dA$  ist infinitesimal, so daß

$$W(A)dA = w[x(A)] \frac{dx}{dA} dA. \tag{B.7}$$

Man kann diese Beziehung noch kürzer in der Form

$$W(A)dA = w(x)dx, \tag{B.8}$$

schreiben, unter Berücksichtigung, daß auf der rechten Seite selbstverständlich  $x = x(A)$  ist.

---

<sup>1</sup>Wir nehmen an, daß die Werte von  $A$  reelle Zahlen sind. Wenn  $A$  eine Größe darstellt, die einem Raum höherer Dimension als 1 angehört ( $A$  ist ja ein Vektor dieses Raumes), so benutzt man die gleiche Methode, indem man nacheinander jede Komponente untersucht.

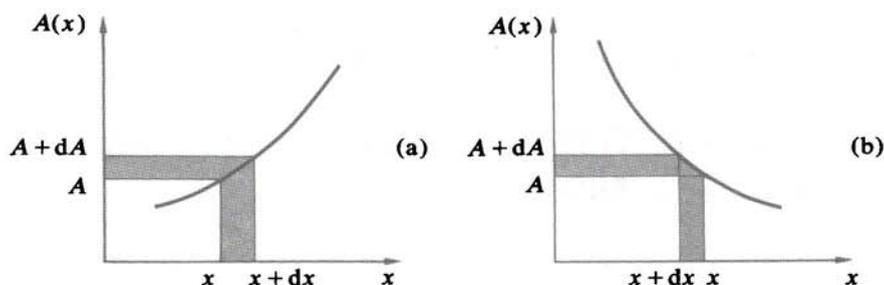


ABBILDUNG 1

Bereich  $dD(A)$  der Werte von  $x$ , so daß  $A \leq A(x) \leq A + dA$  ist für eine wachsende Funktion  $A(x)$  (Figur a) und für eine fallende Funktion  $A(x)$  (Figur b).

**b. Streng fallende Funktion.** Wenn  $A(x)$  eine streng fallende Funktion ist, dann bleibt die Relation zwischen  $A$  und  $x$  eindeutig, aber die Funktion  $x(A)$  ist eine *fallende Funktion von  $A$*  :

$$A = A(x) \text{ fallend} \iff x = x(A) \text{ fallend.} \quad (\text{B.9})$$

Die Abbildung 1.b zeigt, daß jetzt der Integrationsbereich  $dD(A)$  von  $(x+dx)$  bis  $x$  geht<sup>2</sup>, so daß

$$W(A)dA = \int_{x(A+dA)}^{x(A)} w(x')dx'. \quad (\text{B.10})$$

Die gleiche Variablensubstitution  $x' = x(A')$  ergibt dann

$$W(A)dA = \int_{A+dA}^A w[x(A')] \frac{dx}{dA'} dA', \quad (\text{B.11})$$

das heißt

$$W(A)dA = -w[x(A)] \frac{dx}{dA} dA, \quad (\text{B.12})$$

was man analog zu (B.8) in der Form

$$W(A)dA = -w(x)dx \quad (\text{B.13})$$

umschreiben kann.

<sup>2</sup>Da  $x$  eine fallende Funktion von  $A$  ist, wird der infinitesimale Zuwachs  $dx$ , geknüpft an  $dA$  ( $dA > 0$ ), hier negativ. Die relative Lage der Integrationsgrenzen in (B.10) gewährleistet schließlich, daß die Wahrscheinlichkeit  $W(A)dA$  positiv ist.

## 2. Allgemeiner Fall einer beliebigen Funktion

Wenn  $A$  keine monotone Funktion von  $x$  ist, dann definiert  $A = A(x)$  nicht nur eine Funktion  $x$  von  $A$ , sondern genau so viele Funktionen  $x_i(A)$  wie die Gleichung

$$A - A(x) = 0 \quad (\text{B.14})$$

Lösungen  $x_i$  für ein gegebenes  $A$  besitzt (siehe Abbildung 2). Jedes  $x_i(A)$

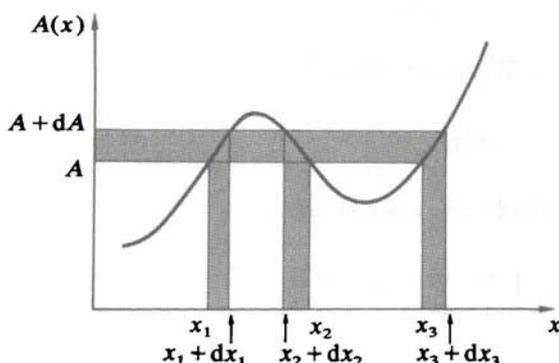


ABBILDUNG 2

Bereich  $dD(A)$  der Werte von  $x$ , so daß  $A \leq A(x) \leq A + dA$  ist für eine beliebige Funktion  $A(x)$ . Hier ist der Bereich  $dD(A)$  in drei Teile zerlegt, denn die Gleichung  $A - A(x) = 0$  hat drei Lösungen in  $x$  für die betrachteten Werte von  $A$ :  $x_1$ ,  $x_2$  und  $x_3$ .

trägt einen Anteil zu  $W(A)dA$  vom Typ (B.7) oder (B.12) bei, vorausgesetzt, daß es sich um eine wachsende oder fallende Funktion von  $A$  handelt. Danach läßt sich die gesuchte Wahrscheinlichkeit angeben als<sup>3, 4</sup>

$$W(A)dA = \sum_i w[x_i(A)] \left| \frac{dx_i}{dA} \right| dA. \quad (\text{B.15})$$

<sup>3</sup>Der Absolutwert der in (B.15) auftretenden Ableitung verbietet eine vereinfachte Schreibweise dieser Formel in Analogie zu (B.8) oder (B.13). Es ist dennoch praktisch, (B.8) und (B.13) als Gedächtnisstütze zu benutzen, um zu (B.15) zu gelangen.

<sup>4</sup>Die Formel (B.15) setzt voraus, daß die Funktion  $A(x)$  keine Treppen hat, das heißt  $A(x)$  nimmt für das endliche Intervall  $[x_1, x_2]$ , zu dem alle  $x$  gehören, einen konstanten Wert  $A_0$  an. Für diesen Fall wäre die Wahrscheinlichkeit, diesen Wert  $A_0$  der Variablen  $A$  zu erhalten, endlich und gleich  $\int_{x_1}^{x_2} w(x)dx$ .

Wir zeigen, wie man dieses Ergebnis aus den Eigenschaften der « Dirac-Funktion » erhalten kann (diese Methode werden wir im Abschnitt II verallgemeinern, um den Fall von Funktionen mit mehreren Zufallsvariablen zu behandeln).

Definieren wir zuerst die *Verteilungsfunktion*  $P(A)$  der Zufallsvariablen  $A$ :  $P(A)$  ist die Wahrscheinlichkeit, daß die Zufallsvariable einen kleineren Wert als den gegebenen Wert  $A$  annimmt.  $P(A)$  berechnet sich nach dem « Additionsgesetz » (Anhang I, §I.2.a) :

$$P(A) = \int_{D(A)} w(x) dx, \quad (\text{B.16})$$

wobei  $D(A)$  der Bereich der  $x$ -Werte ist, so daß

$$A(x) \leq A. \quad (\text{B.17})$$

Unter Benutzung der « Heaviside-Funktion »  $\theta(x)$  :

$$\theta(x) := \begin{cases} 0 & \text{für } x < 0 \\ 1 & \text{für } x \geq 0, \end{cases} \quad (\text{B.18})$$

läßt sich (B.16) in die Form umschreiben<sup>5</sup>

$$P(A) = \int_{-\infty}^{+\infty} w(x) \theta[A - A(x)] dx. \quad (\text{B.19})$$

Andererseits ist die Wahrscheinlichkeit  $dP(A)$ , daß  $A$  zum infinitesimalen Intervall  $[A, A + dA]$  gehört, offensichtlich

$$dP(A) = P(A + dA) - P(A), \quad (\text{B.20})$$

so daß die Wahrscheinlichkeitsdichte  $W(A)$  als die Ableitung der Verteilungsfunktion erscheint :

$$W(A) = \frac{dP(A)}{dA}. \quad (\text{B.21})$$

Mit der Ableitung von  $P(A)$  nach  $A$  in (B.19) erhält man

$$W(A) = \int_{-\infty}^{+\infty} w(x) \delta[A - A(x)] dx. \quad (\text{B.22})$$

---

<sup>5</sup>Wenn  $w(x)$  nur in dem endlichen Intervall  $[x_1, x_2]$  definiert ist, so kann man es erweitern, indem man die Funktion außerhalb dieses Intervalles Null setzt.

Tatsächlich ist die Ableitung der Heaviside-Funktion  $\theta(x)$  nichts anderes als die Dirac-Funktion  $\delta(x)$ . Um  $W(A)$  zu berechnen, benötigt man daher nur die Kenntnis<sup>6</sup> von

$$\delta[f(x)] = \sum_i \frac{1}{|f'(x_i)|} \delta(x - x_i), \quad (\text{B.23.a})$$

wobei  $f'(x_i)$  die Ableitung von  $f(x)$  an der Stelle  $x_i$  ist; die Summe erstreckt sich über die Gesamtheit der  $x_i$ , da gilt

$$f(x_i) = 0. \quad (\text{B.23.b})$$

Man erhält schließlich unter Benutzung von (B.22) und (B.23)

$$W(A) = \sum_i \frac{1}{|A'(x_i)|} w(x_i), \quad (\text{B.24.a})$$

wobei sich die Summe über alle  $x_i$  erstreckt, so daß

$$A(x_i) - A = 0. \quad (\text{B.24.b})$$

Man kann sich leicht davon überzeugen, daß (B.24) mit (B.15) identisch ist.

## II. Statistische Verteilung einer Größe Funktion von mehreren Zufallsvariablen

Die physikalische Größe  $A$  ist jetzt eine Funktion von mehreren Zufallsvariablen  $(x, y, z, \dots)$ , deren Verteilung durch die Wahrscheinlichkeitsdichte  $w(x, y, z, \dots)$  charakterisiert wird. Um die Schreibweise zu erleichtern, benutzen wir die vektorielle Notation (Anhang I, §I.1.b) :

$$w(x, y, z, \dots) dx dy dz \dots := w(\vec{r}) d\vec{r} \quad (\text{B.25})$$

und

$$A(x, y, z, \dots) := A(\vec{r}). \quad (\text{B.26})$$

---

<sup>6</sup>Siehe zum Beispiel C.D.L., Anhang II.

## 1. Verteilungsfunktion

Die Verteilungsfunktion  $P(A)$  ist die Wahrscheinlichkeit dafür, daß die Zufallsvariable  $A(\vec{r})$  einen Wert annimmt, der kleiner ist als der Wert  $A$ ; sie ist durch eine Beziehung analog zu (B.16) gegeben :

$$P(A) = \int_{D(A)} w(\vec{r}) d\vec{r}, \quad (\text{B.27})$$

wobei sich das Integral über das « Volumen »  $D(A)$  erstreckt, das als eine Punktmenge  $\vec{r}$  so definiert wird, daß gilt

$$A(\vec{r}) \leq A. \quad (\text{B.28})$$

Auch hier kann man die Heaviside-Funktion benutzen, um (B.27) in die äquivalente Form

$$P(A) = \int w(\vec{r}) \theta[A - A(\vec{r})] d\vec{r}, \quad (\text{B.29})$$

umzuschreiben, wobei sich das Integral jetzt über alle Werte von  $\vec{r}$  erstreckt.

## 2. Wahrscheinlichkeitsdichte

Die Wahrscheinlichkeitsdichte  $W(A)$  berechnet sich wie im vorhergehenden Abschnitt aus der Verteilungsfunktion :

$$W(A) = \frac{dP(A)}{dA}. \quad (\text{B.30})$$

Indem man explizit die Ableitung nach  $A$  in (B.29) ausdrückt, erhält man

$$W(A) = \int w(\vec{r}) \delta[A - A(\vec{r})] d\vec{r}. \quad (\text{B.31})$$

Aber diesmal kann das Integral nicht so leicht berechnet werden wie bei einer Funktion mit einer Variablen. Tatsächlich gestattet die Dirac-Funktion, eine Integration über eine der  $n$  Komponenten des Vektors  $\vec{r}$  auszuführen, aber es bleibt noch ein « Flächen- » Integral über die restlichen  $(n-1)$ -Variablen zu berechnen. Dieses Flächenintegral läßt sich schreiben<sup>7</sup>

$$W(A) = \int_{s(A)} ds \frac{w(\vec{r})}{|\nabla A(\vec{r})|}, \quad (\text{B.32})$$

<sup>7</sup>Die Wahrscheinlichkeit  $dP(A)$ , daß die Zufallsvariable  $A(\vec{r})$  zum Intervall  $[A, A + dA]$  gehört, ist

$$dP(A) = \int_{dD(A)} d\vec{r} w(\vec{r}), \quad (1)$$

wobei  $|\nabla A(\vec{r})|$  der Betrag des Gradientenvektors von  $A(\vec{r})$  ist; das Integral erstreckt sich über die Fläche  $s(A)$  der Gleichung

$$A(\vec{r}) = A. \quad (\text{B.33})$$

Um die Wahrscheinlichkeitsdichte  $W(A)$  zu berechnen, kann man da-

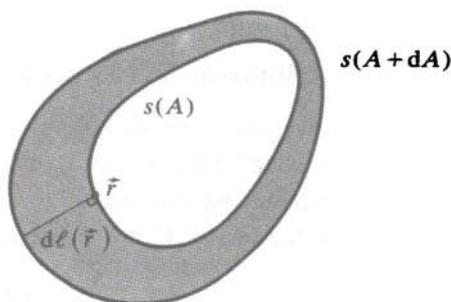


ABBILDUNG 3

Menge  $dD(A)$  der Punkte  $\vec{r}$ , so daß  $A \leq A(\vec{r}) \leq A + dA$ , begrenzt durch die Flächen  $s(A)$  und  $s(A + dA)$ ;  $d\ell(\vec{r})$  ist der Abstand zwischen diesen beiden Flächen im Punkt  $\vec{r}$ .

her entweder das Volumenintegral (B.27) berechnen, um die Verteilungswahrscheinlichkeit  $dD(A)$  das Volumen zwischen der Fläche  $s(A)$  der Gleichung

$$A(\vec{r}) = A \quad (2)$$

und der Fläche  $s(A + dA)$  mit der Gleichung

$$A(\vec{r}) = A + dA \quad (3)$$

ist.

Es sei nun  $d\ell(\vec{r})$  der Abstand zwischen diesen beiden Flächen, gemessen in orthogonaler Richtung von  $s(A)$  zum Punkt  $\vec{r}$  (siehe Abbildung 3). Das Volumenelement  $dV$  der Formel (1) kann als das Produkt von  $d\ell(\vec{r})$  mit dem Flächenelement  $ds$  von  $s(A)$  im Punkt  $\vec{r}$  geschrieben werden, so daß

$$dP(A) = \int_{s(A)} ds d\ell(\vec{r}) w(\vec{r}). \quad (4)$$

Nun ist aber  $d\ell(\vec{r})$  mit  $dA$  durch

$$dA = |\nabla A| d\ell(\vec{r}) \quad (5)$$

verknüpft, denn  $\nabla A$  ist orthogonal zur Fläche  $s(A)$ . Es genügt daher (5) in (4) einzusetzen, um (B.32) zu erhalten.

funktion zu bestimmen und daraus  $W(A)$  durch Ableiten ermitteln, oder direkt die Gleichungen (B.31) oder (B.32) lösen. Wir bemerken (wenn die Rechnung korrekt durchgeführt wurde!), daß das für die Wahrscheinlichkeitsdichte  $W(A)$  erhaltene Resultat automatisch normiert ist<sup>8</sup> :

$$\int_{-\infty}^{+\infty} W(A) dA = 1. \quad (\text{B.34})$$

### 3. Sonderfall : $A(\vec{r})$ als Summe von unabhängigen Zufallsvariablen

Um die Berechnungen nicht unnötig zu erschweren, nehmen wir an, daß  $A(\vec{r})$  eine Summe von *zwei* unabhängigen Zufallsvariablen  $x$  und  $y$  ist. Die Erweiterung auf eine beliebige Anzahl von Variablen ist evident.

Wir versuchen nun, die Wahrscheinlichkeitsdichte  $W(A)$  der Funktion

$$A(\vec{r}) \equiv A(x, y) = x + y \quad (\text{B.35})$$

mit den beiden *unabhängigen* Zufallsvariablen  $x$  und  $y$  zu berechnen. Die Wahrscheinlichkeitsdichte  $w(r)$  ist nun (Anhang I, §I.2.b) das Produkt aus den Wahrscheinlichkeitsdichten  $u(x)$  und  $v(x)$  der beiden Variablen  $x$  und  $y$  :

$$w(x, y) = u(x) \cdot v(y). \quad (\text{B.36})$$

Um das Integral (B.31) zu berechnen, ist es praktisch, die folgende Integraldarstellung der « Dirac-Funktion » zu benutzen (Fußnote 6) :

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk e^{-ikx}. \quad (\text{B.37})$$

Man erhält somit

$$W(A) = \frac{1}{2\pi} \int dx \int dy u(x)v(y) \int_{-\infty}^{+\infty} dk e^{ik(x+y-A)}, \quad (\text{B.38})$$

---

<sup>8</sup>Dieses Resultat läßt sich leicht, ausgehend von (B.31), verifizieren :

$$\begin{aligned} \int W(A) dA &= \int dA \int d\vec{r} w(\vec{r}) \delta[A - A(\vec{r})] \\ &= \int d\vec{r} w(\vec{r}) \int dA \delta[A - A(\vec{r})] \\ &= \int d\vec{r} w(\vec{r}). \end{aligned}$$

und wenn wir die Reihenfolge der Integrationen vertauschen,

$$W(A) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk e^{-ikA} \int_{-\infty}^{+\infty} dx u(x) e^{ikx} \int_{-\infty}^{+\infty} dy v(y) e^{iky}. \quad (\text{B.39})$$

Diese Formel ist einfach zu interpretieren, wenn wir den Begriff der *charakteristischen Funktion* einführen : Für eine beliebige Zufallsvariable  $z$  mit der Wahrscheinlichkeitsdichte  $w(z)$ , ist die charakteristische Funktion  $\tilde{w}(k)$  per Definition *der Mittelwert von  $e^{ikz}$* , das heißt (bis auf einen Faktor genau) die *Fourier-Transformierte* von  $w(z)$  :

$$\tilde{w}(k) := \int dz e^{ikz} w(z). \quad (\text{B.40})$$

Natürlich kann  $w(z)$  auch, ausgehend von der charakteristischen Funktion, durch eine inverse Fourier-Transformation ermittelt werden :

$$w(z) = \frac{1}{2\pi} \int dk e^{-ikz} \tilde{w}(k). \quad (\text{B.41})$$

Die Formel (B.39) drückt nun die folgende Regel aus : Die charakteristische Funktion  $\tilde{W}(k)$  der physikalischen Größe  $A$  ist das Produkt der charakteristischen Funktionen der unabhängigen Zufallsvariablen  $x$  und  $y$  :

$$\tilde{W}(k) = \tilde{u}(k) \cdot \tilde{v}(k). \quad (\text{B.42})$$

### III. Beispiele

#### 1. Projektion eines Zufallsvektors mit konstantem Betrag

Es sei  $\vec{A}$  ein Vektor der Ebene mit konstantem Betrag  $A_0$  und zufälliger Richtung. Wir nehmen alle Richtungen als gleichwahrscheinlich an; das heißt, daß die entsprechende Wahrscheinlichkeitsdichte  $w(\varphi)$  konstant ist, wenn  $\varphi$  ( $0 \leq \varphi < 2\pi$ ) den Winkel zwischen dem Vektor  $\vec{A}$  und einer gegebenen Richtung  $Ox$  darstellt. Die Normierungsbedingung ergibt

$$w(\varphi) = \frac{1}{2\pi}, \quad (0 \leq \varphi < 2\pi). \quad (\text{B.43})$$

Was ist nun die Wahrscheinlichkeit dafür, daß die Projektion des Vektors  $\vec{A}$  auf die Achse  $Ox$  zwischen  $A_x$  und  $A_x + dA_x$  liegt?

Die Abbildung 4 zeigt, daß es zwei Werte für  $\varphi$  gibt :  $\varphi_1$  und  $\varphi_2 = 2\pi - \varphi_1$ , die einem gegebenen Wert von  $A_x$  entsprechen :

$$A_x = A_0 \cos \varphi_1 = A_0 \cos \varphi_2. \quad (\text{B.44})$$

Die Anwendung von (B.15) ergibt sofort

$$W(A_x)dA_x = \frac{1}{2\pi} \left[ \frac{1}{|A_0 \sin \varphi_1|} + \frac{1}{|A_0 \sin \varphi_2|} \right] dA_x. \quad (\text{B.45})$$

Unter Berücksichtigung von

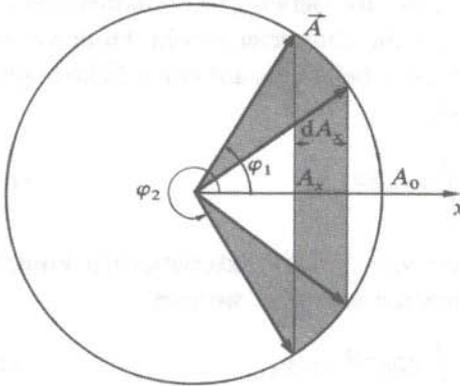


ABBILDUNG 4

Vektor  $\vec{A}$  mit konstantem Betrag  $A_0$  und zufälliger Richtung. Dem Wert  $A_x$  seiner Projektion auf die Achse  $Ox$  entsprechen zwei Werte der Polarwinkel  $\varphi_1$  und  $\varphi_2$ , ( $\varphi_2 = 2\pi - \varphi_1$ ).

$$|\sin \varphi_1| = |\sin \varphi_2| = \left[ 1 - \left( \frac{A_x}{A_0} \right)^2 \right]^{1/2}, \quad (\text{B.46})$$

erhält man schließlich

$$W(A_x)dA_x = \frac{dA_x}{\pi[A_0^2 - A_x^2]^{1/2}}, \quad (-A_0 \leq A_x \leq A_0). \quad (\text{B.47})$$

Die Wahrscheinlichkeitsdichte  $W(A_x)$  (in Abbildung 5 dargestellt) wird unendlich für  $A_x = \pm A_0$ . Dieses Ergebnis war aus Abbildung 4 vorhersehbar: In der Umgebung dieser Extremalwerte entspricht einem kleinen Zuwachs  $dA_x$  ein bedeutender Zuwachs von  $d\varphi$ . Offensichtlich bleibt  $W(A_x)$  integrierbar; insbesondere läßt sich leicht die Normierungsbedingung verifizieren, indem man das zweite Glied von (B.47) über das Intervall  $[-A_0, A_0]$  integriert.

## 2. Kinetische Energie eines Teilchens mit zufälliger Geschwindigkeit

Wir betrachten ein freies Teilchen der Masse  $m$ , dessen Geschwindigkeitskomponenten  $v_x, v_y, v_z$  auf drei orthogonalen Achsen zufällig verteilt sind

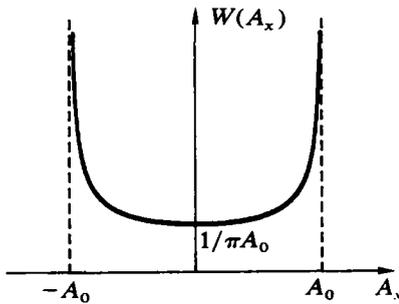


ABBILDUNG 5

Wahrscheinlichkeitsdichte  $W(A_x)$  für die Projektion  $A_x$  des Zufallsvektors  $\vec{A}$  aus Abbildung 4.

mit einer Wahrscheinlichkeitsdichte

$$w(v_x, v_y, v_z) = \left(\frac{\alpha}{\pi}\right)^{3/2} e^{-\alpha(v_x^2 + v_y^2 + v_z^2)}. \quad (\text{B.48})$$

Was ist die Wahrscheinlichkeit  $dP(E) = W(E)dE$  dafür, daß die kinetische Energie  $E$  dieses Teilchens zwischen  $E$  und  $E + dE$  liegt?

In diesem sehr einfachen Fall wird der Leser ohne Schwierigkeiten<sup>9</sup> zeigen können, daß die Beziehungen (B.30), (B.31) oder (B.32) zu folgendem Ergebnis führen :

$$dP(E) = W(E)dE = \frac{2}{\sqrt{\pi}} \left(\frac{2\alpha}{m}\right)^{3/2} \sqrt{E} e^{\frac{2\alpha E}{m}} dE, \quad (\text{B.49})$$

was die Normierungsbedingung

$$\int_0^\infty W(E)dE = 1. \quad (\text{B.50})$$

erfüllt.

<sup>9</sup>Es genügt, auf sphärische Koordinaten im Geschwindigkeitsraum überzugehen unter Berücksichtigung von

$$E = \frac{1}{2}m(v_x^2 + v_y^2 + v_z^2) = \frac{1}{2}mv^2.$$

# Die Binomialverteilung im Grenzfall großer Zahlen

---

I. Hypothesen

2. Entwicklung von  $\ln P$  um den wahrscheinlichsten Wert

II. Erstellung des Näherungsausdrucks

3. Resultat : Gauß-Approximation der Binomialverteilung

1. Näherungsweise Bestimmung des wahrscheinlichsten Wertes von  $n$

III. Diskussion der Approximation

---

Im Anhang I behandeln wir zwei besondere statistische Verteilungen : die Binomialverteilung und die Gauß-Verteilung. Wir werden hier zeigen, daß die Binomialverteilung sich im Bereich großer Zahlen wie eine Gauß-Verteilung verhält.

## I. Hypothesen

Wir betrachten ein beliebiges Problem mit diskreten Wahrscheinlichkeiten. Aus der Menge der möglichen Ereignisse wählen wir ein besonderes aus, das wir mit (+) kennzeichnen, und das die Wahrscheinlichkeit  $P_+$  hat; das Gegenereignis wird mit (-) gekennzeichnet, seine Wahrscheinlichkeit ist per Definition  $P_- = 1 - P_+$ . Wenn man  $N$  unabhängige Ziehungen aus diesen Ereignissen durchführt, dann ist die Wahrscheinlichkeit  $P(N, n)$  dafür, daß  $n$  dieser Experimente das Ergebnis (+) und die  $(N - n)$  anderen das Ergebnis (-) liefern, durch das Binomialgesetz gegeben (Anhang I, §II.1) :

$$P(N, n) = \frac{N!}{n!(N - n)!} P_+^n P_-^{N-n}. \quad (\text{C.1})$$

Der Mittelwert der Anzahl  $n$  ist einfach

$$\bar{n} = NP_+, \quad (\text{C.2})$$

und seine Standardabweichung

$$\Delta n = \sqrt{NP_+P_-}. \quad (\text{C.3})$$

Wir nehmen hier  $N$  als *sehr groß gegenüber 1* an. Wir nehmen außerdem an, daß die Wahrscheinlichkeit  $P_+$  weder zu klein ist, noch zu nahe an 1 liegt<sup>1</sup>, so daß  $n$  ebenfalls *groß gegenüber 1 und weit von  $N$  entfernt ist* :

$$\bar{n} = NP_+ \gg 1; \quad N - \bar{n} = N(1 - P_+) \gg 1. \quad (\text{C.4})$$

Diese beiden Bedingungen können zu einer zusammengefaßt werden :

$$NP_+P_- \gg 1. \quad (\text{C.5})$$

Wir interessieren uns allein für *Werte, die nicht zu weit von  $\bar{n}$  entfernt liegen*; genauer, wir beschränken das Variationsintervall von  $n$  auf einige Standardabweichungen  $\Delta n$  von  $\bar{n}$  :

$$|n - \bar{n}| \lesssim a\Delta n, \quad a \text{ in der Größenordnung von einigen Einheiten} \quad (\text{C.6})$$

(wir beschränken uns auf diesen Bereich, denn die Wahrscheinlichkeit  $P(N, n)$  wird schnell sehr klein, wenn  $n$  sich von  $\bar{n}$  entfernt; siehe Abschnitt III).

Wir schlagen schließlich noch vor,  *$n$  wie eine kontinuierliche Variable zu behandeln*. Eine solche Näherung ist gerechtfertigt, wenn die zu untersuchende Funktion sich wenig zwischen  $n - 1$  und  $n$  ändert, das heißt wenn gilt

$$|P(N, n) - P(N, n - 1)| \ll P(N, n); \quad (\text{C.7})$$

dann kann man die Differenz zwischen  $n$  und  $n - 1$  als infinitesimal betrachten. Diese Bedingung ist tatsächlich erfüllt, wenn die voranstehenden Hypothesen verifiziert werden<sup>2</sup>.

<sup>1</sup>Wenn die Wahrscheinlichkeit  $P_+$  (oder  $P_-$ ) sehr klein gegenüber 1 ist, dann gibt es ein Zwischengebiet, in dem  $N$  groß ist, ohne daß  $NP_+$  (oder  $NP_-$ ) es ist. Der Näherungsausdruck, den wir hier für die Binomialverteilung angeben werden, ist in diesem Bereich dann nicht mehr gültig. Man kann dennoch zeigen, daß die Binomialverteilung darin einer « *Poisson-Verteilung* » nahe kommt (siehe zum Beispiel W. Feller, *An Introduction to Probability Theory and its Applications*, John Wiley, New York (1950), S. 153). Offensichtlich wird die Bedingung  $NP_+P_- \gg 1$  stets für genügend große Werte von  $N$  erfüllt, wenn  $P_+$  fest ist, selbst wenn es sehr klein ist oder sehr nahe bei 1 liegt; die Gauß-Näherung wird dann also gültig.

<sup>2</sup>Man findet leicht, ausgehend von (C.1), das Verhältnis

$$\frac{P(N, n)}{P(N, n - 1)} = \frac{N - n + 1}{n} \frac{P_+}{P_-}. \quad (1)$$

Die uns interessierenden Werte von  $n$  sind so, daß

$$n = \bar{n} + \nu, \quad (2)$$

mit

$$|\nu| \ll N \quad (3)$$

## II. Erstellung des Näherungsausdrucks

Wir werden hier am Beispiel der Binomialverteilung eine Untersuchungsmethode explizit darstellen, die wir noch öfters in diesem Buch auf analoge Fälle anwenden werden. Diese Methode benutzt die *Eigenschaften der logarithmischen Funktion*.

### 1. Näherungsweise Bestimmung des wahrscheinlichsten Wertes von $n$

Wir betrachten den Logarithmus der Wahrscheinlichkeit  $P(N, n)$  :

$$\ln P(N, n) = \ln N! - \ln n! - \ln(N - n)! + n \ln P_+ + (N - n) \ln P_- . \quad (\text{C.8})$$

Diese Operation hat folgende praktische Bedeutung : Die Fakultät einer großen Zahl (man denke an die Avogadro-Konstante!) ist enorm groß; das Logarithmieren gestattet, mit Größenordnungen umzugehen, die unserem Urteilsvermögen und Gefühl zugänglicher sind<sup>3</sup>.

Für ein *konstantes*  $N$  wird der wahrscheinlichste Wert  $n_m$  von  $n$  derjenige sein, der  $P(N, n)$  zum Maximum macht. Aber da der Logarithmus eine wachsende Funktion ist, kann man  $n_m$  bestimmen, indem man fordert, daß  $\ln P(N, n)$  ein Maximum annimmt. Um den Ausdruck (C.8) abzuleiten,

(mit der Einschränkung (C.6),  $|\nu|$  ist höchstens in der Größenordnung von einigen  $\sqrt{N}$ ). Trägt man dem Wert von  $\bar{n}$  Rechnung, so erhält man schließlich

$$\frac{P(N, n)}{P(N, n-1)} = \frac{\bar{n}(1 - P_+) - (\nu - 1)P_+}{(\bar{n} + \nu)P_-} = \frac{1 - (\nu - 1)/NP_-}{1 + \nu/NP_+} = 1 + \mathcal{O}\left(\frac{\nu}{N}\right), \quad (4)$$

so daß die Bedingung (C.7) erfüllt ist.

<sup>3</sup>Für große  $n$  wird  $n!$  durch die *Stirling-Formel* :

$$n! \simeq n^n e^{-n} \sqrt{2\pi n} \quad (1)$$

angenähert (wenn man diese Formel anwendet, so ist der Fehler kleiner als 2%, sobald  $n$  die 5 überschreitet). Daher ist

$$\ln n! \simeq n \ln n - n + \frac{1}{2} \ln(2\pi n) \quad (2)$$

ungefähr in der Größenordnung von  $n$ . Übrigens wird der letzte Term der Formel (2) sehr schnell vernachlässigbar gegenüber dem voranstehenden, wenn  $n$  groß wird, so daß gilt

$$\ln n! \simeq n \ln n - n \quad \text{für } n \gtrsim 50 \quad (3)$$

(für  $n = 50$  ist der Fehler kleiner als 1,5%; er beträgt nicht mehr als  $6 \cdot 10^{-4}$  für  $n = 1000$ ; für  $n = 10^{23}$  muß man  $\ln(2\pi n) \simeq 55$  mit  $\ln n! \simeq 52 \cdot 10^{23}$  vergleichen). Wir unterstreichen noch, daß der Faktor  $\sqrt{2\pi n}$  in der Formel (1) nicht weggelassen werden darf, selbst wenn sein Logarithmus in (2) vernachlässigt werden kann (siehe Fußnote 6).

berücksichtige man, daß<sup>4</sup>

$$\frac{d}{dn} \ln n! \simeq \ln n. \quad (\text{C.9})$$

Man erhält schließlich

$$\frac{\partial}{\partial n} \ln P(N, n) \simeq \ln \frac{N-n}{n} + \ln \frac{P_+}{P_-}. \quad (\text{C.10})$$

Der wahrscheinlichste Wert  $n_m$  ist daher näherungsweise durch

$$\ln \frac{n_m}{N-n_m} \simeq \ln \frac{P_+}{1-P_+}, \quad (\text{C.11})$$

gegeben, das heißt

$$n_m \simeq NP_+. \quad (\text{C.12})$$

Wir werden später sehen (Formel (C.16.b)), daß mit  $n_m$   $\ln P(N, n)$  tatsächlich ein Maximum annimmt.

Wenn man (C.12) mit der Formel (C.2) vergleicht, sieht man, daß

$$n_m \simeq \bar{n} : \quad (\text{C.13})$$

---

<sup>4</sup>Per Definition ist die Ableitung

$$\frac{d}{dn} [\ln n!] = \lim_{\delta n \rightarrow 0} \frac{\ln(n + \delta n)! - \ln n!}{\delta n}. \quad (1)$$

Wenn  $n$  wie eine kontinuierliche Variable behandelt werden kann, dann ist  $\delta n = 1$  genügend klein, um den Wert der Ableitung zu liefern, so daß gilt

$$\frac{d}{dn} \ln n! \simeq \ln [(n+1)!] - \ln(n!) \simeq \ln n. \quad (2)$$

Man beachte übrigens, daß dieses Resultat ebenfalls aus der Formel (3) der Fußnote 3 hervorgeht.

Für große Zahlen fallen der Wert größter Wahrscheinlichkeit und der Mittelwert von  $n$  zusammen<sup>5</sup>.

## 2. Entwicklung von $\ln P$ um den wahrscheinlichsten Wert

Wir entwickeln jetzt  $\ln P(N, n)$  in einer Taylor-Reihe in der Umgebung von  $n = n_m$ . Hier zeigt sich der entscheidende Vorteil des Logarithmus einer Funktion: Die Fakultäten von großen Zahlen sind nicht nur sehr groß, sondern ändern sich auch sehr schnell; *ihr Logarithmus variiert sehr viel langsamer, und seine Entwicklung gestattet eine gute Näherung in einem beträchtlich größeren Wertebereich* der Variablen<sup>6</sup>.

Die Entwicklung, die wir benutzen wollen, schreibt sich

$$\ln P(N, n) = \ln P(N, n_m) + A_1(n - n_m) + \frac{1}{2}A_2(n - n_m)^2 + \frac{1}{3!}A_3(n - n_m)^3 + \dots, \quad (\text{C.14})$$

mit

$$A_p = \left[ \frac{\partial^p}{\partial n^p} \ln P(N, n) \right]_{n=n_m}. \quad (\text{C.15})$$

Die aufeinanderfolgenden Ableitungen von  $\ln P(N, n)$  berechnen sich ausgehend von (C.10); unter Berücksichtigung von (C.12) erhält man

$$A_1 = 0, \quad (\text{C.16.a})$$

$$A_2 = -\frac{1}{NP_+P_-}, \quad (\text{C.16.b})$$

$$A_3 = \frac{1}{N^2} \frac{P_- - P_+}{P_+^2 P_-^2}, \dots \quad (\text{C.16.c})$$

<sup>5</sup>Es ist leicht,  $n_m$  exakt zu berechnen. Tatsächlich ist das Verhältnis

$$\frac{P(N, n)}{P(N, n-1)} = \frac{(N-n+1)P_+}{nP_-} \quad (1)$$

größer oder gleich 1 für

$$n \leq \bar{n} + P_+. \quad (2)$$

Der wahrscheinlichste Wert  $n_m$  ist daher gleich einer *ganzen Zahl, die direkt kleiner als  $\bar{n} + P_+$  ist* (wenn  $\bar{n} + P_+$  ganzzahlig ist, behält  $P(N, n)$  den gleichen Wert für  $n_m + 1 = \bar{n} + P_+$  wie für  $n_m$  und beginnt nur für  $n = n_m + 2$  zu fallen). Man stellt daher fest, daß der Unterschied zwischen dem exakten Wert von  $n_m$  und dem Mittelwert  $\bar{n}$  stets kleiner als 1 ist.

<sup>6</sup>Wir untersuchen unter welchen Bedingungen es vorteilhafter ist, den Logarithmus einer Funktion zu entwickeln.

Also ist in der Entwicklung (C.14) der Term erster Ordnung gleich Null, da  $n_m$  eine Extremwertstelle der Funktion ist. Der Term zweiter Ordnung ist negativ ( $A_2 < 0$ ), was bedeutet, daß der Extremwert ein Maximum ist.

Für die Werte von  $n$ , die die Bedingung (C.6) erfüllen, sind die Terme von höherer Ordnung als 2 vernachlässigbar. Unter der Voraussetzung, daß

$$|n - n_m| \lesssim a \Delta n = a \sqrt{NP_+P_-}, \quad (\text{C.17})$$

gilt, kann man den Term dritter Ordnung durch

$$A_3 |n - n_m|^3 \lesssim a' \frac{1}{N^2} N^{3/2} = \frac{a'}{\sqrt{N}} \ll 1, \quad (\text{C.18})$$

nach oben abschätzen, wobei  $a'$  in der Größenordnung von einigen Einheiten ist. Die gleiche halb-quantitative Überlegung zeigt, daß der Term zweiter Ordnung in der Größenordnung von 1 ist, und für die Werte der aufeinander folgenden Koeffizienten  $A_p$  berechnet man, daß die Terme der Ordnung  $p$  mit wachsendem  $p$  immer kleiner werden.

### 3. Resultat : Gauß-Approximation der Binomialverteilung

Die Entwicklung (C.14) erlaubt,  $\ln P(N, n)$  durch

$$\ln P(N, n) \simeq \ln P(N, n_m) - \frac{1}{2\sigma^2} (n - n_m)^2, \quad (\text{C.19})$$

zu approximieren nach (C.12) und (C.13) mit

$$n_m \simeq \bar{n} = NP_+, \quad (\text{C.20})$$

und nach (C.16.b) und (C.3) mit

$$\sigma^2 = NP_+P_- = (\Delta n)^2. \quad (\text{C.21})$$

Dies liefert (vgl. Fußnote 6) einen Näherungsausdruck für die Wahrscheinlichkeit  $P(N, n)$  :

$$P(N, n) \simeq C_N \exp \left[ -\frac{(n - \bar{n})^2}{2\sigma^2} \right]. \quad (\text{C.22})$$

---

Wir betrachten eine Funktion  $F(x)$  (überall positiv) und ihren Logarithmus  $f(x)$  :

$$f(x) = \ln F(x); \quad (\text{1.a})$$

$$F(x) = \exp f(x). \quad (\text{1.b})$$

Wir nehmen an, daß es in der Entwicklung von  $f(x)$  um einen Wert  $x_0$  der Variablen,

ausreicht, die ersten Terme zu berücksichtigen, um eine gute Näherung von  $f(x)$  zu erhalten. Um uns eine Vorstellung davon zu machen, entwickeln wir bis zur einschließlich 2. Ordnung :

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + \frac{1}{2}(x - x_0)^2 f''(x_0) + \varepsilon, \quad (2)$$

mit der Hypothese

$$|\varepsilon| \ll |f(x)| \quad (3)$$

(in der Praxis muß aus Gründen der Selbstkonsistenz  $\varepsilon$  gegenüber jedem der Terme, die man mitnimmt, vernachlässigbar sein). Indem wir die unvollständige Entwicklung (2) in die Formel (1.b) einsetzen, wird ein Näherungsausdruck für  $F(x)$  erhalten, den wir in der Form

$$F(x) \simeq e^{f(x_0)} \exp \left[ (x - x_0)f'(x_0) + \frac{1}{2}(x - x_0)^2 f''(x_0) \right]. \quad (4)$$

schreiben.

- (i) Diese Approximation ist nur gültig, wenn der vernachlässigte Term

$$e^\varepsilon \simeq 1 \quad (5)$$

erfüllt. Vor dem Einsetzen einer unvollständigen Entwicklung von  $f(x)$  in die Formel (1.b) muß man sich vergewissern, daß diese Entwicklung bis zu einer genügend hohen Ordnung durchgeführt wird, damit

$$|\varepsilon| \ll 1. \quad (6)$$

erfüllt ist. Dies ist sicherlich nicht notwendig, damit  $f(x)$  gut angenähert wird (die Bedingung (3) ist schon ausreichend); indem man die Exponentialfunktion benutzt, führt man die Bedingung (6) ein.

- (ii) Man könnte auch direkt die Funktion  $F(x)$  entwickeln. Die entsprechende Taylor-Formel schreibt sich zu

$$F(x) = e^{f(x_0)} \left\{ 1 + (x - x_0)f'(x_0) + \frac{1}{2}(x - x_0)^2 \left[ f''(x_0) + [f'(x_0)]^2 \right] + \frac{1}{3!}(x - x_0)^3 \left[ f'''(x_0) + 3f'(x_0)f''(x_0) + [f'(x_0)]^3 \right] + \dots \right\}. \quad (7)$$

Aus mathematischer Sicht ist dieser Ausdruck identisch mit dem, den man erhält, wenn man zuerst  $f(x)$  entwickelt : Man erkennt in der Klammer von (7) die ersten Terme der Entwicklung des zweiten Exponenten aus (4). Beide Entwicklungen liefern daher äquivalente Approximationen, wenn alle Terme (der erste ausgenommen) klein gegenüber 1 sind :

$$\begin{aligned} |(x - x_0)f'(x_0)| &\ll 1, \\ \frac{1}{2}(x - x_0)^2 |f''(x_0)| &\ll 1, \dots \end{aligned} \quad (8)$$

Im Wertebereich von  $x$ , die diese Bedingungen erfüllen, kann man ebenso gut die Entwicklung (7) wie die Formel (4) benutzen, um  $F(x)$  zu approximieren.

- (iii) Aber für bestimmte Fälle gibt es einen Bereich, der viel größer als  $(x - x_0)$  ist und in dem die ersten Terme der Entwicklung (2) groß oder in der Größenordnung von 1 und die folgenden klein gegenüber 1 sind. In diesem Fall liefert die Formel (4) noch eine ausgezeichnete Näherung für  $F(x)$ , während die Entwicklung (7) nicht mehr brauchbar ist, denn ihre nachfolgenden Terme sind a priori alle von der gleichen Größenordnung.

Die Konstante

$$C_N = P(N, n_m) \quad (\text{C.23})$$

kann durch Normierung bestimmt werden<sup>7</sup>, was schließlich

$$P(N, n) \simeq \frac{1}{\sqrt{2\pi(\Delta n)^2}} \exp \left[ -(n - \bar{n})^2 / 2(\Delta n)^2 \right] \quad (\text{C.24})$$

ergibt.

*Für große Zahlen stimmt die Binomialverteilung praktisch mit einer Gauß-Verteilung überein, deren Parameter mit dem Mittelwert  $\bar{n}$  und der Standardabweichung  $\Delta n$  der Binomialverteilung zusammenfallen.*

### III. Diskussion der Approximation

Um den Näherungsausdruck (C.24) für die Binomialverteilung zu erhalten, haben wir *a priori* die Bedingung (C.6) aufgestellt, die die zu akzeptierenden Werte von  $n$  begrenzt. Wir untersuchen den Gültigkeitsbereich der Gauß-Approximation näher.

Diese Approximation, die wir durch Abbrechen der Entwicklung (C.14) von  $\ln P(N, n)$  und anschließend durch Bildung des Exponenten dieser abgebrochenen Entwicklung erhalten haben, wird gültig, *wenn die vernachlässigten Terme klein gegenüber 1 und dem Term 2. Ordnung sind, der beibehalten wurde* (vgl. Fußnote 6). Betrachten wir den Term 3. Ordnung (vgl. (C.16.c)). Er ist klein gegenüber 1 (im Absolutwert), wenn gilt

$$|n - \bar{n}| \ll (NP_+P_-)^{2/3}; \quad (\text{C.25})$$

er ist vernachlässigbar gegenüber dem Term 2. Ordnung, wenn gilt

$$|n - \bar{n}| \ll NP_+P_- . \quad (\text{C.26})$$

---

<sup>7</sup>Die Variable  $n$ , die wir als kontinuierlich behandeln, liegt zwischen 0 und  $N$ . Aber die zu integrierende Gauß-Funktion verschwindet völlig, sobald  $|n - \bar{n}|$  einige  $\Delta n$  überschreitet. Ihr Integral von 0 bis  $N$  unterscheidet sich praktisch nicht von einem Integral von  $-\infty$  bis  $+\infty$ , so daß sich die Normierungsbedingung schreiben läßt

$$\int_0^N C_N e^{-(n-\bar{n})^2/2\sigma^2} dn \simeq C_N \int_{-\infty}^{+\infty} e^{-(n-\bar{n})^2/2\sigma^2} dn = 1 .$$

Sie ergibt schließlich

$$C_N = \frac{1}{\sqrt{2\pi\sigma^2}} .$$

Die Approximation wäre ohne Interesse, wenn sie auf Werte von  $n$  beschränkt wäre, die kaum vom Mittelwert  $\bar{n}$  abweichen; deshalb interessieren wir uns nur für Fälle, in denen gilt

$$NP_+P_- \gg 1. \quad (\text{C.27})$$

Nun folgt aber die Beziehung (C.26) aus (C.25); daher legt die erste dieser Bedingungen das Intervall der Werte von  $n$  fest, in der die Gauß-Verteilung die Binomialverteilung korrekt approximiert<sup>8</sup>. Es läßt sich leicht zeigen, daß die Terme von höherer Ordnung als 3 ebenfalls vernachlässigbar sind, *erst recht* in diesem Bereich.

Die Bedingung (C.25) ist nicht die gleiche, die wir ursprünglich in (C.6)

#### TABELLE I

**Numerischer Vergleich zwischen  $(NP_+P_-)^{2/3}$  und  $\Delta n$ , das heißt zwischen den Bedingungen (C.25) und (C.6), die den Gültigkeitsbereich von  $n$  der Gauß-Approximation begrenzen.**

$NP_+P_-$	$(NP_+P_-)^{2/3}$	$(NP_+P_-)^{1/2} = \Delta n$
50	13,6	7,07
100	21,5	10
1 000	100	31,6
10 000	464	100
$10^{10}$	$4,64 \cdot 10^6$	$10^5$
$10^{20}$	$2,15 \cdot 10^{13}$	$10^{10}$

angesetzt hatten. Man stellt fest, daß in der Realität  $(NP_+P_-)^{2/3}$  von gleicher Größenordnung ist wie  $\Delta n = (NP_+P_-)^{1/2}$  solange  $NP_+P_-$  kleiner als 10 000 bleibt; in diesem Bereich kann der Wert von  $a$  in (C.6) die Zahl 2 nicht überschreiten. Erst für sehr viel größere Werte von  $NP_+P_-$  (mindestens in der Größenordnung von  $10^{10}$ ) wird  $(NP_+P_-)^{2/3}$  deutlich größer als  $\Delta n$ , so

<sup>8</sup>Im besonderen Fall,

$$P_+ = P_- = \frac{1}{2}, \quad (1)$$

ist der Term 3. Ordnung identisch Null. Betrachtet man den Term 4. Ordnung, so wird die Bedingung (C.25) ersetzt durch

$$|n - \bar{n}| \ll (NP_+P_-)^{3/4}. \quad (2)$$

Es ist verständlich, daß die Gauß-Näherung in diesem besonderen Fall noch besser wird: Die Binomialverteilung ist dann genauso wie die Gauß-Funktion symmetrisch bezüglich  $n$ .

TABELLE II

Numerischer Vergleich zwischen der Binomialverteilung  $P(N, n)$  und der Gauß-Approximation  $w_G(n)$  für  $N = 100\,000$  und  $P_+ = 0,1$  ( $\bar{n} = 10\,000$  und  $\Delta n \simeq 100$ ).

$n$	$P(N, n)$	$w_G(n)$
$9\,500(\bar{n} - 5\Delta n)$	$3,237 \cdot 10^{-9}$	$3,908 \cdot 10^{-9}$
9600	$5,304 \cdot 10^{-7}$	$5,799 \cdot 10^{-7}$
9700	$2,745 \cdot 10^{-5}$	$2,833 \cdot 10^{-5}$
9800	$4,537 \cdot 10^{-4}$	$4,557 \cdot 10^{-4}$
$9\,900(\bar{n} - \Delta n)$	$2,419 \cdot 10^{-3}$	$2,413 \cdot 10^{-3}$
9920	$2,955 \cdot 10^{-3}$	$2,947 \cdot 10^{-3}$
9940	$3,451 \cdot 10^{-3}$	$3,443 \cdot 10^{-3}$
9960	$3,854 \cdot 10^{-3}$	$3,847 \cdot 10^{-3}$
9980	$4,116 \cdot 10^{-3}$	$4,113 \cdot 10^{-3}$
$10\,000(\bar{n})$	$4,205 \cdot 10^{-3}$	$4,205 \cdot 10^{-3}$
10020	$4,109 \cdot 10^{-3}$	$4,113 \cdot 10^{-3}$
10040	$3,841 \cdot 10^{-3}$	$3,847 \cdot 10^{-3}$
10060	$3,435 \cdot 10^{-3}$	$3,443 \cdot 10^{-3}$
10080	$2,939 \cdot 10^{-3}$	$2,947 \cdot 10^{-3}$
$10\,100(\bar{n} + \Delta n)$	$2,406 \cdot 10^{-3}$	$2,413 \cdot 10^{-3}$
10200	$4,576 \cdot 10^{-4}$	$4,557 \cdot 10^{-4}$
10300	$2,922 \cdot 10^{-5}$	$2,833 \cdot 10^{-5}$
10400	$6,319 \cdot 10^{-7}$	$5,799 \cdot 10^{-7}$
$10\,500(\bar{n} + 5\Delta n)$	$4,674 \cdot 10^{-9}$	$3,908 \cdot 10^{-9}$

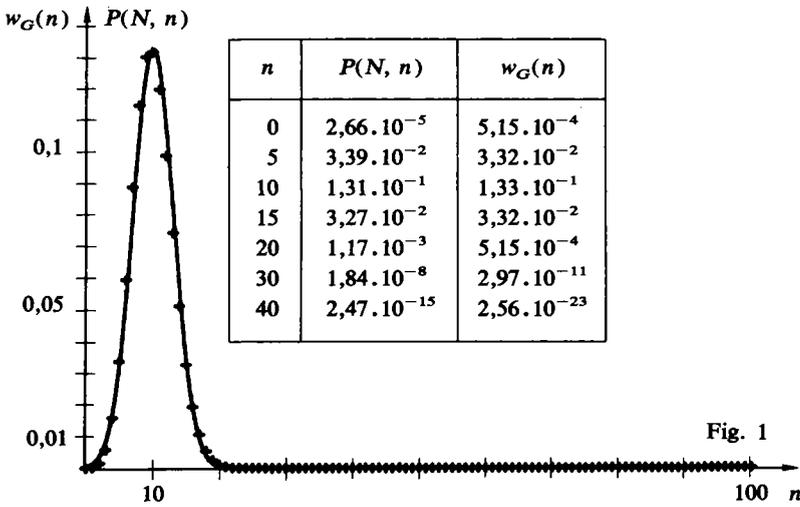


Fig. 1

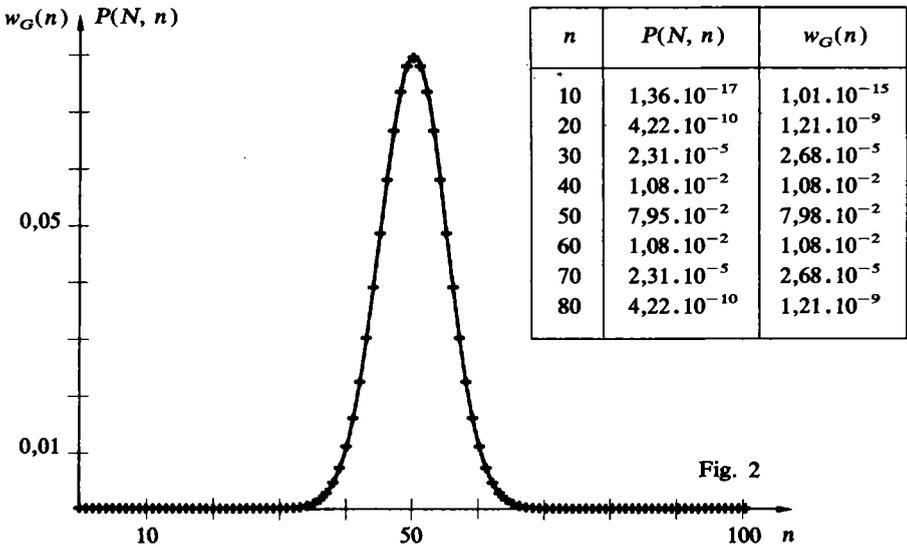


Fig. 2

ABBILDUNGEN 1 und 2

Vergleich zwischen der Binomialverteilung  $P(N, n)$  (sternförmige Punkte) und ihrer Gauß-Approximation (ausgezogene Kurve) für  $N = 100$ . Einige typische numerische Werte werden umrahmt angegeben. Die Abbildung 1 entspricht den Werten  $P_+ = 0,1$  mit  $\bar{n} = 10$  und  $\Delta n = 3$  ( $NP_+P_-$  beträgt nur 9); für die Abbildung 2 sind  $P_+ = 1/2$ ,  $\bar{n} = 50$  und  $\Delta n = 5$ ,  $NP_+P_-$  beträgt nur 25.

daß die durch die Bedingung (C.25) festgelegten Grenzen den durch (C.6) definierten Bereich bedeutend überschreiten. Sicherlich hängen die Grenzen des Gültigkeitsbereiches der Gauß-Approximation von der Genauigkeit ab, die man von dieser Approximation fordert.

Man beachte außerdem, daß die Wahrscheinlichkeit  $P(N, n)$  sehr klein wird (im Absolutwert und im relativen Wert) im Vergleich zu  $P(N, \bar{n})$ , sobald  $n$  den durch (C.6) und (C.25) definierten Bereich verläßt: Die Tabelle II zeigt zum Beispiel, daß für  $N = 100\,000$  und  $P_+ = 0,1$  die Binomialwahrscheinlichkeit auf 7 Tausendstel ihres Maximalwertes fällt, wenn  $|n - \bar{n}| 300$  beträgt, das heißt ungefähr das 3-fache von  $\Delta n$ .

Halten wir also fest: *In der Praxis ist die Gauß-Approximation für  $NP_+P_-$  groß gegenüber 1 und in einem Intervall von einigen  $\Delta n$  um  $\bar{n}$  herum gültig, das heißt für solche Werte von  $n$ , die eine nennenswerte Wahrscheinlichkeit haben*<sup>9</sup>.

Um die Gültigkeit der Gauß-Approximation zu überprüfen, kann man sie numerisch für einige typische Fälle mit den exakten Werten der Binomialwahrscheinlichkeit vergleichen. In Tabelle II ist  $N$  relativ groß, jedoch kleiner als die Avogadro-Konstante; man stellt vor allem fest, daß die exakten und angenäherten Werte in einem Intervall von mehreren  $\Delta n$  sehr dicht beieinander liegen, was einer Variation dieser Wahrscheinlichkeiten von mehreren Größenordnungen entspricht: Folglich überschreitet die Differenz zwischen exaktem und angenähertem Wert keine 20% in den Bereichen, wo die Wahrscheinlichkeit um einen Faktor  $10^6$  im Vergleich zu ihrem Maximalwert gesunken ist. Wir haben in den Abbildungen 1 und 2 einen sehr viel kleineren Wert von ( $N = 100$ ) gewählt, damit die Unterschiede zwischen  $P(N, n)$  und  $w_G(n)$  wahrgenommen werden können:  $NP_+P_-$  erreicht knapp 9 in der Abbildung 1, und knapp 25 in der Abbildung 2. Die Gauß-Approximation ist in diesen beiden Fällen offensichtlich noch vernünftig; im zweiten Fall erscheinen noch nicht einmal die Unterschiede in einer Abbildung mit normalem Maßstab. Dieser numerische und graphische Vergleich zeigt die *bemerkenswerte Genauigkeit* der Gauß-Approximation für die Binomialverteilung.

---

<sup>9</sup>Wenn man sich für sehr wenig wahrscheinliche Werte von  $n$  interessiert, die sehr verschieden von  $\bar{n}$  sind (nahe an 0 oder an  $N$  zum Beispiel), so ist die Gauß-Approximation kein Hilfsmittel mehr; man muß auf den Anfangsausdruck der Binomialverteilung zurückgreifen.

# Zufallsbewegung und Brownsche Bewegung

---

I. Zufallsbewegung in einer Dimension

1. Schritte von konstanter Länge
2. Schritte von variabler Länge

II. Graphische Darstellung der Zufallsbewegung in einer Dimension

1. Die graphische Darstellung
2. Reflexionstheorem
3. Anwendung

III. Zufallsbewegung in drei Dimensionen und Brownsche Bewegung

1. Isotrope Zufallsbewegung in drei Dimensionen
2. Ein einfaches Modell für die Brownsche Bewegung
3. Experimentelle Verifizierung

---

Die Brownsche Bewegung wurde 1827 von dem englischen Botaniker R. Brown entdeckt. Er beobachtete unter dem Mikroskop eine Suspension von Pollenkörnern in Wasser. Brown stellte fest, daß die Körner von einer ständigen ungeordneten Bewegung angeregt wurden. « Sie kommen und gehen, drehen sich, steigen, fallen, steigen wieder, ohne zur Ruhe zu gelangen » (Jean Perrin). Die Hypothese, daß es sich hierbei um « winzige rudimentäre Tierchen » handelt, wurde sehr schnell fallen gelassen, und das Phänomen blieb unerklärt.

1905 äußerte Einstein<sup>1</sup> die Idee, daß diese Bewegung durch die Stöße der Moleküle der Flüssigkeit auf die Körner in der Suspension hervorgerufen wird. Er nahm als Hypothese an, daß die Richtungsänderungen infolge dieser Stöße zufällig sind, und schlug eine Theorie der Brownschen Bewegung vor, deren experimentelle Konsequenzen 1908 von Jean Perrin verifiziert wurden. Die Existenz der Brownschen Bewegung war ein entscheidendes Argument für die Physiker, die ihre Kollegen von der atomaren Natur der Materie zu überzeugen versuchten. Jean Perrin schrieb : « ... es wird daher schwierig, die objektive Realität der Moleküle zu leugnen ... Die molekulare Bewegung wird uns sichtbar gemacht. Die Brownsche Bewegung ist davon ein getreues Abbild, oder besser, sie ist bereits eine molekulare Bewegung ».

---

<sup>1</sup> M. Smoluchovski hat unabhängig davon 1906 eine ähnliche, wenn auch weniger ausführliche Theorie veröffentlicht.

Die von Einstein vorgeschlagene Theorie besteht darin, die Untersuchung der Brownschen Bewegung auf eine Zufallsbewegung zurückzuführen, was wir heute auch bildhaft als « Zufallswanderung »\* bezeichnen. Dieses Problem, für das es noch andere Beispiele gibt, hat zu wichtigen mathematischen Entwicklungen Anlaß gegeben : Wir begegnen dem Prozeß der Zufallsbewegung in einer Dimension, zwei, drei Dimensionen; die Bewegung kann in einem unendlichen oder begrenzten Raum stattfinden; die « Schritte » können von konstanter oder variabler Länge sein, alle Richtungen können gleichwahrscheinlich oder einige von ihnen können bevorzugt sein; die Situationen sind daher mehr oder weniger komplex. Wir beschränken uns hier auf einige Beispiele<sup>2</sup>.

## I. Zufallsbewegung in einer Dimension

### 1. Schritte von konstanter Länge

Wir betrachten zunächst den einfachsten Fall. Ein Wanderer bewegt sich mit Schritten von konstanter Länge  $\ell$  auf einer unendlichen Straße, die wir durch eine orientierte Gerade schematisch darstellen. Ein Schritt kann in positiver oder negativer Richtung mit den Wahrscheinlichkeiten  $P_+$  bzw.  $P_-$  ausgerichtet sein. Die Schritte sind statistisch unabhängig : Die Richtung eines Schrittes ist unabhängig von der des vorhergehenden; man sagt, daß der Wanderer « die Erinnerung » an seine vorherige Bewegung « verloren » hat, oder daß es sich um einen Markow-Prozeß handelt. Was ist die Wahrscheinlichkeit  $\mathcal{P}(N, s)$  dafür, daß sich der Wanderer nach  $N$  Schritten im Punkt  $A$  der Abszisse  $s\ell$  ( $s$  ganzzahlig,  $|s| \leq N$ ) in bezug auf den Startpunkt  $D$  befindet?

Dieses Problem ist mit dem von Kopf oder Zahl identisch : Vor jedem Schritt könnte die Wahl der positiven oder negativen Richtung durch das Ergebnis (Kopf oder Zahl) des Werfens eines Geldstückes bestimmt werden. Die Verteilung  $\mathcal{P}(N, s)$  wird daher binomial sein.

---

\*Anmerkung des Übersetzers : im Französischen « marche au hasard », im Englischen « random walk ».

<sup>2</sup>Die Methoden und Resultate der Zufallsbewegung werden in verschiedenen Gebieten der Physik benötigt, wenn man eine physikalische Größe als Zufallsvariable zu betrachten hat, die sukzessive voneinander unabhängige Werte annimmt, zumindest in erster Näherung : Diffusion von Teilchen in einer Umgebung, die aus anderen Teilchen besteht, Sedimentation von Staub in einer Flüssigkeit, oder Ermittlung des Gravitationsfeldes, das von den Sternen verursacht wird (hier wird die Verteilung der Sterne als zufällig betrachtet - siehe Literaturstelle 1). Aber die Zufallsbewegung kann auch bei Problemen eingesetzt werden, bei denen keine Größe von vornherein zufällig ist, wie zum Beispiel in der Potentialtheorie (Literaturstelle 2).

Wir bezeichnen als *Strecke* die geordnete Menge der  $N$  aufeinanderfolgenden Positionen des Wanderers. Jede Strecke mit  $n_+$  Schritten in positiver Richtung (und  $n_- = (N - n_+)$  in negativer Richtung) führt den Wanderer zum Abszissenpunkt

$$n_+ \ell - (N - n_+) \ell = (2n_+ - N) \ell. \quad (\text{D.1})$$

Es gibt daher zahlreiche verschieden Strecken, die nach  $N$  Schritten von  $D$  aus zum Abszissenpunkt  $s\ell$  führen : Es sind alle die, die eine Anzahl von Schritten  $n_+$  in positive Richtung enthalten, so daß gilt

$$(2n_+ - N) \ell = s \ell, \quad (\text{D.2.a})$$

das heißt<sup>3</sup>

$$n_+ = \frac{N + s}{2}. \quad (\text{D.2.b})$$

Diese verschiedenen Strecken unterscheiden sich durch den Platz der  $n_+$  positiven Schritte in der Sequenz der  $N$  Schritte voneinander; ihre Anzahl ist gleich  $C_N^{n_+}$ . \* Sie sind alle gleichwahrscheinlich, die Wahrscheinlichkeit eines jeden von ihnen ist

$$(P_+)^{n_+} \cdot (P_-)^{n_-} = (P_+)^{(N+s)/2} \cdot (P_-)^{(N-s)/2}. \quad (\text{D.3})$$

Man erhält schließlich

$$\begin{aligned} \mathcal{P}(N, s) &= C_N^{(N+s)/2} \cdot (P_+)^{(N+s)/2} \cdot (P_-)^{(N-s)/2} \\ &= \frac{N!}{\left(\frac{N+s}{2}\right)! \left(\frac{N-s}{2}\right)!} (P_+)^{(N+s)/2} \cdot (P_-)^{(N-s)/2}. \end{aligned} \quad (\text{D.4})$$

Indem man die Ergebnisse (44) und (47) von Anhang I sowie die Beziehung (D.2.b) zwischen  $s$  und  $n_+$  verwendet, erhält man leicht den Mittelwert und die Standardabweichung von  $s$  :

$$\bar{s} = N(P_+ - P_-), \quad (\text{D.5})$$

$$\Delta s = (\overline{s^2} - \bar{s}^2)^{1/2} = \sqrt{4NP_+P_-}. \quad (\text{D.6})$$

Wir stellen uns vor, daß eine große Anzahl von Wanderern gemeinsam vom gleichen Startpunkt aufbrechen und betrachten den besonderen Fall, daß

<sup>3</sup>Die Gleichung (D.2.b) hat nur dann eine Lösung, wenn  $N$  und  $s$  beide gerade oder beide ungerade sind (denn die Zahl  $n_+$  muß ganzzahlig sein). Im entgegengesetzten Fall gilt  $\mathcal{P}(N, s) = 0$ .

\* Anmerkung des Übersetzers : Das französische Symbol  $C_N^{n_+}$  liest sich «  $CNn_+$  » und entspricht dem Binomialkoeffizienten  $\binom{N}{n_+}$ .

$P_+ = P_- = 1/2$  gilt. Nach  $N$  Schritten verteilen sich die Wanderer symmetrisch zu ihrem Ausgangspunkt entlang der Straße, was erklärt, daß der Mittelwert ihrer Abszisse Null ist. Aber jeder Wanderer hat die Tendenz, sich weiter vom Ursprung zu entfernen : Die Standardabweichung ist nicht Null; sie wächst mit der Quadratwurzel aus der Anzahl  $N$  von gemachten Schritten.

## 2. Schritte von variabler Länge

Wir betrachten jetzt den Fall, in dem die aufeinanderfolgenden, immer noch unabhängigen Schritte keine konstante Länge mehr haben : Es sei  $w(x)dx$  die Wahrscheinlichkeit dafür, daß ein gegebener Schritt die Abszisse des Wanderers um eine Größe, die zwischen  $x$  und  $x + dx$  liegt, vergrößert. Was ist die Wahrscheinlichkeitsdichte  $W_N(X)$  dafür, daß der Wanderer sich nach  $N$  Schritten an der Stelle  $X$  auf der Abszisse befindet?

Wenn  $N$  genügend groß ist, kann man den zentralen Grenzwertsatz benutzen<sup>4</sup> :

$$W_N(X) = \frac{1}{\sqrt{2\pi N\sigma^2}} \exp \left[ -\frac{(X - N\bar{x})^2}{2N\sigma^2} \right], \quad (\text{D.7})$$

wobei  $\bar{x}$  und  $\sigma$  der Mittelwert bzw. die Standardabweichung der Zufallsvariablen  $x$  sind.

---

<sup>4</sup>Die Wahrscheinlichkeitsverteilung  $w(X)$  der Summe

$$X = \sum_{j=1}^N x_j \quad (1)$$

von  $N$  unabhängigen Zufallsvariablen  $x_j$  ist für  $N$  sehr groß gegenüber 1 eine Gauß-Verteilung um

$$\bar{X} = \sum_{j=1}^N \bar{x}_j \quad (2)$$

zentriert und der Breite  $\Delta X$ , so daß gilt

$$(\Delta X)^2 = \sum_{j=1}^N (\Delta x_j)^2 \quad (3)$$

(siehe Übung I.12).

## II. Graphische Darstellung der Zufallsbewegung in einer Dimension

Im Fall *konstanter Schrittlänge* erleichtert eine graphische Darstellung die Untersuchung und Lösung des Problems der Zufallsbewegung. Wir werden sie jetzt beschreiben und zur Aufstellung eines wichtigen Theorems benutzen; anschließend zeigen wir eine Anwendung.

### 1. Die graphische Darstellung

Wir stellen die Strecke des Wanderers in einem Diagramm dar, das seine Position als Funktion der Zeit angibt : Wir tragen auf der Abszisse die Anzahl  $N$  der seit einem Anfangszeitpunkt ausgeführten Schritte auf und auf der Ordinate das ganzzahlige  $s$ , das den Wanderer bezüglich seines Startpunktes markiert. Die durchlaufene Strecke wird nun durch eine unterbrochene Linie dargestellt, die die mit den aufeinander folgenden Positionen verknüpften Punkte verbindet, und die wir einen  $\ll$  Weg  $\gg$  nennen. Wenn ein Weg durch einen Punkt  $M$  mit den Koordinaten  $(N, s)$  verläuft, dann bedeutet das, daß sich der Wanderer nach  $N$  Schritten im Abszissenpunkt  $s\ell$  befindet (siehe Figur 1). Mit jeder Strecke ist nur ein einziger Weg verknüpft; dies gilt auch umgekehrt.

### 2. Reflexionstheorem

Wir betrachten auf der graphischen Darstellung einer Zufallsbewegung eine horizontale Gerade  $\Delta$  und zwei Punkte  $A$  und  $B$ , die auf der gleichen Seite dieser Geraden liegen (Abb. 2). Einige Wege, die von  $A$  nach  $B$  verlaufen, kreuzen oder berühren  $\Delta$ , andere erreichen sie nicht.

Wir werden das folgende Theorem beweisen : *die Anzahl der Wege, die von  $A$  nach  $B$  verlaufen und die Gerade  $\Delta$  schneiden oder berühren, ist gleich der gesamten Anzahl von Wegen, die vom Punkt  $A'$ , der symmetrisch zu  $A$  bezüglich  $\Delta$  liegt, zum Punkt  $B$  verlaufen.*

Es sei ein Weg gegeben, der von  $A$  nach  $B$  verläuft und der  $\Delta$  schneidet oder berührt,  $C$  sei der Schnitt- bzw. Berührungspunkt. Wir ordnen ihm einen Weg zu, der  $A'$  mit  $B$  verbindet und auf folgende Weise konstruiert wird : Von  $C$  bis  $B$  fällt der neue Weg mit dem alten zusammen; für den vorhergehenden Teil nach  $C$  sind beide Wege in bezug auf  $\Delta$  symmetrisch. Umgekehrt schneidet jeder Weg, der von  $A'$  nach  $B$  verläuft, notwendigerweise  $\Delta$ , und wir können ihm auf die gleiche Weise einen Weg zuordnen, der von  $A$  nach  $B$  unter Berührung von  $\Delta$  verläuft. Das beweist das hier oben

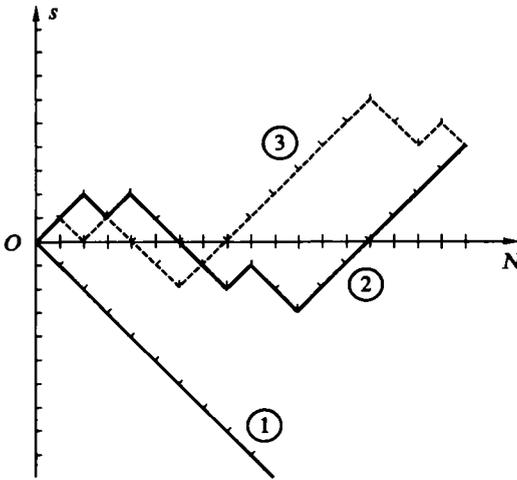


ABBILDUNG 1

**Graphische Darstellung der Zufallsbewegung :** Auf der Abszisse ist die Anzahl der seit dem Start ausgeführten Schritte dargestellt; auf der Ordinate wird der Abszissenwert des Wanderers angegeben ( $\ell$  wird als Einheit genommen). Weg 1 : Alle Schritte erfolgen in negativer Richtung. Wege 2 und 3 : Zwei verschiedene Strecken, die nach 18 Schritten vom Ursprung  $O$  aus zu einem Punkt mit dem Abszissenwert  $4\ell$  führen.

ausgesprochene Ergebnis, das man als Theorem der Reflexion bezeichnet<sup>5</sup>.

### 3. Anwendung

Wir suchen die Wahrscheinlichkeit  $\tilde{P}(N, s)$  dafür, daß ein Wanderer nach  $N$  Schritten konstanter Länge  $\ell$  mit gleichen Wahrscheinlichkeiten in positiver oder negativer Richtung im Abszissenpunkt  $s\ell$  zum ersten Mal vorbeikommt.

Mit Hilfe der graphischen Darstellung suchen wir die Anzahl von Wegen, die von dem Punkt  $O(0, 0)$  zu dem Punkt  $A(N, s)$  verlaufen, ohne vorher die horizontale Gerade  $\Delta$  mit dem Abszissenwert  $s$  zu berühren (Abb. 3). Alle diese Wege verlaufen gezwungenermaßen durch den Punkt  $B(N - 1, s - 1)$ . Insgesamt gibt es  $C_{N-1}^{(N+s-2)/2}$  Wege von  $O$  nach  $B$ ; unter diesen ist die Anzahl derer, die  $\Delta$  schneiden oder berühren, nach dem Reflexionstheorem gleich

<sup>5</sup>Achtung : Die Wahrscheinlichkeiten, von  $A$  nach  $B$  unter Berührung von  $\Delta$  und von  $A'$  nach  $B$  zu gehen, sind nicht gleich, denn ein Weg, der von  $A'$  nach  $B$  verläuft, enthält nicht die gleiche Anzahl von Schritten in positiver Richtung wie ein Weg, der von  $A$  nach  $B$  verläuft (diese Wahrscheinlichkeiten sind nur gleich für  $P_+ = P_- = 1/2$ ).

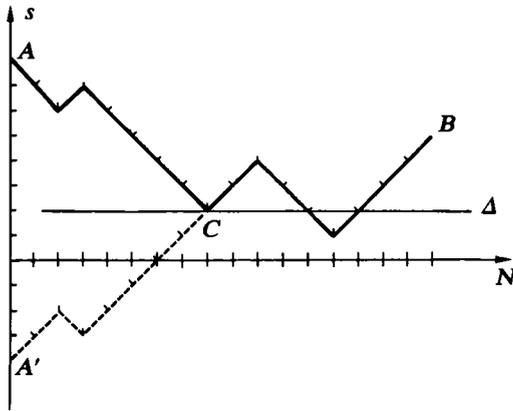


ABBILDUNG 2

Zwei miteinander verknüpfte Wege zur Demonstration des Reflexionstheorem.

der Anzahl von Wegen, die von  $O'(0, 2s)$  nach  $B$  verlaufen, somit<sup>6</sup>  $C_{N-1}^{(N+s)/2}$ ; durch Subtraktion erhalten wir die uns interessierende Anzahl von Wegen. Da alle Wege  $N$  Schritte enthalten, ist die Wahrscheinlichkeit für jeden<sup>7</sup>  $(1/2)^N$ , so daß die gesuchte Wahrscheinlichkeit  $\tilde{P}(N, s)$

$$\begin{aligned}\tilde{P}(N, s) &= \left(\frac{1}{2}\right)^N \left\{ C_{N-1}^{(N+s-2)/2} - C_{N-1}^{(N+s)/2} \right\} \\ &= \left(\frac{1}{2}\right)^N \frac{s}{N} C_N^{(N+s)/2}\end{aligned}\quad (\text{D.8})$$

beträgt. Man kann zeigen<sup>8</sup>, daß die Summe über alle Werte von  $N$  der Wahrscheinlichkeit  $\tilde{P}(N, s)$  gleich 1 ist :

$$\sum_{N=s}^{\infty} \left(\frac{1}{2}\right)^N \frac{s}{N} C_N^{(N+s)/2} = 1 \quad \forall s, \quad (\text{D.9})$$

<sup>6</sup>Die Anzahl der Wege von  $O'(0, 2s)$  nach  $B(N-1, s-1)$  ist tatsächlich gleich der Anzahl der Wege von  $O(0, 0)$  nach  $B'(N-1, s-1-2s)$  oder auch von  $O(0, 0)$  nach  $B''(N-1, s+1)$  (siehe Abbildung 3).

<sup>7</sup>Wenn die Wahrscheinlichkeiten der positiven und negativen Schritte nicht gleich sind sondern  $P_+$  und  $P_-$  betragen, dann hat jeder dieser Wege eine Wahrscheinlichkeit von  $(P_+)^{(N+s)/2} (P_-)^{(N-s)/2}$ .

<sup>8</sup>Siehe Literaturstelle 4, S. 350.

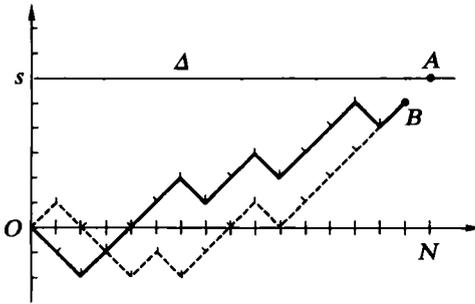


ABBILDUNG 3

**Erste Passage in  $s$  nach  $N$  Schritten : Alle günstigen Wege gehen durch  $B$ , und man kann sie nur auf eine einzige Art verlängern, um  $A$  zu erreichen.**

das bedeutet, daß der Wanderer zuletzt stets an einem beliebigen Punkt der Geraden vorbeikommt<sup>9</sup>. Dagegen ist der Mittelwert der Anzahl von erforderlichen Schritten, um den betrachteten Punkt zu erreichen, nicht bestimmt (die entsprechende Reihe konvergiert nicht).

### III. Zufallsbewegung in drei Dimensionen und Brownsche Bewegung

#### 1. Isotrope Zufallsbewegung in drei Dimensionen

Ein Körper bewege sich im dreidimensionalen Raum durch eine Folge von « Schritten »; diese sind statistisch unabhängig und für alle Richtungen gleichwahrscheinlich. Die Länge  $\ell$  der Schritte ist variabel und durch eine Wahrscheinlichkeitsdichte  $w(\ell)$  charakterisiert, die von der Richtung des Schrittes unabhängig ist. Wir wollen die Position des Körpers untersuchen, nachdem er  $N$  Schritte ausgeführt hat; wir nehmen  $N$  als groß an.

Die Wahrscheinlichkeit dafür, daß der erste Schritt den Körper von seinem Startpunkt (den wir als Ursprung nehmen) zu einem Punkt führt, der

<sup>9</sup>Man findet für den Fall, daß die Wege in positiver oder negativer Richtung nicht gleichwahrscheinlich sind, daß die Wahrscheinlichkeit, an einem positiven Punkt der Abszisse vorbeizukommen, 1 beträgt, solange  $P_+ > P_-$  gilt und die Anzahl  $N$  von Schritten nicht begrenzt ist. Man findet ein analoges Resultat für die Zufallsbewegung in 2 Dimensionen : Wenn alle Richtungen gleichwahrscheinlich sind, erreicht der Wanderer mit einer Wahrscheinlichkeit gleich 1 jeden beliebigen Punkt der Ebene. Anders ausgedrückt : « Alle Wege führen nach Rom ». Dies gilt nicht in 3 Dimensionen.

innerhalb eines Volumenelementes um den Punkt  $M$  mit den sphärischen Koordinaten  $(r, \theta, \varphi)$  liegt, beträgt

$$dP(r, \theta, \varphi) = w(r) dr \frac{d\Omega}{4\pi}, \quad (\text{D.10})$$

wobei

$$d\Omega = \sin \theta d\theta d\varphi \quad (\text{D.11})$$

das Flächenelement in Richtung  $(\theta, \varphi)$  ist. Tatsächlich ist  $w(r)dr$  die Wahrscheinlichkeit dafür, daß die Schrittlänge  $\ell$  gleich  $r$  ist mit  $dr$ , und  $d\Omega/4\pi$  ist die Wahrscheinlichkeit dafür, daß seine Richtung bis auf  $d\Omega$  gleich  $(\theta, \varphi)$  ist.

Der nächste Schritt ist sicherlich durch eine analoge Verteilung charakterisiert, seine Länge und seine Richtung sind von den vorhergehenden Schritten unabhängig; aber sein Ausgangspunkt ist der Endpunkt des ersten Schrittes. Um diese Schwierigkeit zu lösen, genügt eine *Projektion auf die drei kartesischen Koordinatenachsen  $Ox, Oy, Oz$* . Die Position des Körpers nach  $N$  Schritten ist dann durch die Koordinaten  $X_N, Y_N, Z_N$  charakterisiert, und zwar so, daß gilt

$$X_N = \sum_{i=1}^N x_i, \quad (\text{D.12})$$

wobei  $x_i$  die Projektion auf  $Ox$  des  $i$ -ten Schrittes ist (analoge Formeln liefern  $Y_N$  und  $Z_N$ ).

Man leitet aus (D.12) leicht den Mittelwert und die Standardabweichung der Position nach  $N$  Schritten ab; es genügt dafür den Mittelwert und die Standardabweichung der Projektion eines Schrittes auf jede der drei Achsen zu kennen. Das Problem ist somit isotrop, und man erhält notwendigerweise

$$\bar{x} = \bar{y} = \bar{z} = 0, \quad (\text{D.13})$$

$$\overline{x^2} = \overline{y^2} = \overline{z^2} = \frac{1}{3} \overline{\ell^2}, \quad (\text{D.14})$$

mit

$$\overline{\ell^2} := \int_0^{+\infty} \ell^2 w(\ell) d\ell. \quad (\text{D.15})$$

Man erhält somit<sup>10</sup>

$$\overline{X_N} = \overline{Y_N} = \overline{Z_N} = 0, \quad (\text{D.16})$$

<sup>10</sup>Es ist bekannt, daß der Mittelwert einer Summe gleich der Summe der Mittelwerte ist :

$$\overline{X_N} = \sum_{i=1}^N \overline{x_i} = N\overline{x} = 0, \quad (1)$$

$$(\Delta X_N)^2 = (\Delta Y_N)^2 = (\Delta Z_N)^2 = \frac{N}{3} \ell^2. \quad (\text{D.17})$$

Auch da sind noch die Mittelwerte der drei Koordinaten Null (da alle Richtungen gleichwahrscheinlich sind) und die Standardabweichungen variieren mit  $\sqrt{N}$ ; insbesondere ist der Mittelwert des Quadrates des Abstandes  $R_N^2 = X_N^2 + Y_N^2 + Z_N^2$  zum Ursprung einfach gleich  $N\ell^2$ .

Wenn  $N$  genügend groß ist, können wir den zentralen Grenzwertsatz benutzen (Fußnote 4) : Die Verteilung der Wahrscheinlichkeit  $W_N(X)$  der Abszisse  $X$  nach  $N$  Schritten ist eine Gauß-Verteilung :

$$W_N(X) = \sqrt{\frac{3}{2\pi N\ell^2}} \exp\left[-\frac{3X^2}{2N\ell^2}\right], \quad N \gg 1. \quad (\text{D.18})$$

Die Verteilungen auf den anderen Achsen sind identisch.

## 2. Ein einfaches Modell für die Brownsche Bewegung

Die Zufallsbewegung kann dazu dienen, die Brownsche Bewegung eines Teilchens in einem Gas oder einer Flüssigkeit darzustellen. Wir nehmen an, daß das Teilchen sich frei bewegt und nur mit den Molekülen des Gases oder der Flüssigkeit bei einem momentanen Zusammenstoß in Wechselwirkung tritt. Durch den Stoß wird es zufällig in eine beliebige Richtung mit einer in alle Richtungen gleichförmig verteilten Wahrscheinlichkeit gelenkt : Seine Trajektorie ist somit eine unterbrochene Linie, deren Einzelsegment einer Bahn zwischen zwei Zusammenstößen entspricht<sup>11</sup>.

sowie

$$\bar{Y}_N = \bar{Z}_N = 0. \quad (2)$$

Die Berechnung der Standardabweichung wird daher auf die Berechnung von  $\bar{X}_N^2$  rückgeführt :

$$(\Delta X_N)^2 = \bar{X}_N^2 - (\bar{X}_N)^2 = \bar{X}_N^2. \quad (3)$$

Um das zu berechnen, schreiben wir

$$X_N^2 = \left(\sum_{i=1}^N x_i\right) \left(\sum_{j=1}^N x_j\right) = \sum_{i,j=1}^N x_i x_j = \sum_{i=1}^N x_i^2 + \sum_{\substack{i,j=1 \\ j \neq i}}^N x_i x_j. \quad (4)$$

Für  $i \neq j$  sind die Größen  $x_i$  und  $x_j$  statistisch unabhängig; der Mittelwert ihres Produktes ist daher gleich dem Produkt ihrer Mittelwerte (vgl. Anhang I, §I.2.b), die hier Null sind. Man erhält schließlich

$$(\Delta X_N)^2 = \bar{X}_N^2 = \sum_{i=1}^N \bar{x}_i^2 = \frac{N}{3} \ell^2. \quad (5)$$

<sup>11</sup>Die Theorie der Brownschen Bewegung ist viel komplexer als die Skizze zeigt, die wir hier angeben. Siehe Literaturstelle 3, S. 560 und Übung IV.5 bis IV.7.

Am Ende der Zeit  $t$ , die groß gegenüber der mittleren Zeit  $\tau$  ist, die zwei aufeinanderfolgende Zusammenstöße voneinander trennt, hat das Teilchen eine große Anzahl  $N$  von Zusammenstößen erlitten, die offensichtlich proportional zu  $t$  sind ( $N = t/\tau$ ). Der Mittelwert der Abszisse des Teilchens in eine feste Richtung ist Null mit einer Standardabweichung, die durch (D.17) gegeben ist und die man in der Form

$$(\Delta X)^2 = 2Dt \quad (\text{D.19})$$

schreiben kann, wobei  $D$  der *Diffusionskoeffizient*<sup>12</sup> der Teilchen in der Suspension, entweder in der Flüssigkeit oder in dem Gas ist. In unserem einfachen Modell gilt

$$D = \frac{1}{6} \frac{\overline{\ell^2}}{\tau}. \quad (\text{D.20})$$

### 3. Experimentelle Verifizierung

Um die Brownsche Bewegung zu beobachten, muß man Teilchen untersuchen, die genügend groß sind, um unter dem Mikroskop sichtbar zu sein, damit die Zusammenstöße der Moleküle auf sie eine wahrnehmbare Bewegung übertragen.

Um die Vorhersagen der Theorie der Brownschen Bewegung zu verifizieren, beobachtete Jean Perrin unter dem Mikroskop die Bewegung von der Größe nach sortierten Kügelchen<sup>13</sup>, die in einer Flüssigkeit suspendiert waren. Er verfolgte ein Kügelchen und zeichnete davon die Position alle dreißig

---

<sup>12</sup>Der Diffusionskoeffizient mißt *die Antwort des Systems auf den Konzentrationsgradienten*. In Abwesenheit eines äußeren Feldes, entspricht das Gleichgewicht einer homogenen Verteilung der Teilchen in der Suspension. Wenn man eine Inhomogenität in dieser Verteilung erzeugt, tritt ein Teilchenfluß auf, der versucht, den Gleichgewichtszustand wieder herzustellen. Wir bezeichnen mit  $c$  die Konzentration der betrachteten Teilchen (Anzahl der Teilchen durch Volumen). Wenn man nun ein Konzentrationsgefälle (Gradient)  $\partial c/\partial z$  in Richtung  $Oz$  aufrecht erhält, so stellt sich ein Teilchenfluß  $J_z$  entlang von  $Oz$  ein (Anzahl der Teilchen, die pro Zeiteinheit eine zu  $Oz$  senkrechte Flächeneinheit durchdringen). Im Grenzfall eines ziemlich kleinen Gradienten beobachtet man eine « lineare Antwort » , das heißt, daß  $J_z$  zu  $\partial c/\partial z$  proportional ist. Der Diffusionskoeffizient ist dann definiert durch

$$J_z := -D \frac{\partial c}{\partial z}$$

(siehe Literaturstelle 3, S. 483).

<sup>13</sup>Die Experimente von Jean Perrin wurden mit Kügelchen von Gummigutt realisiert (Radien von  $0,2 \mu\text{m}$  bis  $0,5 \mu\text{m}$  und Massen von  $4,8 \cdot 10^{-14} \text{ g}$  bis  $6,00 \cdot 10^{-13} \text{ g}$ ), sowie mit Kügelchen von Kitt (Radien von  $0,5 \mu\text{m}$  bis  $5,5 \mu\text{m}$  und Massen von  $6,5 \cdot 10^{-13} \text{ g}$  bis  $7,5 \cdot 10^{-10} \text{ g}$ ).

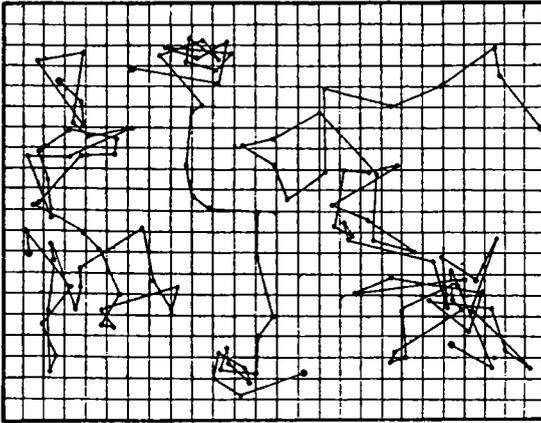


ABBILDUNG 4

Aufeinanderfolgende Positionen von 3 in einer Flüssigkeit suspendierten Kügelchen. Die Eintragung der Punkte wurde von J. Perrin alle 30 Sekunden vorgenommen (16 Kästchenlängen entsprechen  $50 \mu\text{m}$ ).

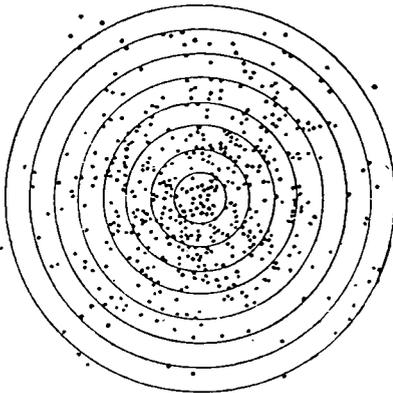


ABBILDUNG 5

Positionen von 500 Kügelchen zu einem Zeitpunkt  $t$  nach ihrem Start von einem als gemeinsam angenommenen Startpunkt.

Sekunden auf. Diese Beobachtung wurde an 500 Kügelchen durchgeführt. Man sieht in Abbildung 4 die aufeinanderfolgenden Positionen von drei von ihnen und in Abbildung 5 die Ortsveränderungen bezüglich des gleichen Ursprungs von 500 Kügelchen nach einer genügend langen Zeit  $t$  (in diesem Experiment beobachtet man natürlich die Projektion der Bewegung in eine Ebene senkrecht zur Blickrichtung des Mikroskops).

Tatsächlich fand Perrin, daß die Standardabweichung des Abszissenwertes  $X$  der Kügelchen proportional zur Quadratwurzel aus der Zeit wuchs. Darüberhinaus befand sich die Anzahl der Punkte, die innerhalb der verschiedenen in Abbildung 5 gezeichneten Ringe lag, in ausgezeichneter Übereinstimmung mit den berechneten Wahrscheinlichkeiten<sup>14</sup>. Anschließend verglich er die experimentellen Werte mit den theoretischen<sup>15</sup> der Größe  $\bar{\rho}/\sqrt{\rho^2}$  ( $\rho$  ist der Abstand vom Startpunkt). Er fand 0,894; angesichts der Anzahl der durchgeführten Beobachtungen ist dies eine ausgezeichnete Übereinstimmung mit dem theoretischen Wert von 0,886.

Eine letzte erfolgreiche experimentelle Verifizierung durch Perrin und

<sup>14</sup>Ausgehend von der Formel (D.18) kann man die Wahrscheinlichkeit  $\omega_N(\rho, \varphi) d\rho d\varphi$  dafür erhalten, daß sich die Projektion des bewegten Körpers auf eine Ebene senkrecht zur  $Oz$ -Achse in der Elementarfläche befindet, die den Punkt mit den Polarkoordinaten  $(\rho, \varphi)$  umgibt :

$$\begin{aligned} \omega_N(\rho, \varphi) d\rho d\varphi &= W_N(X) dX \cdot W_N(Y) dY \\ &= \frac{3}{2\pi N \bar{\ell}^2} \exp \left[ -\frac{3(X^2 + Y^2)}{2N \bar{\ell}^2} \right] dX dY \\ &= \frac{3}{2\pi N \bar{\ell}^2} \exp \left[ -\frac{3\rho^2}{2N \bar{\ell}^2} \right] \rho d\rho d\varphi. \end{aligned} \quad (1)$$

Daraus leitet man die Wahrscheinlichkeitsdichte  $w_N(\rho)$  dafür ab, daß der Körper sich in einer Entfernung  $\rho$  vom Ursprung befindet :

$$w_N(\rho) = \int_0^{2\pi} d\varphi \omega_N(\rho, \varphi) = \frac{3}{N \bar{\ell}^2} \rho \exp \left[ -3\rho^2/2N \bar{\ell}^2 \right]. \quad (2)$$

<sup>15</sup>Die Formeln der Fußnote 14 ergeben leicht den Mittelwert von  $\rho$  und seine Standardabweichung :

$$\bar{\rho} = \int_0^\infty \rho w_N(\rho) d\rho = \sqrt{\frac{\pi N \bar{\ell}^2}{6}}, \quad (1)$$

$$\overline{\rho^2} = \int_0^\infty \rho^2 w_N(\rho) d\rho = \frac{2}{3} N \bar{\ell}^2, \quad (2)$$

woraus

$$\frac{\bar{\rho}}{(\overline{\rho^2})^{1/2}} = \frac{\sqrt{\pi}}{2} \simeq 0,886. \quad (3)$$

folgt.

Brillouin wurde zu ihren Lebzeiten als entscheidend betrachtet. Sie bestimmten den Diffusionskoeffizienten  $D$ , der in der Formel (D.19) auftritt, und verglichen ihn mit seinem theoretischen Ausdruck<sup>16</sup>. Sie leiteten daraus einen Wert für die Boltzmann-Konstante ab, der mit Hilfe der Gaskonstanten für ideale Gase

$$R := kN_A = 8,31 \text{ Joule/Grad}, \quad (\text{D.21})$$

die Avogadro-Konstante  $N_A$  ergibt. Sie fanden

$$5,5 \cdot 10^{23} < N_A < 8,0 \cdot 10^{23}, \quad (\text{D.22})$$

in sehr guter Übereinstimmung mit den durch andere Methoden bestimmten Werten.

Ergänzende und weiterführende *Literatur* :

1. Artikel von S. Chandrasekar in A. Wax, *Selected Papers on Noise and Stochastic Processes*, Dover, New York (1954).
2. Reuben Hersh und Richard J. Griego, « Brownian Motion and Potential Theory », *Scientific American*, März 1969, S. 66.
3. F. Reif, *Fundamentals of Statistical and Thermal Physics*, McGraw-Hill, New York (1965), Kapitel 12 und 15.
4. W. Feller, *An Introduction to Probability Theory and its Applications*, John Wiley, New York (1950).

---

<sup>16</sup>Für kugelförmige Teilchen ist der Diffusionskoeffizient durch das Stokesche Gesetz gegeben :

$$D = \frac{kT}{6\pi\eta a},$$

wobei  $k$  die Boltzmann-Konstante,  $T$  die absolute Temperatur,  $\eta$  der Viskositätskoeffizient der Flüssigkeit und  $a$  der Radius der Brownschen Teilchen ist (siehe Übung IV.5).

# Zustandsdichte eines idealen Gases

---

I. Kugelvolumen und Kugeloberfläche  
im  $n$ -dimensionalen Raum

II. Berechnung der Zustandsdichte  
eines idealen Gases

---

Die Berechnung der Zustandsdichte eines Systems von  $N$  freien, unabhängigen und unterscheidbaren Teilchen (ideales Gas) wurde im Teil I, Abschnitt III.B.2.b durchgeführt. Das Ergebnis

$$\rho_N(E) = C(N)V^N \left(\frac{E}{N}\right)^{\frac{3N}{2}-1} \quad (\text{E.1})$$

läßt den numerischen Koeffizienten  $C(N)$  unbestimmt. Wir werden hier dieses Ergebnis durch eine andere Methode berechnen und den Koeffizienten  $C(N)$  explizit bestimmen.

Wir brauchen für diese Berechnung den Ausdruck für Volumen und Oberfläche einer Kugel in einem Raum von beliebiger Dimension  $n$ ; damit werden wir beginnen.

## I. Kugelvolumen und Kugeloberfläche im $n$ -dimensionalen Raum

Es sei ein euklidischer Raum in  $n$  Dimensionen gegeben, dessen Punkte sich in einem orthonormierten Koordinatensystem durch  $x_i$  ( $i = 1, 2, \dots, n$ ) darstellen lassen. Die Kugel mit dem Radius  $r$  ist definiert durch ihre Gleichung :

$$x_1^2 + x_2^2 + \dots + x_n^2 = r^2. \quad (\text{E.2})$$

Das Volumen der durch diesen Radius begrenzten Kugel ist durch die Gleichung

$$V_n(r) = \alpha_n r^n \quad (\text{E.3})$$

gegeben, wobei  $\alpha_n$  ein Koeffizient ist, der von der Dimension  $n$  des Raumes, aber nicht vom Radius  $r$  der Kugel abhängt. Die Oberfläche der Kugel erhält man daraus durch Ableitung :

$$S_n(r) = n\alpha_n r^{n-1}. \quad (\text{E.4})$$

Um  $\alpha_n$  zu bestimmen, berechnen wir das Integral

$$I_n = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} dx_1 dx_2 \dots dx_n e^{-(x_1^2+x_2^2+\dots+x_n^2)} \quad (\text{E.5})$$

nach zwei verschiedenen Methoden. Die erste besteht darin als Produkt aus  $n$  identischen Integralen

$$I_1 = \int_{-\infty}^{+\infty} dx e^{-x^2} = \sqrt{\pi} \quad (\text{E.6})$$

zu ermitteln :

$$I_n = (\pi)^{n/2}. \quad (\text{E.7})$$

Aber man kann auch, um  $I_n$  zu berechnen, den Raum in konzentrische Kugelschichten zerlegen, denn die zu integrierende Funktion hängt nur von  $r$  ab. Das Elementarvolumen zwischen zwei Schichten vom Radius  $r$  und  $r + dr$  beträgt  $dV = S - n(r)dr$  und somit

$$I_n = \int_0^{\infty} dr n \alpha_n r^{n-1} e^{-r^2}. \quad (\text{E.8})$$

Der Variablensubstitution  $u = r^2$  liefert

$$I_n = \frac{1}{2} \int_0^{\infty} du n \alpha_n u^{n/2-1} e^{-u}. \quad (\text{E.9})$$

Mit Einführung der Gammafunktion<sup>1</sup>

$$\Gamma(x) := \int_0^{\infty} e^{-t} t^{x-1} dt, \quad (\text{E.10})$$

erhält man

$$I_n = \frac{1}{2} n \alpha_n \Gamma\left(\frac{n}{2}\right) = \alpha_n \Gamma\left(\frac{n}{2} + 1\right). \quad (\text{E.11})$$

<sup>1</sup>Die Gammafunktion  $\Gamma(x)$  ist bekannt (siehe zum Beispiel L. Schwartz, *Méthodes mathématiques pour les sciences physiques*, Hermann, Paris (1961), S.344). Aus ihrer Definition (E.10) leitet man leicht durch partielle Integration

$$\Gamma(x+1) = x\Gamma(x) \quad (1)$$

ab. Darüberhinaus gilt

$$\Gamma(1) = 1, \quad (2)$$

und für ganzzahliges  $n$

$$\Gamma(n+1) = n!. \quad (3)$$

Schließlich brauchen nur noch (E.7) und (E.11) gleichgesetzt zu werden, um  $\alpha_n$  zu bestimmen :

$$\alpha_n = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2} + 1\right)}. \quad (\text{E.12})$$

Beachten wir, daß für  $n \gg 1$ ,  $\Gamma(n/2 + 1)$  durch die Stirling-Formel

$$\Gamma(x + 1) \simeq x^x e^{-x} \sqrt{2\pi x}; \quad (\text{E.13})$$

angenähert werden kann, so gilt

$$\alpha_n \simeq \left(\frac{2\pi e}{n}\right)^{n/2} \frac{1}{\sqrt{\pi n}} \quad \text{für } n \gg 1. \quad (\text{E.14})$$

## II. Berechnung der Zustandsdichte eines idealen Gases

Wir betrachten  $N$  freie, unabhängige und unterscheidbare Teilchen, die in einem Kasten makroskopischer Größe eingeschlossen sind. Die  $N$  Teilchen der Masse  $m$  sind untereinander identisch, und der Kasten mit dem Volumen  $V$  ist ein Parallelepiped mit den Kantenlängen  $L_x$ ,  $L_y$ , und  $L_z$ .<sup>2</sup>

Der Hamiltonoperator  $H$  des Systems ist nun die Summe von  $N$  Termen, von denen sich jeder auf ein Teilchen bezieht :

$$H = \sum_{i=1}^N \frac{\vec{P}_i^2}{2m}. \quad (\text{E.15})$$

Wir wählen periodische Randbedingungen. Ein stationärer Zustand des Systems ist dann durch die Menge von  $N$  dreidimensionalen Vektoren charakterisiert :

$$\{\vec{k}_1, \vec{k}_2, \dots, \vec{k}_N\}, \quad (\text{E.16})$$

$\hbar\vec{k}_i$  ist der Impuls des Teilchens ( $i$ ). Die Energie dieses Zustandes beträgt

$$E = \sum_{i=1}^N \frac{\hbar^2 \vec{k}_i^2}{2m}, \quad (\text{E.17})$$

und die Komponenten  $k_{ix}$ ,  $k_{iy}$ ,  $k_{iz}$  eines jeden Vektors  $\vec{k}_i$  müssen ganzzahlige Vielfache (positiv, negativ oder Null) von  $2\pi/L_x$ ,  $2\pi/L_y$  bzw.  $2\pi/L_z$  sein

<sup>2</sup>Tatsächlich ist die Form des Rezipienten ohne Bedeutung, solange die de Broglie-Wellenlänge eines Teilchens gegenüber den Abmessungen des Rezipienten vernachlässigbar ist (siehe Ergänzung I.A, §III.2).

(siehe I, §III.B.2.b).

Statt den Zustand des Systems von  $N$  Teilchen festzuhalten, kann man daher auch einen Vektor  $\vec{K}$  im  $3N$ -dimensionalen Raum festlegen; dieser Vektor ist durch seine Komponenten

$$\vec{K} := \{k_{1x}, k_{1y}, k_{1z}, \dots, k_{Ny}, k_{Nz}\} \quad (\text{E.18})$$

definiert. Die einzigen zulässigen Vektoren  $\vec{K}$  haben ihre Enden in einem der Knoten des  $3N$ -dimensionalen Raunggitters.

Die Beweisführung ist in diesem Fall mit der von Fußnote 19 aus Teil I identisch. Die Energie  $E$  des Systems von  $N$  Teilchen schreibt sich zu

$$E = \frac{\hbar^2 \vec{K}^2}{2m}, \quad (\text{E.19})$$

das Volumen des Elementargitters beträgt

$$v_{3N} = \left( \frac{2\pi}{L_x} \cdot \frac{2\pi}{L_y} \cdot \frac{2\pi}{L_z} \right)^N = \frac{(2\pi)^{3N}}{V^N}; \quad (\text{E.20})$$

die Anzahl von Zuständen  $\rho_N(E)dE$  mit der Energie zwischen  $E$  und  $E+dE$  ist daher

$$\rho_N(E)dE = \frac{S_{3N}(K)dK}{v_{3N}} = \frac{3N\alpha_{3N}V^N K^{3N-1}}{(2\pi)^{3N}} dK. \quad (\text{E.21})$$

$\alpha_{3N}$  ist durch (E.12) gegeben,  $K$  und  $dK$  werden mit Hilfe von (E.19) als Funktion von  $E$  und  $dE$  ausgedrückt; man erhält schließlich

$$\rho_N(E) = \frac{3N}{2} \left( \frac{m}{2\pi\hbar^2} \right)^{3N/2} \frac{V^N}{\Gamma\left(\frac{3N}{2} + 1\right)} E^{3N/2-1}. \quad (\text{E.22})$$

Dieses Resultat stimmt mit (E.1) überein, und zwar mit

$$C(N) = \frac{3}{2} \left( \frac{mN}{2\pi\hbar^2} \right)^{3N/2} \cdot \frac{1}{\Gamma\left(\frac{3N}{2} + 1\right)}. \quad (\text{E.23})$$

In allen praktischen Fällen  $N \gg 1$  kann man die Stirling-Formel (E.13) zur Berechnung von  $C(N)$  benutzen :

$$C(N) \simeq \frac{1}{2} \sqrt{\frac{3}{N\pi}} \left( \frac{me}{3\pi\hbar^2} \right)^{3N/2}. \quad (\text{E.24})$$

*Achtung* : Für das Ergebnis wurde vorausgesetzt, daß die  $N$  Teilchen unterscheidbar sind. Für identische und daher nicht unterscheidbare Teilchen läßt sich für den Fall, in dem die Maxwell-Boltzmann-Näherung gültig ist, ein angenähertes Resultat erhalten, indem man  $C(N)$  durch  $N!$  dividiert (s. auch Ergänzung II.C und Teil III, §IV.A.3).

# Elektronenzustände in Festkörpern

---

I. Beschreibung des Kristallgitters

1. Bravais-Gitter
2. Elementarzelle
3. Reziprokes Gitter

II. Elektronenzustände in einem Kristall

1. Periodisches Potential und Bloch-Theorem. Energiebänder

2. Die Randwertbedingungen von Born-van Karman

3. Zustandsdichte. Allgemeiner Ausdruck

III. Einfache Beispiele zur Ermittlung der Zustandsdichte

1. Das Modell der freien Elektronen
  2. Das Modell der starken Bindungen
- 

Wir wollen einige Begriffe über Festkörper und insbesondere über ihre Elektronenzustände einführen. Wir beschränken uns dabei auf die zum Verständnis erforderlichen elementaren Betrachtungen. Wir werden stets die einfachsten Beispiele angeben<sup>1</sup>.

## I. Beschreibung des Kristallgitters

### 1. Bravais-Gitter

Ein Kristall ist eine Gesamtheit von im Raum regulär angeordneten *identischen atomaren Objekten*. Sie bilden ein periodisches Gitter in drei Dimensionen. Die elementaren Objekte, die sich regelmäßig wiederholen, können aus einem einzigen Atom, einer Atomgruppe, einem oder mehreren Ionen oder Molekülen bestehen.

Man bezeichnet als *Bravais-Gitter des Kristalls* eine Gesamtheit von Punkten, die die zugrundeliegende Geometrie der periodischen Struktur wiedergibt und damit einer Verallgemeinerung der elementaren Objekte dient: Man wählt einen besonderen Punkt aus; das Bravais-Gitter ist dann die Gesamtheit von Positionen, die dieser besondere Punkt in der Aufeinanderfolge von identischen Objekten, aus denen der Kristall besteht, bildet. Das

---

<sup>1</sup>Für weitere Einzelheiten verweisen wir auf die Spezialliteratur, die am Ende dieser Ergänzung aufgeführt wird.

Bravais-Gitter eines bestimmten Kristalls kann durch ein System von drei unabhängigen Vektoren  $\vec{a}_1, \vec{a}_2, \vec{a}_3$  erzeugt werden. Seine Punkte werden durch die Vektoren  $\vec{R}$  in der Form

$$\vec{R} = n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3, \quad (\text{F.1})$$

gekennzeichnet, wobei  $n_1, n_2, n_3$  drei beliebige ganze Zahlen sind, die positiv, negativ oder Null sein können<sup>2</sup>. Jeder dieser Punkte, die Gitterpunkte des Gitters genannt werden, ist mit den grundlegenden elementaren Objekten verknüpft<sup>3,4</sup>.

Die Abbildung 1 zeigt als einfachstes Bravais-Gitter das einfache kubische Gitter; es wird durch drei orthogonale Vektoren gleichen Betrages erzeugt.

## 2. Elementarzelle

Als Elementarzelle eines Kristalls wird der Bereich bezeichnet, der durch eine Translation der Gesamtheit der Vektoren des Bravais-Gitters den gesamten Raum ohne Überdeckung oder Auslassung ausfüllt. Jede Elementarzelle enthält somit einen einzigen Gitterpunkt des Bravais-Gitters.

<sup>2</sup>Das System  $(\vec{a}_1, \vec{a}_2, \vec{a}_3)$  der nicht koplanaren Elementarvektoren, die ein Bravais-Gitter erzeugen, ist nicht das einzige. Es gibt eine Unzahl von verschiedenen Auswahlmöglichkeiten.

<sup>3</sup>So wie es definiert wurde, erstreckt sich ein Bravais-Gitter und damit der Kristall, den es darstellt, bis ins Unendliche. Reale Festkörper sind endlich, aber ihre Ausmaße sind so, daß die Mehrzahl der in ihnen enthaltenen Atome zu weit von der Oberfläche entfernt sind, um das Vorhandensein dieser Begrenzung zu merken; die Approximation des unendlichen Kristalls ist daher oft ausreichend. Wenn es dennoch erforderlich wird, die Begrenzung des Kristalls zu berücksichtigen, so ordnet man ihm aus praktischen Gründen die möglichst einfachste Form zu. Man begrenzt das Bravais-Gitter auf Vektoren  $\vec{R}$  (Formel (F.1)), die zum Beispiel

$$0 \leq n_1 < N_1,$$

$$0 \leq n_2 < N_2,$$

$$0 \leq n_3 < N_3,$$

erfüllen, wobei  $N_1, N_2, N_3$  drei sehr große ganze Zahlen sind, deren Produkt  $N_1 \cdot N_2 \cdot N_3$  gleich der Anzahl  $N$  von Gitterpunkten in dem betrachteten realen Kristall ist.

<sup>4</sup>Ein realer Kristall zeigt auch Fehlstellen, die von dem Bravais-Gitter nicht berücksichtigt werden. Wir nehmen an, daß diese Fehlstellen selten sind und nicht die Eigenschaften des Kristalls beeinflussen. Wir beschränken uns daher auf *ideale Kristalle*. Aber selbst in diesem Fall besetzen infolge von thermischer Bewegung die Atome oder Ionen des Kristalls nicht genau die Positionen, die durch das Bravais-Gitter definiert werden; sie weichen davon auf zufällige Weise ab. Sie weichen jedoch nur wenig ab, wenn die Temperatur deutlich niedriger als der Schmelzpunkt des Kristalls ist. Wir vernachlässigen hier diesen Effekt, auf den wir im Teil III (Ergänzung III.E) zurückkommen.

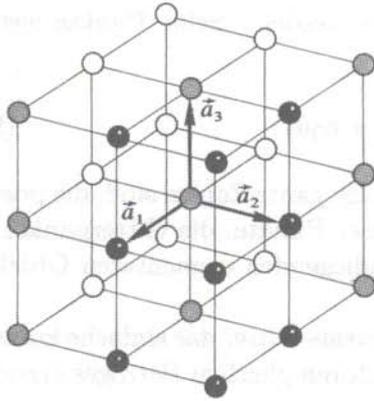


ABBILDUNG 1

**Struktur des einfachen kubischen Gitters. Es kann durch die drei orthogonalen und betragsgleichen Vektoren ( $\vec{a}_1, \vec{a}_2, \vec{a}_3$ ) erzeugt werden.**

Um einen kristallinen Festkörper zu definieren, genügt es daher, das entsprechende Bravais-Gitter und die Position der zugrundeliegenden Atome, Ionen oder Moleküle im Inneren der Elementarzelle anzugeben.

Es gibt unendlich viele Auswahlmöglichkeiten für die Elementarzelle eines gegebenen Kristalls (Abb. 2), aber alle möglichen Elementarzellen haben das gleiche Volumen.

Es ist praktisch, eine Elementarzelle auszuwählen, die die gleichen Symmetrieelemente besitzt wie das Bravais-Gitter. Die einfachste ist die *Wigner-Seitz-Zelle*: Die Elementarzelle nach Wigner-Seitz um einen gegebenen Gitterpunkt herum besteht aus der Menge der Punkte des Raumes, die diesem Gitterpunkt näher sind als allen anderen Gitterpunkte des Gitters. Dies ist das kleinste Polyeder, das durch Ebenen begrenzt wird, die in den Mitten der Verbindungslinien zwischen dem betrachteten Gitterpunkt mit allen anderen liegen (Abb. 3). Es läßt sich leicht zeigen, daß die Wigner-Seitz-Zelle eines einfachen kubischen Kristalls ein Würfel ist.

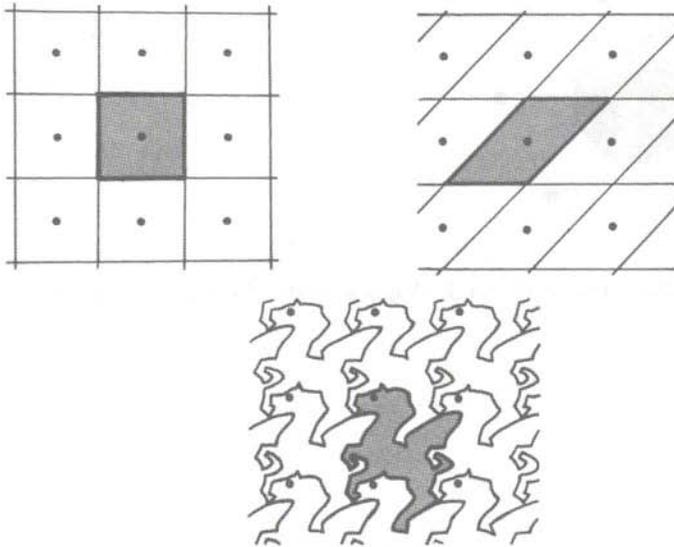


ABBILDUNG 2

Zweidimensionales quadratisches Gitter mit drei Auswahlmöglichkeiten der Elementarzelle (die dritte stammt von M. C. Escher).

### 3. Reziprokes Gitter

Ist ein Bravais-Gitter gegeben, dann wird sein *reziprokes Gitter* durch die Enden der Vektoren  $\vec{K}$  so aufgebaut, daß gilt

$$\exp(i\vec{K} \cdot \vec{R}) = 1 \quad (\text{F.2})$$

für alle  $\vec{R}$ , die mit dem ursprünglichen Gitter verknüpft sind (was auch *direktes Gitter* genannt wird). Das reziproke Gitter ist wiederum ein Bravais-Gitter. Tatsächlich sind für ein System von Elementarvektoren  $(\vec{a}_1, \vec{a}_2, \vec{a}_3)$ , die das direkte Gitter erzeugen, alle Vektoren des reziproken Gitters von der Form

$$\vec{K} = m_1 \vec{b}_1 + m_2 \vec{b}_2 + m_3 \vec{b}_3, \quad (\text{F.3})$$

wobei  $m_1, m_2, m_3$  drei ganze Zahlen sind, die positiv, negativ oder Null sein können, und  $\vec{b}_1, \vec{b}_2, \vec{b}_3$  drei nicht koplanare Vektoren darstellen, die durch

$$\vec{b}_1 := 2\pi \frac{\vec{a}_2 \times \vec{a}_3}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)}, \quad (\text{F.4.a})$$

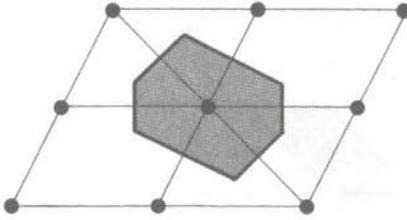


ABBILDUNG 3

Konstruktion der Elementarzelle nach Wigner-Seitz für ein zweidimensionales Bravais-Gitter.

$$\vec{b}_2 := 2\pi \frac{\vec{a}_3 \times \vec{a}_1}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)}, \quad (\text{F.4.b})$$

$$\vec{b}_3 := 2\pi \frac{\vec{a}_1 \times \vec{a}_2}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)}, \quad (\text{F.4.c})$$

definiert sind.

Somit ist das reziproke Gitter eines einfachkubischen Kristalls selbst wieder einfachkubisch. Es besitzt eine würfelförmige Elementarzelle mit den Seitenlängen  $2\pi/a$  ( $a$  ist die Seitenlänge der kubischen Elementarzelle des direkten Gitters).

Im allgemeinen beträgt im Vektorraum  $\vec{K}$  das Volumen der Elementarzelle des reziproken Gitters<sup>5</sup>  $(2\pi)^3/v$ , wobei  $v$  das Volumen der Elementarzelle des direkten Gitters ist.

Die Elementarzelle von Wigner-Seitz des reziproken Gitters wird *erste Brillouin-Zone* des Kristalls genannt.

<sup>5</sup>Um dieses Resultat zu beweisen, muß man das gemischte Produkt  $(\vec{b}_1 \times \vec{b}_2) \cdot \vec{b}_3$  mit Hilfe der Formel (F.4) und mit der Formel des doppelten vektoriellen Produktes

$$\vec{v}_1 \times (\vec{v}_2 \times \vec{v}_3) = (\vec{v}_1 \cdot \vec{v}_3) \vec{v}_2 - (\vec{v}_1 \cdot \vec{v}_2) \vec{v}_3$$

berechnen.

## II. Elektronenzustände in einem Kristall

### 1. Periodisches Potential und Bloch-Theorem. Energiebänder

Die Elektronen eines Festkörpers sind einem periodischen Potential ausgesetzt, das von der Gesamtheit der atomaren Kerne verursacht wird. Sie wechselwirken gleichfalls untereinander; dennoch betrachtet man diese Elektronen wie *unabhängige Teilchen*, die einem « mittleren Potential » unterworfen sind, das von den Kernen und den anderen Elektronen des Kristalls gebildet wird<sup>6</sup>. Dieses mittlere Potential hat in jeder Elementarzelle die gleichen Werte, so daß jedes Elektron sich in einem *periodischen Potential*  $V(\vec{r})$  bewegt :

$$V(\vec{r}) = V(\vec{r} + \vec{R}). \quad (\text{F.5})$$

Das *Bloch-Theorem*, das wir hier ohne Beweis zugrundelegen (siehe Literaturstelle 1, Kapitel 8 und Literaturstelle 2, Kapitel 9), liefert die allgemeine Form der stationären räumlichen Wellenfunktionen<sup>7</sup>  $\psi(\vec{r})$ , die Lösungen der stationären Schrödinger-Gleichung

$$\left[ -\frac{\hbar^2}{2m}\Delta + V(\vec{r}) \right] \psi(\vec{r}) = \varepsilon \psi(\vec{r}) \quad (\text{F.6})$$

für ein Elektron in dem periodischen Potential  $V(\vec{r})$  sind. Diese Wellenfunktionen können stets als *Produkt einer ebenen Welle mit einer periodischen Funktion* ausgedrückt werden, deren Periodizität die des Bravais-Gitters des Kristalls ist<sup>8</sup> :

$$\psi_{\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} u_{\vec{k}}(\vec{r}), \quad (\text{F.7.a})$$

<sup>6</sup>Es ist nicht von vornherein klar, daß man die Coulomb-Wechselwirkung zwischen den Elektronen korrekt durch ein mittleres Potential beschreiben kann. Die Näherung ist jedoch sehr gut, und zahlreiche Eigenschaften der Festkörper erklären sich recht einfach mit dem Modell von unabhängigen Elektronen.

Dennoch muß man beachten, daß dieses mittlere Potential, mit dem man die Elektronenwellenfunktionen berechnen kann, ausgehend von den gleichen Funktionen bestimmt werden muß : Ein gegebenes Elektron « fühlt » im Mittel (anders als die Kernanziehung) die Ladungsverteilung, die von allen anderen Elektronen verursacht wird, einer Ladungsverteilung, die von ihren Wellenfunktionen abhängt. Das Problem muß also auf « selbstkonsistente » Art und Weise gelöst werden.

<sup>7</sup>Der Hamiltonoperator von (F.6) hängt nicht vom Spin ab. In dieser Näherung ist der Zustand eines Elektrons das Tensorprodukt aus seinem Raumzustand und seinem Spinzustand (Anhang II, §III.2.b).

<sup>8</sup>Somit besitzt die gequantelte Wahrscheinlichkeitsdichte  $|\psi_{\vec{k}}(\vec{r})|^2 = |u_{\vec{k}}(\vec{r})|^2$  die Symmetrie des Kristalls : Wie es sich gehört, nimmt  $|\psi_{\vec{k}}(\vec{r})|^2$  die gleichen Werte in jeder Elementarzelle an. Die Wellenfunktion  $\psi_{\vec{k}}(\vec{r})$  selbst variiert von einer Zelle zur anderen, aber nur mittels eines Faktors  $e^{i\vec{k}\cdot\vec{r}}$  (dies ist eine allgemeine Eigenschaft von Wellenfunktionen unter periodischen Bedingungen).

mit

$$u_{\vec{k}'}(\vec{r} + \vec{R}) = u_{\vec{k}}(\vec{r}). \quad (\text{F.7.b})$$

Wir merken noch an, daß man sich darauf beschränken kann,  $\vec{k}$  aus der ersten Brioullin-Zone zu wählen. Alle  $\vec{k}'$  außerhalb davon können geschrieben werden zu

$$\vec{k}' = \vec{k} + \vec{K}, \quad (\text{F.8})$$

wobei  $\vec{K}$  ein Vektor des reziproken Gitters und  $\vec{k}$  ein Vektor der ersten Brioullin-Zone ist. Eine Wellenfunktion, die das Produkt von  $\exp[i\vec{k} \cdot \vec{r}]$  mit der Funktion  $u(\vec{r})$  ist, ist auch gleich dem Produkt von  $\exp[i\vec{k} \cdot \vec{r}]$  mit der nach (F.2) und (F.7.b) periodischen Funktion  $u(\vec{r}) \exp[i\vec{k} \cdot \vec{r}]$ .

Für einen unendlichen Kristall kann  $\vec{k}$  beliebige Werte im Inneren der ersten Brioullin-Zone annehmen; die Zustände von benachbarten  $\vec{k}$  haben ähnliche Eigenschaften, ihre Energien  $\varepsilon(\vec{k})$  ändern sich kontinuierlich mit  $\vec{k}$ . Wenn  $\vec{k}$  die erste Brioullin-Zone beschreibt, dann erstreckt sich  $\varepsilon(\vec{k})$  über einen Energiebereich, den man als *erlaubtes Energieband bezeichnet*. Ein Kristall verfügt über mehrere erlaubte Energiebänder<sup>9</sup>, die oftmals durch

<sup>9</sup>In einem isolierten Atom haben die verschiedenen Niveaus, die ein Elektron besetzen kann, auch unterschiedliche Energien (mit Zwischenräumen in der Größenordnung von Elektronenvolt). Wenn man zwei identische Atome einander nähert, dann können die Elektronen von einem zum anderen « springen », was zu dem Phänomen der *Quantenresonanz* Anlaß gibt und zu einer Verdopplung der Niveaus führt: Jeder atomare Zustand wird somit an den entsprechenden Zustand des anderen Atoms gekoppelt; diese Kopplung erzeugt zwei Zustände molekularen Typs (das heißt sie vereinigen die beiden Kerne) mit benachbarten aber verschiedenen Energien. In einem Kristall betreffen die Kopplung und die Resonanz eine große Anzahl  $N$  von Atomen, so daß jeder der anfänglichen atomaren Zustände eine Gruppe von  $N$  verschiedenen, aber sehr nahe beinanderliegenden Energiezuständen entstehen läßt; diese bilden ein *erlaubtes Band* (die  $N$  Niveaus, wobei  $N$  in der Größenordnung der Avogadro-Konstanten ist, sind in einem Band von der Breite in der Größenordnung von Elektronenvolt konzentriert). Man erhält, indem man die Darstellung (F.7) der Wellenfunktionen in die Schrödinger-Gleichung einsetzt, die folgende Relation

$$\left[ \frac{\hbar^2}{2m} (-i\nabla + \vec{k})^2 + V(\vec{r}) \right] u_{\vec{k}}(\vec{r}) = \varepsilon(\vec{k}) u_{\vec{k}}(\vec{r}), \quad (1)$$

mit der Bedingung

$$u_{\vec{k}}(\vec{r}) = u_{\vec{k}}(\vec{r} + \vec{R}). \quad (2)$$

Für festes  $\vec{k}$  suche man die Menge der Eigenfunktionen  $u_{\vec{k}}(\vec{r})$  und der Eigenwerte  $\varepsilon(\vec{k})$  eines hermiteschen Operators mit (infolge von Beziehung (2)) periodischen Randwertbedingungen in einem auf die Elementarzelle beschränkten Gebiet. Man zeige allgemein, daß bei einem solchen Problem die Eigenwerte ein diskretes Spektrum  $\varepsilon_n(\vec{k})$  ( $n = 1, 2, \dots$ ) bilden. Wenn  $\vec{k}$  für festes  $n$  die erste Brioullin-Zone durchläuft, dann erzeugt  $\varepsilon_n(\vec{k})$  ein *erlaubtes Band* mit ganzzahligem  $n$  (Abb. 4). Wenn man berücksichtigt, daß die Atomanzahl  $N$  im Kristall endlich ist, umfaßt dieses erlaubte Band  $N$  verschiedene Niveaus (siehe Abschnitt 2).

nicht erlaubte Bänder voneinander getrennt werden, also Energiebereichen, für die kein Zustand (das heißt keine Lösung der Schrödinger-Gleichung (F.6)) existiert.

Somit wird die räumliche Wellenfunktion eines Elektrons in einem Kri-

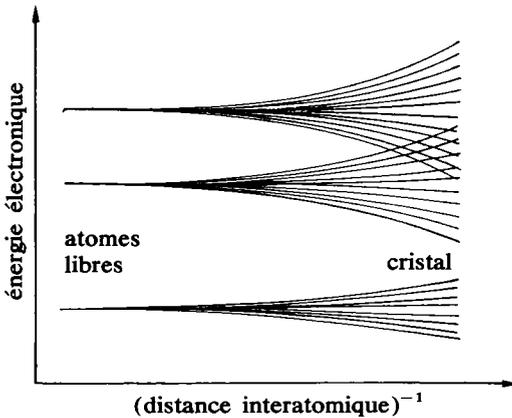


ABBILDUNG 4

Schematische Zuordnung zwischen den Energieniveaus in einem isolierten Atom und in einem Kristall.

stall durch ihren Vektor  $\vec{k}$  (der aus der ersten Brioullin-Zone ausgewählt wird) und dem Band  $n$ , zu dem sie gehört, spezifiziert. Man schreibt diese Funktion zu

$$\psi_{n,\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} u_{n,\vec{k}}(\vec{r}), \quad (\text{F.9})$$

und bezeichnet  $\varepsilon_n(\vec{k})$  als zugehörige Energie.

## 2. Die Randwertbedingungen von Born-von Karman

Wie im Fall der freien Teilchen beschränkt die endliche Größe des Kristalls die Werte von  $\vec{k}$  auf eine diskrete Menge; die Randwertbedingungen, denen die Oberfläche des Kristalls unterworfen ist, bestimmen diese erlaubten Werte von  $\vec{k}$ . Gewöhnlich wählt man die Born-von Karman-Randbedingungen, die eine Verallgemeinerung der periodischen Bedingungen, wie sie im Anhang II, Abschnitt II.4.c für ein Teilchen in einem Parallelepiped untersucht werden, darstellen.

Für einen Kristall ist es praktisch, ein Volumen zu benutzen, das eine ganzzahlige Anzahl von Elementarzellen enthält (vgl. Fußnote 3); man setzt

daher an

$$\psi_{n,\vec{k}}(\vec{r} + N_i \vec{a}_i) = \psi_{n,\vec{k}}(\vec{r}) \quad (i = 1, 2, 3), \quad (\text{F.10})$$

wobei  $\vec{a}_1, \vec{a}_2, \vec{a}_3$  drei Elementarvektoren sind, die das Bravais-Gitter definieren, und  $N_1, N_2, N_3$  drei sehr große ganze Zahlen von der Größenordnung  $N^{1/3}$  ( $N = N_1 N_2 N_3$  ist die Gesamtzahl der Elementarzellen in dem betrachteten endlichen Kristall).

Die Beziehung (F.10) liefert in Verbindung mit der Form (F.7) der Wellenfunktionen

$$\exp \left[ i N_j \vec{k} \cdot \vec{a}_j \right] = 1 \quad (j = 1, 2, 3); \quad (\text{F.11})$$

das heißt, daß die erlaubten Werte von  $\vec{k}$  diskret und von der Form

$$\vec{K} = \frac{p_1}{N_1} \vec{b}_1 + \frac{p_2}{N_2} \vec{b}_2 + \frac{p_3}{N_3} \vec{b}_3, \quad (\text{F.12})$$

sind, wobei die  $p_i$  ganze Zahlen sind, die positiv, negativ oder Null sein können, und die  $\vec{b}_i$  durch (F.4) definiert sind.

Wie im Abschnitt III.B.2 des Teils I berechnen wir das Elementarvolumen  $\Delta V_{\vec{k}}$  des reziproken Raumes, das jedem erlaubten Wert von  $\vec{k}$  zugeordnet ist. Es ist das Volumen des Parallelepipeds, das durch die Vektoren  $\vec{b}_1/N_1, \vec{b}_2/N_2, \vec{b}_3/N_3$  aufgespannt wird :

$$\Delta V_{\vec{k}} = \frac{1}{N} \vec{b}_1 \cdot (\vec{b}_2 \times \vec{b}_3). \quad (\text{F.13})$$

Nun ist aber das Spatprodukt  $\vec{b}_1 \cdot (\vec{b}_2 \times \vec{b}_3)$  gleich dem Volumen der Elementarzelle des reziproken Gitters, das heißt dem Volumen der ersten Brioullin-Zone; die Anzahl der erlaubten Wellenvektoren  $\vec{k}$  (in der ersten Brioullin-Zone) ist daher einfach gleich der Anzahl  $N$  der Gitterpunkte im Kristall<sup>10</sup>.

Wenn darüberhinaus  $v$  das Volumen der Elementarzelle des direkten Gitters ist, so beträgt das der ersten Brioullin-Zone im reziproken Raum  $(2\pi)^3/v$ . Daraus folgt, daß

$$\Delta V_{\vec{k}} = \frac{(2\pi)^3}{Nv}; \quad (\text{F.14})$$

und da  $Nv$  gleich dem Kristallvolumen  $V$  ist, wird

$$\Delta V_{\vec{k}} = \frac{(2\pi)^3}{V}. \quad (\text{F.15})$$

*Das Elementarvolumen mit erlaubten Werten von  $\vec{k}$  ist daher das gleiche wie für ein freies Teilchen, das in einem Kasten von gleichem Volumen wie der Kristall eingeschlossen ist.*

<sup>10</sup>Dieses Ergebnis war vorauszusetzen; es entspricht einfach der Erhaltung der Anzahl von Freiheitsgraden (vgl. voranstehende Fußnote).

### 3. Zustandsdichte. Allgemeiner Ausdruck

Zahlreiche makroskopische Eigenschaften des Kristalls werden erhalten, indem man Mittelwerte der Elektronenzustände berechnet (Teil I, §III.B). Im Modell der unabhängigen Elektronen bezieht man sich häufig auf Mittelwerte von Größen  $A$ , die nur einen Elektronenzustand beeinflussen, und zwar in der Form

$$\bar{A} = \sum_{\sigma=\pm 1/2} \sum_n \sum_{\vec{k} \in \text{B.Z.}} A_{n,\sigma}(\vec{k}) P_{n,\sigma}(\vec{k}), \quad (\text{F.16})$$

wobei  $P_{n,\sigma}(\vec{k})$  die Wahrscheinlichkeit dafür ist, daß ein Elektron sich in einem Zustand des Bandes  $n$  befindet, der durch den Wellenvektor  $\vec{k}$ , der Projektion  $\sigma$  des Spins und  $A_{n,\sigma}(\vec{k})$ , dem Wert, den die Größe  $A$  in diesem Zustand annimmt, charakterisiert wird; die Summe über  $\vec{k}$  ist auf die erlaubten Werte von  $\vec{k}$  beschränkt, die zur ersten Brioullin-Zone (B.Z.) gehören.

Für Kristalle von makroskopischer Größe sind die erlaubten Werte von  $\vec{k}$  eng benachbart, und die Summe über  $\vec{k}$  kann durch ein Integral ersetzt werden; unter Berücksichtigung von (F.15) erhält man

$$\bar{A} = \sum_{\sigma=\pm 1/2} \sum_n \frac{V}{(2\pi)^3} \int_{\text{B.Z.}} d^3k A_{n,\sigma}(\vec{k}) P_{n,\sigma}(\vec{k}). \quad (\text{F.17})$$

Oft ist die zu integrierende Größe von  $\sigma$  unabhängig und hängt von  $n$  und von  $\vec{k}$  nur über die Energie  $\varepsilon_n(\vec{k})$  ab :

$$A_{n,\sigma}(\vec{k}) P_{n,\sigma}(\vec{k}) \equiv F \left[ \varepsilon_n(\vec{k}) \right]. \quad (\text{F.18})$$

Wenn dies der Fall ist, kann die Integration über  $\vec{k}$  durch eine Integration über die Energie ersetzt werden : Indem wir die Elektronenzustandsdichte für das Band  $n$  einführen, definiert durch<sup>11</sup>

$$\rho_n(\varepsilon) := \frac{2V}{(2\pi)^3} \int_{\text{B.Z.}} d^3k \delta[\varepsilon - \varepsilon_n(\vec{k})], \quad (\text{F.19})$$

<sup>11</sup>Der Faktor 2, der in der Definition von  $\rho_n(\varepsilon)$  auftritt, stammt von den beiden Spin-zuständen des Elektrons. Wenn der Hamiltonoperator vom Spin abhängt (in Gegenwart eines magnetischen Feldes zum Beispiel), dann gibt es zwei verschiedene Energieniveaus für  $\vec{k}$  und ein gegebenes  $n$  :  $\varepsilon_{n,\sigma}(\vec{k})$  mit  $\sigma = \pm 1/2$ . Man definiert dann eine *Zustandsdichte für jede Spinrichtung* (siehe zum Beispiel Ergänzung VI.A, §II.2.c) :

$$\rho_{n,\sigma}(\varepsilon) = \frac{V}{(2\pi)^3} \int_{\text{B.Z.}} d^3k \delta[\varepsilon - \varepsilon_{n,\sigma}(\vec{k})].$$

folgt tatsächlich

$$\bar{A} = \sum_n \int d\varepsilon \rho_n(\varepsilon) F(\varepsilon); \quad (\text{F.20})$$

die Gesamtzustandsdichte für die Elektronen des Kristalls ist

$$\rho(\varepsilon) := \sum_n \rho_n(\varepsilon), \quad (\text{F.21})$$

und (F.20) wird dann in der gebräuchlichen Form

$$\bar{A} = \int d\varepsilon \rho(\varepsilon) F(\varepsilon). \quad (\text{F.22})$$

geschrieben.

Man kann den Ausdruck (F.19) von  $\rho_n(\varepsilon)$  umformen, indem man die Integration in bezug auf einen der Komponenten des Vektors  $\vec{k}$  durchführt; das Volumenintegral wird dann in ein Flächenintegral transformiert<sup>12,13</sup> :

$$\rho_n(\varepsilon) = \frac{V}{4\pi^3} \int_{S_n(\varepsilon)} \frac{dS}{|\vec{\nabla}_{\varepsilon}(\vec{k})|}, \quad (\text{F.23})$$

wobei  $S_n(\varepsilon)$  die im Raum der  $\vec{k}$  durch die Gleichung  $\varepsilon_n(\vec{k}) = \varepsilon$  definierte Oberfläche ist, und die auf die erste Brioullin-Zone beschränkt ist.

### III. Einfache Beispiele zur Ermittlung der Zustandsdichte

Die Kenntnis der Bänderstruktur, das heißt der Funktionen  $\varepsilon_n(\vec{k})$  für  $n = 1, 2, \dots$  gestattet schließlich, die Elektronenzustandsdichten eines Festkörpers zu berechnen. Die Hauptschwierigkeit besteht in der *gemeinsamen* Bestimmung der Wellenfunktionen  $\psi_{n,\vec{k}}$  der entsprechenden Energien  $\varepsilon_n(\vec{k})$  und dem damit verknüpften periodischen Potential  $V(\vec{r})$  (vgl. Fußnote 6). Dieses Problem ist schwierig und übersteigt bei weitem den Rahmen dieses Buches. Wir beschränken uns auf zwei sehr einfache Modelle, mit denen man zumindest qualitativ zahlreiche praktische Fälle behandeln kann.

<sup>12</sup>Der Übergang von (F.19) nach (F.23) wird in der Fußnote 7 der Ergänzung I.B explizit durchgeführt.

<sup>13</sup>Der Integrand in (F.23) divergiert an den Stellen, wo  $\vec{\nabla}_{\varepsilon}(\vec{k})$  verschwindet. Im dreidimensionalen Fall sind diese Singularitäten im allgemeinen integrierbar und liefern für die Zustandsdichte endliche Werte. Dennoch divergiert die Ableitung  $d\rho_n/d\varepsilon$  an diesen Stellen, die man als « van Hove-Singularitäten » bezeichnet. Dies ereignet sich insbesondere an den Rändern eines erlaubtes Band : Die Zustandsdichten verschwinden hier mit unendlicher Ableitung (siehe Literaturstelle 1, Kapitel 8 und 2, Ergänzung C).