

Reif
Statistische Physik und
Theorie der Wärme

Frederick Reif

Statistische Physik und Theorie der Wärme

Bearbeitung und wissenschaftliche Redaktion
der deutschsprachigen Ausgabe
W. Muschik

übersetzt von
K.-P. Charlé · W. Muschik · H.U. Zimmer · J. Zwanzger

Dritte, durchgesehene Auflage



Walter de Gruyter · Berlin · New York 1987

Titel der Originalausgabe: *Fundamentals of Statistical and Thermal Physics*
McGraw-Hill Book Company, Copyright © 1965 by McGraw-Hill, Inc.

Autor der Originalausgabe
Frederick Reif
Professor of Physics
University of California, Berkeley

Bearbeitung und wissenschaftliche Redaktion der deutschsprachigen Ausgabe
Prof. Dr. W. Muschik
Institut für Theoretische Physik
Technische Universität Berlin

1. Auflage 1975, erschienen unter dem Titel:
Grundlagen der Physikalischen Statistik
und der Physik der Wärme

2. Auflage 1985
3. Auflage 1987

CIP-Kurztitelaufnahme der Deutschen Bibliothek

Reif, Frederick:
Statistische Physik und Theorie der Wärme / Frederick Reif. Bearb.
u. wiss. Red. d. dt.-sprachigen Ausg. W. Muschik. Übers. von K.-P.
Charlé . . . - 3., durchges. Aufl. - Berlin ; New York : de Gruyter, 1987. -
Einheitssacht.: Fundamentals of statistical and thermal physics
<dt.>
ISBN 3-11-011383-X
NE: Muschik, Wolfgang [Bearb.]

Copyright © 1975, 1984, 1987 by Walter de Gruyter & Co., Berlin 30. Alle Rechte, insbesondere das
Recht der Vervielfältigung und Verbreitung sowie der Übersetzung, vorbehalten. Kein Teil des
Werkes darf in irgendeiner Form (durch Fotokopie, Mikrofilm oder ein anderes Verfahren) ohne
schriftliche Genehmigung des Verlages reproduziert oder unter Verwendung elektronischer Systeme
verarbeitet, vervielfältigt oder verbreitet werden. Printed in Germany.

Satz: Compositorsatz Verena Boldin, Aachen. - Druck: Gerike GmbH, Berlin.
Bindearbeiten: Dieter Mikolaj, Berlin. - Einbandentwurf: Thomas Bonnie, Hamburg.

Vorwort zur dritten Auflage

Diese dritte Auflage wurde kritisch durchgesehen, entdeckte Fehler wurden berichtigt. Sie ist ein nahezu unveränderter Nachdruck der zweiten Auflage.

Geplante Ergänzungen hätten eine weitergehende Umarbeitung des Buches erforderlich gemacht, was aber, ohne den didaktischen Charakter des Buches zu stören, bei begrenztem Umfang nicht einfach ist. Dies soll daher einer nächsten Auflage vorbehalten bleiben.

Möge auch die dritte Auflage eine so freundliche Aufnahme finden wie die beiden Voraufgaben.

Berlin, im Juni 1987

W. Muschik

Vorwort zur zweiten Auflage

In die zweite Auflage wurden einige Ergänzungen aufgenommen: so wurde unter anderem der stochastische Prozeß genauer definiert, für die Herleitung der Spektraldichten der FOURIERSche Satz verwendet und der Abschnitt über die WIENER-CHINTSCHIN-Relationen neu formuliert. Ein Kapitel⁺ 15.17a über die BROWNSche Bewegung als stochastischer Prozeß ist ergänzend zur ersten Auflage hinzugefügt worden. Außerdem wurde ein Symbolverzeichnis angefertigt und das Register erweitert. Die Literaturangaben wurden teilweise auf einen neueren Stand gebracht. Wo immer entdeckt, sind Druckfehler der ersten Auflage berichtigt worden und in dieser Hinsicht gilt mein Dank jenen Lesern, die sich daran beteiligt haben. Der Titel der zweiten Auflage ist gegenüber der ersten verändert worden: Er wurde mehr dem deutschen Sprachgebrauch angepaßt. Möge auch die zweite Auflage des bewährten Buches von F. Reif, das inzwischen durch die von K.-P.CHARLÈ und H.U. ZIMMER bearbeitete Aufgabensammlung ergänzt wurde, allen Benutzern eine Hilfe sein.

Berlin, im August 1984
Institut für Theoretische Physik
Technische Universität Berlin

W. Muschik

Vorwort zur deutschen Übersetzung

In den letzten Jahren hat die Statistische Physik eine Entwicklung erfahren, die noch vor kurzer Zeit nicht abzusehen war. Insbesondere die Theorie der Phasenübergänge und der Stabilität, die Theorie dissipativer Systeme sowie algebraische und feldtheoretische Methoden der Statistischen Thermodynamik haben neben einer steigenden Anzahl von Anwendungen der Statistischen Physik in nichtphysikalischen Disziplinen – wie der Biologie – mehr und mehr Bedeutung bekommen.

Daher erscheint es geboten, den Studenten möglichst frühzeitig mit den grundlegenden Methoden der Statistischen Physik vertraut zu machen, und zwar in einem Umfang, der über die übliche Einführung hinausgeht, wie sie im Kurs „Theoretische Physik“ gegeben wird. Da es in der deutschsprachigen Literatur bisher nur sehr wenige Lehrbücher gibt, die die phänomenologische und die statistische Thermodynamik gemeinsam aus einer Wurzel, nämlich aus dem Grundpostulat der Statistischen Thermodynamik, herleiten, ist es angebracht, ein didaktisch hervorragendes Lehrbuch dieser Richtung aus der vielfältigen angelsächsischen Literatur in die deutsche Sprache zu übersetzen. Die Wahl fiel auf Reifs Lehrbuch „Fundamentals of statistical and thermal physics“. Für wen das Buch geschrieben wurde, welche Voraussetzungen zu seinem Studium erforderlich sind und wie es benutzt werden sollte, entnimmt man am besten dem nachstehenden Vorwort zur amerikanischen Ausgabe.

Daß das Buch in relativ kurzer Zeit übersetzt werden konnte, verdanke ich meinen Mitarbeitern, den Herren Dr. rer. nat. K.-P. Charlé, Dipl.Phys. U. Zimmer und Dipl.Phys. J. Zwanzger. Mit ihnen wurde die Übersetzung im Team angefertigt, die Bearbeitung einiger Kapitel bis ins einzelne besprochen und oftmals nach ihren Vorschlägen vorgenommen. Einige Abschnitte wurden umgestaltet (z.B. + 13.3 Bahnintegralmethode), andere wurden völlig neu geschrieben (z.B. + A.7 Diracsche Deltafunktion). Ein Abschnitt (+ 15.19 Skizze thermodynamischer Theorien irreversibler Prozesse) wurde dem Original hinzugefügt, weil in einem einführenden Werk meiner Meinung nach die phänomenologische Nichtgleichgewichtsthermodynamik nicht völlig fehlen darf. Das Glossar und die Kapitelzusammenfassungen sind ebenfalls ergänzend hinzugefügt worden, um dem Studenten die Übersicht zu erleichtern. Wo immer es nötig erschien, wurde der Text mit erläuternden oder hinweisenden Fußnoten versehen. Alle Zusätze zum amerikanischen Original und stärker bearbeitete Stellen sind mit einem Kreuz + gekennzeichnet. Der Text wurde so frei wie möglich übersetzt. Dabei hält sich die Übersetzung so eng wie nötig an die amerikanische Ausgabe, um nicht die didaktischen Bestrebungen des Autors zu verfälschen.

Herrn O. Bibl.-Rat Dipl.Ing. P. Detje bin ich für die Überprüfung der von Reif angegebenen Literaturstellen dankbar. Herr Detje hat diese Zusammenstellung auf den neuesten Stand gebracht und mich bei der ergänzenden Auswahl neuerer Literaturangaben beraten. Frau Oberstudienrat H. Scheffler war mir dankenswerterweise bei der Übersetzung des Vorworts zur amerikanischen Ausgabe und der Einleitung behilflich. Nicht zuletzt möchte ich – auch im Namen des Teams – Herrn Dr. R. Weber, Verlag de Gruyter, für das verständnisvolle Entgegenkommen danken, das unsere Arbeit wesentlich erleichtert hat. Auf ihn geht auch die Auswahl des Buches zur Übersetzung zurück.

Berlin, im Sept. 1975
Institut für Theoretische Physik
der Technischen Universität

W. Muschik

Vorwort zur amerikanischen Ausgabe

In diesem Buch werden einige grundlegende physikalische Begriffe und Methoden erörtert, die für die Beschreibung solcher Systeme geeignet sind, die aus sehr vielen Teilchen bestehen. Es ist insbesondere beabsichtigt, die Gebiete der Thermodynamik, der statistischen Mechanik und der kinetischen Theorie von einem vereinheitlichten und modernen Gesichtspunkt aus darzustellen. Demgemäß weicht die Darstellung von der historischen Entwicklung ab, in der die Thermostatik die erste dieser Disziplinen war, die als ein unabhängiges Gebiet entstand. Die historische Entwicklung der Begriffe, die sich mit Wärme, Arbeit und der kinetischen Theorie der Materie befassen, ist interessant und lehrreich, aber sie gibt nicht die einfachste oder durchsichtigste Darstellung dieser Gegenstände wieder. Ich habe daher die historische Betrachtungsweise zugunsten einer Darstellung aufgegeben, die die wesentliche Einheit des behandelten Themas hervorhebt und das physikalische Verständnis zu entwickeln bestrebt ist, indem der mikroskopische Inhalt der Theorie betont wird.

Atome und Moleküle sind Gebilde, die in der modernen Wissenschaft so nachgewiesen sind, daß ein Zweifel an ihnen unangemessen erscheint. Aus diesem Grund habe ich es absichtlich vorgezogen, die ganze Erörterung auf der Voraussetzung zu gründen, daß alle makroskopischen Systeme letztlich aus Atomen bestehen, die den Gesetzen der Quantenmechanik gehorchen. Eine Kombination dieser mikroskopischen Beschreibung mit einigen statistischen Postulaten führt dann leicht zu einigen sehr allgemeinen Schlußfolgerungen auf einer rein *makroskopischen* Ebene. Diese Schlußfolgerungen sind unbestreitbar *unabhängig* von irgendwelchen besonderen Modellen, die über die Natur oder Wechselwirkung der Teilchen in den betrachteten Systemen vorausgesetzt werden könnten. Diese Schlußfolgerungen besitzen daher die unumschränkte Allgemeingültigkeit der klassischen Gesetze der Thermostatik. Sie sind darüberhinaus allgemeiner als diese, da sie klarlegen, daß die makroskopischen Parameter eines Systems von Natur aus statistisch sind und Schwankungen zeigen, die unter geeigneten Bedingungen berechenbar und wahrnehmbar sind. Trotz des mikroskopischen Ausgangspunktes enthält das Buch daher viele allgemeine Argumentationen auf einer rein makroskopischen Ebene – wahrscheinlich ungefähr ebensoviel wie ein Text über klassische Thermostatik – aber der mikroskopische Inhalt der makroskopischen Argumente bleibt in allen Abschnitten sichtbar. *Falls* man überdies gewillt ist, spezifische mikroskopische Modelle für die Teilchen eines Systems vorauszusetzen, dann ist es auch einleuchtend, wie man makroskopische Größen auf der Grundlage dieser mikroskopischen Information berechnen kann. Schließlich bilden die statistischen Begriffe, die gebraucht werden, um Gleichgewichtssituationen zu erörtern, eine angemessene Vorbereitung für die Behandlung der Nichtgleichgewichtssysteme.

Diese Art der Betrachtung hat sich in meiner eigenen Lehrpraxis als nicht schwieriger erwiesen als die übliche, die mit der klassischen Thermostatik beginnt. Die Thermostatik nach rein makroskopischen Grundsätzen zu entwickeln, ist begrifflich schwierig. Die zu benutzenden Beweisführungen sind oft bedenklich und von einer Art, die manchen Physikstudenten unnatürlich erscheint und die Bedeutung des Grundbegriffes der Entropie sehr schwer begreiflich macht. Ich habe es vorgezogen, auf die Spitzfindigkeiten traditioneller Argumentation, die auf ausgeklügelten Kreisprozessen beruht, zu verzichten und an ihre Stelle einige elementare statistische Begriffe zu setzen. Folgende Vorteile werden hierdurch erreicht:

- a) Statt viel Zeit aufzuwenden, verschiedenartige Argumente zu diskutieren, die auf der Existenz spezieller Wärmemaschinen beruhen, kann man die Studenten in einem frühen Stadium mit statistischen Methoden bekanntmachen, die von großer und in der ganzen Physik wiederkehrender Bedeutung sind.
- b) Die mikroskopische Art der Betrachtung gewährt bessere physikalische Einsicht in viele Erscheinungen und führt zu einem leichten Verständnis der Bedeutung der Entropie.
- c) Ein großer Teil der modernen Physik befaßt sich mit der Erklärung makroskopischer Erscheinungen durch mikroskopische Begriffe. Es erscheint daher nützlich, sich an eine Darstellung zu halten, die zu jeder Zeit die Wechselbeziehung zwischen den mikroskopischen und makroskopischen Stufen der Beschreibung betont. Die traditionelle Lehre behandelt Thermostatik und statistische Mechanik als verschiedene Gebiete und hat deshalb auch das Wissen der Studenten ebenso zergliedert und sie somit schlecht dafür vorbereitet, neuere Ideen wie z.B. Spintemperatur oder negative Temperatur als folgerichtig und natürlich zu akzeptieren.
- d) Da die vereinheitlichte Darstellung zweckmäßiger ist – sowohl begrifflich als auch in zeitlicher Hinsicht – erlaubt sie, mehr Material und modernere Themen zu diskutieren.

Der grundlegende Plan des Buches ist folgender: Das erste Kapitel führt in einige grundlegende wahrscheinlichkeitstheoretische Begriffe ein. Statistische Vorstellungen werden dann auf Teilchensysteme im Gleichgewicht angewandt, um die grundlegenden Begriffe der statistischen Mechanik zu entwickeln, und um davon ausgehend, die rein makroskopische allgemeine Darstellung der Thermostatik abzuleiten. Die *makroskopischen* Aspekte der Theorie werden dann ausführlich diskutiert und veranschaulicht; das gleiche wird dann für die *mikroskopischen* Aspekte der Theorie getan. Einige komplizierte Gleichgewichtssituationen – wie z.B. Phasenübergänge und Quantengase werden als nächstes behandelt. Danach werden Nichtgleichgewichtssituationen besprochen, und die Transporttheorie in verdünnten Gasen unter verschiedenen Aspekten erörtert. Schließlich befaßt sich das letzte Kapitel mit einigen allgemeinen Fragen der irreversiblen Prozesse und der Fluktuationstheorie. Mehrere Anhänge bringen einige nützliche mathematische Tatsachen.

Das Buch ist hauptsächlich als Einführungskurs in die statistische Physik und in die Theorie der Wärme für Studenten aller Semester gedacht. Das Buch beruht auf **vervielfältigten Aufzeichnungen**, die von mir und mehreren meiner Kollegen mehr als zwei Jahre lang für eine Vorlesungsreihe an der Universität von Kalifornien in Berkeley benutzt worden sind. Vorkenntnisse in Wärmelehre oder Thermostatik werden nicht vorausgesetzt; die notwendigen Vorbedingungen sind die Einführungsvorlesungen in Physik und Atomphysik. Die Vorlesung über Atomphysik soll nur die Voraussetzung für eine ausreichende Grundkenntnis in moderner Physik sichern:

- a) zu wissen, daß die Quantenmechanik Systeme mit den Begriffen „Quantenzustand“ und „Wellenfunktion“ beschreibt,
- b) die Energieniveaus eines harmonischen Oszillators und die quantenmechanische Beschreibung eines freien Teilchens in einem Kasten zu kennen, und
- c) von der Heisenbergschen Unbestimmtheitsrelation und dem Paulischen Ausschließungsprinzip gehört zu haben.

Das sind im wesentlichen alle Quantenbegriffe, die benötigt werden.

Das hier enthaltene Material ist umfangreicher als in einer einsemestrigen Anfängervorlesung behandelt werden kann. Dies wurde absichtlich getan, um

- a) jene grundlegenden Begriffe in die Behandlung einzuschließen, die dem Studierenden den Zugang zu vorgeschritteneren Arbeiten erleichtern,
- b) Studenten mit einigem Wissensdurst es zu ermöglichen, über das Minimum hinaus zu einem gegebenen Thema zu lesen,
- c) dem Dozenten einige Möglichkeiten zu bieten, zwischen alternativen Themen auszuwählen und
- d) um die augenblickliche Umgestaltung des Grundkurses Physik vorwegzunehmen, die die Studenten in höheren Semestern besser als jetzt befähigen soll, sich mit vorgeschrittenem Material zu befassen.

In der Praxis habe ich erfolgreich die ersten 12 Kapitel (Kapitel 10 und die meisten mit einem Sternchen versehenen Abschnitte auslassend) in einer einsemestrigen Vorlesung behandelt, Kapitel 1 enthält eine umfassendere Darstellung der Grundbegriffe der Wahrscheinlichkeitsrechnung, wie sie für das Verständnis der folgenden Kapitel gebraucht wird. Außerdem sind die Kapitel so angelegt, daß es leicht möglich ist, nach den ersten acht Kapiteln einige zugunsten anderer auszulassen, ohne Schwierigkeiten zu haben.

Das Buch sollte auch für den Gebrauch in einer einführenden Vorlesung für höhere Semester geeignet sein, wenn man die mit einem Sternchen versehenen Abschnitte und die drei letzten Abschnitte einschließt, die vorgeschrittenere Themen enthalten. Man kann in der Tat bei Studenten, die klassische Thermostatik studiert haben, aber in ihrem Anfangsstudium keine wesentliche Bekanntschaft mit den Begriffen der statistischen Mechanik gemacht hatten, nicht hoffen, in einer

einsemestrigen Vorlesung für Fortgeschrittene beträchtlich mehr Stoff zu behandeln als hier gebracht wird. Einer meiner Kollegen hat das Material deshalb in unserem Berkeley Kurs für Fortgeschrittene benutzt (eine Vorlesung, die bis jetzt meistens von Studenten mit der genannten Vorbildung besucht wird.)

Durch das ganze Buch hindurch habe ich versucht, die Schlußfolgerungen wohlbegründet zu halten und mich um Einfachheit der Darstellung bemüht. Es ist nicht mein Ziel gewesen, Genauigkeit im formalen mathematischen Sinn anzustreben. Ich habe dagegen versucht, die grundlegenden physikalischen Begriffe in den Vordergrund zu stellen und sie mit Sorgfalt zu erörtern. Im Verlauf des Schreibens ist das Buch länger geworden als es sonst hätte werden können, denn ich habe nicht gezögert, das Verhältnis der Worte zu den Formeln zu vergrößern, um erläuternde Beispiele zu geben oder verschiedene Arten der Problembetrachtung darzustellen, so oft ich meinte, daß dies das Verständnis vertiefen würde. Mein Ziel ist es gewesen, physikalische Einsicht und wichtige Methoden der Argumentation zu betonen, und ich rate sehr ernsthaft, daß der Student auf diese Aspekte des Themas Gewicht legen soll, anstatt zu versuchen, verschiedene in sich bedeutungslose Formeln auswendig zu lernen. Um zu vermeiden, daß sich der Leser in belanglose Details verliert, habe ich es unterlassen, den allgemeineren Fall eines Problems darzustellen und habe stattdessen versucht, relativ einfache Fälle durch wirksame und leicht zu verallgemeinernde Methoden zu behandeln. Das Buch soll nicht enzyklopädisch sein; es ist nur beabsichtigt, ein grundlegendes Gerüst aus einigen fundamentalen Begriffen zu geben, das geeignet ist, dem Studenten in seiner künftigen Arbeit zu helfen. Unnötig zu sagen, daß eine Auswahl getroffen werden mußte. Zum Beispiel hielt ich es für wichtig, die Boltzmann-Gleichung einzuführen, aber widerstand der Versuchung, die Anwendung der Onsagerschen Reziprozitätsbeziehungen auf verschiedene irreversible Phänomene wie z.B. thermoelektrische Effekte zu erörtern ⁺⁾ .

Es ist nützlich, wenn ein Leser Material von nebensächlicher Bedeutung von dem unterscheiden kann, welches wesentlich für den Hauptfaden der Beweisführung ist. Um auf den Stoff von untergeordneter Bedeutung hinzuweisen, ist folgendermaßen verfahren worden:

- a) Abschnitte, die durch ein Sternchen gekennzeichnet sind, enthalten vorgeschritteneres oder ausführlicheres Material; sie können ausgelassen werden (und sollten wahrscheinlich bei einer ersten Lesung übergangen werden), ohne daß beim Übergang zu den folgenden Abschnitten irgendeine Schwierigkeit auftritt.
- b) Viele Bemerkungen, Beispiele und Ausarbeitungen sind in den ganzen Text eingestreut und durch einen dünnen Strich am Rande hervorgehoben. Umgekehrt sind schwarze am Rand befindliche Hinweiszeichen benutzt worden, um wichtige Resultate hervorzuheben und die Bezugnahme auf diese zu erleichtern.

⁺⁾ Es gibt gute Gründe, dieser Versuchung zu erliegen (s. 15.19).

Das Buch enthält ungefähr 230 Aufgaben, die als ein wesentlicher Teil des Textes betrachtet werden sollten. Es ist unerlässlich, daß der Student einen beträchtlichen Anteil dieser Aufgaben löst, wenn er ein tieferes Verständnis des Stoffes erlangen will, und nicht nur eine beiläufige Kenntnis.

Ich bin mehreren meiner Kollegen für viele wertvolle Kritiken und Anregungen zu Dank verpflichtet. Im besonderen möchte ich Prof. Eyvind H. Wichmann danken, der eine ältere Version des ganzen Manuskriptes mit peinlich genauer Sorgfalt las, Prof. Owen Chamberlain, Prof. John J. Hopfield, Dr. Allan N. Kaufman und Dr. John M. Worlock. Unnötig zu sagen, daß keinem dieser Herren die Schuld für die Fehler der endgültigen Fassung gegeben werden sollte.

Lobende Anerkennung gebührt auch Herrn Roger F. Knacke für die Ausarbeitung der Lösungen zu den Aufgaben. Schließlich bin ich besonders meiner Sekretärin, Fräulein Beverly West dankbar, ohne deren Eifer und außerordentliche Geschicklichkeit, Seiten mit äußerst schlecht leserlicher Handschrift in ein perfekt getipptes Manuskript umzuwandeln, dieses Buch niemals hätte geschrieben werden können.

Es ist gesagt worden, daß „ein Autor niemals ein Buch beendet, er verläßt es nur.“ Ich bin dahingekommen, die Wahrheit dieser Feststellung klar zu würdigen, und ich fürchte den Tag zu erleben, an dem ich – wenn ich auf das gedruckte Manuskript sehe – erkennen muß, daß so vieles hätte besser getan und klarer erklärt werden können. Wenn ich das Buch dennoch verlasse, so ist es in der bescheidenen Hoffnung, daß es anderen trotz seiner Mängel nützlich sein möge.

F. Reif

Inhalt

1. Einführung in die statistische Methode	1
Zufallsbewegung und Binomialverteilung	6
1.1 Elementare statistische Begriffe und Beispiele	6
1.2 Das einfache Problem der eindimensionalen Zufallsbewegung	9
1.3 Mittelwerte	14
1.4 Berechnung von Mittelwerten für das Problem der Zufallsbewegung	17
1.5 Wahrscheinlichkeitsverteilung für großes N	21
1.6 Gaußsche Wahrscheinlichkeitsverteilungen	25
Allgemeine Diskussion der Zufallsbewegung	29
1.7 Wahrscheinlichkeitsverteilungen mit mehreren Variablen	29
1.8 Bemerkungen zu kontinuierlichen Wahrscheinlichkeitsverteilungen	31
1.9 Allgemeine Berechnung von Mittelwerten für die Zufallsbewegung	37
*1.10 Berechnung der Wahrscheinlichkeitsverteilung	41
1.11 Wahrscheinlichkeitsverteilung für großes N	44
Ergänzende Literatur	46
Aufgaben	47
2. Statistische Beschreibung von Vielteilchensystemen	55
Statistische Formulierung des mechanischen Problems	57
2.1 Beschreibung des Systemzustandes	57
2.2 Statistisches Ensemble	61
2.3 Grundlegende Postulate	63
2.4 Berechnung der Wahrscheinlichkeit	70
2.5 Zustandsdichte	71
Wechselwirkung zwischen makroskopischen Systemen	77
2.6 Thermische Wechselwirkung	77
2.7 Mechanische Wechselwirkung	79
2.8 Allgemeine Wechselwirkungen	84
2.9 Quasistatische Prozesse	85
2.10 Quasistatische Arbeit durch Druck	88
2.11 Exakte (vollständige) und „nichtexakte“ Differentiale	90
Ergänzende Literatur	95
Aufgaben	95
3. Statistische Thermodynamik	101
Irreversibilität und die Annäherung an das Gleichgewicht	102
3.1 Gleichgewichtsbedingungen und äußere Zwänge	102
3.2 Reversible und irreversible Prozesse	106

Thermische Wechselwirkung zwischen makroskopischen Systemen	110
3.3 Verteilung der Energie auf Systeme im Gleichgewicht	110
3.4 Die Annäherung an das thermische Gleichgewicht	116
3.5 Temperatur	118
3.6 Wärmereservoir	123
3.7 Das Maximum der Wahrscheinlichkeitsverteilung	124
Allgemeine Wechselwirkung zwischen makroskopischen Systemen	129
3.8 Abhängigkeit der Zustandsdichte von äußeren Parametern	129
3.9 Wechselwirkende Systeme im Gleichgewicht	131
3.10 Eigenschaften der Entropie	135
Zusammenstellung der grundlegenden Ergebnisse	140
3.11 Hauptsätze und fundamentale statistische Beziehungen	140
3.12 Statistische Berechnung thermodynamischer Größen	142
Ergänzende Literatur	145
Aufgaben	145
4. Makroskopische Parameter und ihre Messung	147
4.1 Arbeit und innere Energie	148
4.2 Wärme	151
4.3 Absolute Temperatur	153
4.4 Wärmekapazität und spezifische Wärme	159
4.5 Entropie	163
4.6 Konsequenzen der absoluten Entropiedefinition	166
4.7 Extensive und intensive Parameter	170
Ergänzende Literatur	171
Aufgaben	172
5. Einige Anwendungen der makroskopischen Thermostatik	175
Eigenschaften idealer Gase	177
5.1 Zustandsgleichung und innere Energie	177
5.2 Spezifische Wärmen	180
5.3 Adiabatische Expansion bzw. Kompression	183
5.4 Entropie	184
Allgemeine Beziehungen für ein homogenes System	185
5.5 Ableitung allgemeiner Beziehungen	185
5.6 Zusammenfassung der Maxwellschen Relationen und der thermo- dynamischen Potentiale	190
5.7 Spezifische Wärme	191
5.8 Entropie und innere Energie	198
Freie Expansion und Drosselexperimente	203
5.9 Freie Expansion eines Gases	203
5.10 Der Drossel- (oder Joule-Thomson-)Prozeß	207
Wärmemaschinen und Kältemaschinen	214
5.11 Wärmemaschinen	214

5.12 Kältemaschinen	220
Ergänzende Literatur	222
Aufgaben	223
6. Grundlegende Methoden und Ergebnisse der statistischen Mechanik	235
Repräsentative Ensemble für Systeme unter verschiedenen Nebenbedingungen	236
6.1 Abgeschlossenes System	236
6.2 System in Kontakt mit einem Wärmereservoir	237
6.3 Einfache Anwendungen der kanonischen Verteilung	241
6.4 System mit fester mittlerer Energie	246
6.5 Berechnungen von Mittelwerten im kanonischen Ensemble	248
6.6 Zusammenhang mit der Thermostatik	250
Näherungsmethoden	255
6.7 Ensembles als Näherungen	255
*6.8 Mathematische Näherungsmethoden	257
Verallgemeinerungen und andere Näherungen	263
*6.9 Großkanonisches und andere Ensemble	263
*6.10 Alternative Herleitung der kanonischen Verteilung	266
Phasenräume	270
+6.11 μ - und Γ -Raum	270
Ergänzende Literatur	272
Aufgaben	273
7. Einfache Anwendungen der statistischen Mechanik	279
Allgemeine Methoden	280
7.1 Zustandssummen und ihre Eigenschaften	280
Das ideale einatomige Gas	282
7.2 Berechnung thermodynamischer Größen	282
7.3 Das Gibbsche Paradoxon	286
7.4 Gültigkeit der klassischen Näherung	290
Der Gleichverteilungssatz	292
7.5 Beweis des Satzes	292
7.6 Einfache Anwendungen	295
7.7 Spezifische Wärmen von Festkörpern	298
Paramagnetismus	302
7.8 Allgemeine Berechnung der Magnetisierung	302
Kinetische Theorie verdünnter Gase im Gleichgewicht	309
7.9 Die Maxwellsche Geschwindigkeitsverteilung	309
7.10 Zugeordnete Geschwindigkeitsverteilungen und Mittelwerte	311
7.11 Anzahl der auf eine Oberfläche aufschlagenden Moleküle	317
7.12 Effusion	321
7.13 Druck- und Impulsübertragung	326

Ergänzende Literatur	330
Aufgaben	331
8. Gleichgewicht zwischen Phasen oder chemischen Verbindungen	339
Allgemeine Gleichgewichtsbedingungen	340
8.1 Isoliertes System	340
8.2 System in Kontakt mit einem Reservoir konstanter Temperatur	344
8.3 System konstanten Drucks in Kontakt mit einem Reservoir konstanter Temperatur	347
8.4 Stabilitätsbedingungen für eine homogene Substanz	349
Gleichgewicht zwischen Phasen	354
8.5 Gleichgewichtsbedingungen und die Clausius-Clapeyronsche Gleichung	354
8.6 Phasenübergänge und Zustandsgleichung	359
Systeme aus mehreren Komponenten; chemisches Gleichgewicht	366
8.7 Allgemeine Beziehungen für ein System aus mehreren Komponenten	366
8.8 Alternative Behandlung des Phasengleichgewichts	369
8.9 Allgemeine Bedingungen für chemisches Gleichgewicht	371
8.10 Chemisches Gleichgewicht zwischen idealen Gasen	374
Ergänzende Literatur	382
Aufgaben	382
9. Quantenstatistik idealer Gase	389
Maxwell-Boltzmann-, Bose-Einstein- und Fermi-Dirac-Statistik	390
9.1 Identische Teilchen und Symmetrie-Bedingungen	390
9.2 Formulierung des statistischen Problems	395
9.3 Die quantenmechanischen Verteilungsfunktionen	397
9.4 Maxwell-Boltzmann-Statistik	403
9.5 Photonen-Statistik	405
9.6 Bose-Einstein-Statistik	407
9.7 Fermi-Dirac-Statistik	411
9.8 Quantenstatistik im klassischen Grenzfall	412
Das ideale Gas im klassischen Grenzfall	415
9.9 Quantenzustände eines einzelnen Teilchens	415
9.10 Auswertung der Zustandssumme	423
9.11 Physikalische Folgerungen aus der quantenmechanischen Abzählung der Zustände	426
*9.12 Die Zustandssummen mehratomiger Moleküle	431
Die Strahlung des schwarzen Körpers	437
9.13 Elektromagnetische Hohlraumstrahlung im thermischen Gleichgewicht	437
9.14 Untersuchung der Strahlung in einem beliebigen Hohlraum	443
9.15 Die von einem Körper bei der Temperatur T emittierte Strahlung.	446

Leitungselektronen in Metallen	454
9.16 Folgerungen aus der Fermi-Dirac-Verteilung	454
*9.17 Quantitative Berechnung der spezifischen Wärme der Elektronen	460
Ergänzende Literatur	464
Aufgaben	464
10. Systeme wechselwirkender Teilchen	473
Festkörper	476
10.1 Gitter- und Normalschwingungen	476
10.2 Die Debyesche Näherung	482
Das nichtideale klassische Gas	489
10.3 Berechnung der Zustandssumme für geringe Dichten	489
10.4 Zustandsgleichung und Virialkoeffizienten	493
10.5 Eine andere Ableitung der van der Waals-Gleichung	497
Ferromagnetismus	499
10.6 Wechselwirkung zwischen Spins	499
10.7 Molekularfeld-Näherung von Weiß	502
Ergänzende Literatur	507
Aufgaben	508
11. Magnetismus und tiefe Temperaturen	511
11.1 Magnetische Arbeit	513
11.2 Magnetisches Kühlen	519
11.3 Messung sehr tiefer Temperaturen	528
11.4 Supraleitfähigkeit	532
Ergänzende Literatur	536
Aufgaben	537
12. Elementare kinetische Theorie der Transportvorgänge	541
12.1 Die Stoßzeit	543
12.2 Stoßzeit und Streuquerschnitt	549
12.3 Viskosität (dynamische Zähigkeit)	554
12.4 Wärmeleitfähigkeit	561
12.5 Diffusion	567
12.6 Elektrische Leitfähigkeit	573
Ergänzende Literatur	575
Aufgaben	575
13. Transporttheorie in der Relaxationszeit-Näherung	579
13.1 Transporterscheinungen und Verteilungsfunktionen	580
13.2 Die stoßfreie Boltzmann-Gleichung	585
⁺ 13.3 Die Bahnintegralmethode	590
13.4 Beispiel: Berechnung der elektrischen Leitfähigkeit	594
13.5 Beispiel: Berechnung der Viskosität	597

13.6 Boltzmann-Gleichung	599
13.7 Äquivalenz von Bahnintegral-Methode und Relaxationszeit-Ansatz	600
13.8 Beispiele zur Anwendung der Boltzmann-Gleichung	602
Ergänzende Literatur	604
Aufgaben	604
14. Die fast exakte Form der Transporttheorie	607
14.1 Zweierstöße	608
14.2 Streuquerschnitte und Symmetrieeigenschaften	612
14.3 Aufstellung der Boltzmann-Gleichung	615
14.4 Bilanzgleichungen für Mittelwerte	619
14.5 Erhaltungssätze und Hydrodynamik	623
14.6 Beispiel: Einfache Untersuchung der elektrischen Leitfähigkeit	626
14.7 Näherungsmethoden zur Lösung der Boltzmann-Gleichung	629
14.8 Beispiel: Berechnung der Viskosität	635
Ergänzende Literatur	642
Aufgaben	643
15. Irreversible Prozesse und Schwankungen	647
Übergangswahrscheinlichkeiten und Mastergleichung	648
15.1 Abgeschlossene Systeme	648
15.2 System in Kontakt mit einem Wärmereservoir	650
15.3 Magnetische Resonanz	654
15.4 Dynamische Kernpolarisation; Overhauser-Effekt	657
Einfache Erörterung der Brownschen Bewegung	661
15.5 Langevinsche Gleichung	661
15.6 Berechnung des Schwankungsquadrats der Verrückung	666
Genauere Untersuchung der Brownschen Bewegung	669
15.7 Beziehung zwischen Dissipation und Fluktuationskraft	669
15.8 Korrelationsfunktionen und Reibungskonstante	672
*15.9 Schwankungsquadrat der Geschwindigkeit	676
*15.10 Korrelationsfunktion der Geschwindigkeit und Schwankungsquadrat der Verrückung	678
Berechnung von Wahrscheinlichkeitsverteilungen	680
*15.11 Fokker-Planck-Gleichung	680
*15.12 Lösung der Fokker-Planck-Gleichung	684
Stochastische Prozesse	686
+15.13 Fourieranalyse	686
15.14 Ensemble- und Zeitmittelwerte	688
+15.15 Wiener-Chintschin-Theorem	690
15.16 Nyquist-Theorem	693
15.17 Nyquist-Theorem und Gleichgewichtsbedingungen	695
+15.17a Brownsche Bewegung als stochastischer Prozeß	700

Allgemeine Erörterung irreversibler Prozesse	702
15.18 Schwankungen und Onsagersche Reziprozitätsbeziehungen	702
⁺ 15.19 Skizze der thermodynamischen Theorien irreversibler Prozesse	709
Ergänzende Literatur	739
Aufgaben	741
 Anhang	 745
A.1 Elementare Summen	746
A.2 Auswertung des Integrals $\int_{-\infty}^{\infty} e^{-x^2} dx$	746
A.3 Auswertung des Integrals $\int_0^{\infty} e^{-x} x^n dx$	748
A.4 Auswertung von Integralen der Form $\int_0^{\infty} e^{-\alpha x^2} x^n dx$	749
A.5 Die Fehlerfunktion	750
A.6 Stirlingsche Formel	752
⁺ A.7 Diracsche Deltafunktion	756
A.8 Die Ungleichung $\ln x \leq x - 1$	762
A.9 Beziehungen zwischen partiellen Ableitungen mehrerer Variablen	763
A.10 Die Methode der Lagrangeschen Multiplikatoren	765
A.11 Berechnung des Integrals $\int_0^{\infty} (e^x - 1)^{-1} x^3 dx$	767
A.12 Das H-Theorem und die Annäherung an das Gleichgewicht	769
A.13 Das Liouvillesche Theorem der klassischen Mechanik	771
Bibliographie	775
Numerische Konstanten	781
Lösungen zu ausgewählten Aufgaben	782
⁺ Glossar	787
Internationales Einheitensystem (SI)	798
⁺ Symbolverzeichnis	805
Sachregister	811

1. Einführung in die statistische Methode

Für das eindimensionale Random-Walk-Problem (Irrflugproblem) ist die Wahrscheinlichkeit $W_N(n)$ dafür, daß nach N Schritten n Schritte in eine Richtung getan wurden, durch die Binomialverteilung $W_N(n) = [N!/n!(N-n)!]p^nq^{N-n}$ gegeben (p Wahrscheinlichkeit dafür, daß ein Schritt nach rechts ausgeführt wird, $q = 1 - p$).

Ihr Mittelwert ist $\bar{n} \equiv \sum_{n=0}^N W_N(n)n = Np$, ihr Schwankungsquadrat ist

$(\Delta n)^2 \equiv \overline{(n - \bar{n})^2} = Npq$. Die relative mittlere quadratische Abweichung beträgt für $p = q = \frac{1}{2}$ $\Delta^* n/\bar{n} \equiv \sqrt{(\Delta n)^2}/\bar{n} = 1/\sqrt{N}$. Für große Schrittzahlen N geht die Binomialverteilung in die diskrete Gaußsche Verteilung über

$W(n) = (2\pi Npq)^{-1/2} \exp [-(n - Np)^2/2Npq]$. Der Mittelwert $\overline{f(n)}$ einer Funktion f einer kontinuierlichen Zufallsveränderlichen n ist in einem Intervall Δn zu

$\overline{f(n)} \equiv \int_{\Delta n} \mathcal{P}(u) f(u) du$ definiert, wobei $\mathcal{P}(u)$ die Dichteverteilung der Veränderlichen u (Wahrscheinlichkeitsdichte) ist. $\mathcal{P}(u)du$ ist die Wahrscheinlichkeit dafür,

daß u im Intervall zwischen u und $u + du$ liegt. Bei Wechsel der Veränderlichen $u = u(\varphi)$ gilt für die Wahrscheinlichkeit dafür, daß φ zwischen φ und $\varphi + d\varphi$ liegt

$W(\varphi)d\varphi = \mathcal{P}(du)|du/d\varphi|d\varphi$. Für kleine Schrittlänge l und große Schrittzahl N geht die diskrete Gaußsche Verteilung in die stetige Gauß-Verteilung mit der Verteilungsdichte $\mathcal{P}(x)dx = (2\pi\sigma^2)^{-1/2} \exp [-(x - \mu)^2/2\sigma^2]dx$ über, wobei

$\mu \equiv (p - q)Nl$ der Mittelwert und $\sigma \equiv 2\sqrt{Npq}l$ die mittlere quadratische Abweichung der Verrückung ist.

Dieses Buch befaßt sich mit Systemen, die aus sehr vielen Teilchen bestehen. Beispiele sind Gase, Flüssigkeiten, feste Körper, elektromagnetische Strahlung (Photonen) usw. In der Tat bestehen die meisten physikalischen, chemischen oder biologischen Systeme aus vielen Teilchen; unser Thema schließt daher einen weiten Teil der Natur ein. Das Studium von Systemen, die aus vielen Teilchen bestehen, ist wahrscheinlich das am intensivsten bearbeitete Gebiet moderner physikalischer Forschung außerhalb des Bereichs der Hochenergiephysik. Auf dem letzteren Gebiet geht es darum, die fundamentalen Wechselwirkungen zwischen Nukleonen, Neutrinos, Mesonen oder anderen Elementarteilchen zu verstehen. Aber bei dem Versuch, feste Körper, Flüssigkeiten, Plasmen, chemische oder biologische Systeme und andere solcher Vielteilchensysteme zu beschreiben, steht man einer ziemlich schwierigen Aufgabe gegenüber, die nicht weniger anspruchsvoll ist. Hier gibt es gute Gründe für die Annahme, daß die bekannten Gesetze der Quantenmechanik die Bewegung der Atome und Moleküle dieser Systeme angemessen beschreiben; da außerdem die Atomkerne in gewöhnlichen chemischen oder biologischen Prozessen nicht gespalten werden und da Gravitationskräfte zwischen den Atomen dieser Systeme im wesentlichen allein von der elektromagnetischen Wechselwirkung verursacht. Jemand, der optimistisch genug ist, könnte daher versucht sein, zu behaupten, daß diese Systeme „im Prinzip verstanden“ werden. Dies wäre jedoch eine ziemlich leere und irreführende Feststellung: Denn trotz der Möglichkeit, die Bewegungsgleichungen für jedes dieser Systeme anzuschreiben, ist die Komplexität eines solchen Vielteilchensystems so groß, daß irgendwelche nützliche Folgerungen oder Vorhersagen aus den Bewegungsgleichungen abzuleiten, fast hoffnungslos erscheint. Die auftretenden Schwierigkeiten sind nicht nur Fragen quantitativen Details, die durch Anwendung größerer und besserer Computer gelöst werden können. Denn wenn auch die Wechselwirkungen zwischen individuellen Teilchen sogar ziemlich einfach sind, kann die Vielschichtigkeit, die aus der Wechselbeziehung einer großen Zahl von Teilchen hervorgeht, ganz unerwartete qualitative Merkmale im Verhalten eines Systems entstehen lassen. Es kann eine sehr gründliche Analyse erfordern, um das Auftreten dieser Merkmale aus der Kenntnis der individuellen Teilchen vorherzusagen. So ist es zum Beispiel eine auffallende und im mikroskopischen Detail schwierig zu verstehende Tatsache, daß einfache Atome, die ein Gas bilden, plötzlich kondensieren können, um eine Flüssigkeit mit sehr verschiedenen Eigenschaften zu bilden. Ebenso ist es eine unvorstellbar schwierige Aufgabe, zu verstehen, wie eine Ansammlung gewisser Molekülsorten zu einem System führen kann, das zur biologischen Entwicklung und Fortpflanzung fähig ist. Vielteilchensysteme sind also auch dann schwer zu durchschauen, wenn die Wechselwirkungen zwischen den individuellen Teilchen wohlbekannt sind, und dies ist nicht nur wegen der komplizierten Berechnungen ein Problem. Das Hauptziel besteht darin, aus den grundlegenden physikalischen Gesetzen neue Begriffe zu entwickeln, die die wesentlichen charakteristischen Merkmale solcher komplexen Systeme beleuchten und somit die Einsicht vermitteln, die es ermöglicht, wichtige Beziehungen zu erkennen und nützliche Voraussagen

zu machen. Wenn die betrachteten Probleme nicht zu komplex und wenn die gewünschte Genauigkeit der Beschreibung nicht zu detailliert ist, kann wirklich ein beträchtlicher Fortschritt durch relativ einfache analytische Methoden erreicht werden.

Die Systeme sollen der Größe nach unterschieden werden. Wir werden ein System „mikroskopisch“ nennen, wenn es annähernd von Atom-Dimensionen oder kleiner ist (etwa von der linearen Abmessung 10 \AA oder weniger). Zum Beispiel könnte das System ein Molekül sein. Andererseits werden wir ein System „makroskopisch“ nennen, wenn es groß genug ist, um im gewöhnlichen Sinne sichtbar zu sein (etwa größer als 10^{-3} mm , so daß es wenigstens mit einem gewöhnlichen Licht-Mikroskop beobachtet werden kann). Das System besteht dann aus sehr vielen Atomen oder Molekülen. Zum Beispiel könnte es ein fester Körper oder eine Flüssigkeit von der Art sein, der wir in unserer täglichen Erfahrung begegnen. Wenn man es mit einem solchen makroskopischen System zu tun hat, befaßt man sich im allgemeinen nicht mit dem detaillierten Verhalten jedes dieser individuellen Teilchen, die das System bilden. Statt dessen ist man gewöhnlich an gewissen makroskopischen Parametern interessiert, die das System als ganzes charakterisieren, z.B. Größen, wie Volumen, Druck, magnetisches Moment, Wärmeleitfähigkeit usw. Wenn alle makroskopischen Parameter eines isolierten Systems nicht mit der Zeit variieren, dann sagt man, daß das System im Gleichgewicht sei. Wenn ein isoliertes System sich nicht im Gleichgewicht befindet, werden sich die Parameter des Systems im allgemeinen verändern, bis sie konstante Werte erreichen, die den Gleichgewichtsbedingungen des Endzustandes entsprechen. Die Theorie dieser Gleichgewichtszustände wird zweifellos einer einfacheren theoretischen Diskussion zugänglich sein, als die der allgemeineren zeitabhängigen Nichtgleichgewichtszustände.

Im letzten Jahrhundert begann man zuerst, makroskopische Systeme (wie Gase, Flüssigkeiten oder feste Körper) von einem makroskopischen phänomenologischen Gesichtspunkt aus systematisch zu erforschen. Die so entdeckten Gesetze bildeten die „Thermodynamik“, die heute besser als „Thermostatik“ bezeichnet wird, weil sie eine Theorie der Gleichgewichtszustände darstellt. In der zweiten Hälfte des letzten Jahrhunderts erlangte die Theorie der Atomstruktur der Materie allgemeine Anerkennung, und man begann, makroskopische Systeme von einem fundamentalen mikroskopischen Standpunkt aus als Systeme zu analysieren, die aus sehr vielen Atomen oder Molekülen bestehen. Die Entwicklung der Quantenmechanik nach 1926 lieferte eine angemessene Theorie für die Beschreibung von Atomen und machte so den Weg für eine Analyse solcher Systeme auf der Grundlage realistischer mikroskopischer Begriffe frei. Außer den modernsten Methoden des „Vielteilchenproblems“ sind mehrere Disziplinen der Physik entwickelt worden, die sich mit Systemen befassen, die aus sehr vielen Teilchen bestehen. Obgleich die Grenzen zwischen diesen Gebieten nicht sehr scharf sind, ist es nützlich, kurz die Ähnlichkeiten und Unterschiede ihrer Betrachtungsweisen aufzuzeigen.

- a) Für ein System im Gleichgewicht kann man versuchen, einige sehr allgemeine Beziehungen aufzuzeigen, die zwischen den makroskopischen Parametern des Systems bestehen. Dies ist der Zugang zur klassischen „Thermostatik“, der historisch ältesten Disziplin. Die Stärke dieser Methode besteht in ihrer großen Allgemeingültigkeit, die es erlaubt, gültige Feststellungen zu treffen, die auf einer Mindestanzahl von Postulaten beruhen, ohne irgendwelche detaillierten Voraussetzungen über die mikroskopischen (d.h. molekularen) Eigenschaften des Systems zu benötigen. Diese Stärke der Methode impliziert auch ihre Schwäche: Nur relativ wenige Feststellungen können auf einer solchen allgemeinen Basis aufgestellt werden, und viele interessante Eigenschaften des Systems bleiben außerhalb des Rahmens dieser Methode.
- b) Für ein System im Gleichgewicht kann man außerdem versuchen, sehr allgemeine Feststellungen zu treffen, die jedoch durchweg auf den mikroskopischen Eigenschaften der Teilchen in dem System und auf den Gesetzen der Mechanik beruhen, die das Verhalten der Teilchen bestimmen. Dies ist die Betrachtungsweise der „Statistischen Mechanik“. Sie ergibt alle Resultate der Thermostatik zuzüglich einer großen Anzahl allgemeiner Relationen, um die makroskopischen Parameter des Systems aus einer Kenntnis seiner mikroskopischen Bestandteile zu berechnen. Diese Methode ist von großer Schönheit und Wirksamkeit.
- c) Wenn das System nicht im Gleichgewicht ist, steht man vor einer viel schwierigeren Aufgabe. Man kann immer noch versuchen, sehr allgemeine Feststellungen über solche Systeme zu erhalten, und dies führt zu den Methoden der „Irreversiblen Thermodynamik“, oder, allgemeiner, zum Studium der „Statistischen Mechanik irreversibler Prozesse“. Aber die Allgemeinheit dieser Methoden ist viel begrenzter als im Falle der Gleichgewichtssysteme.
- d) Man kann versuchen, die Wechselwirkungen aller Teilchen des Systems im Detail zu studieren und so Parameter von makroskopischer Bedeutung zu errechnen. Dies ist die Methode der „Kinetischen Theorie“. Sie ist im Prinzip immer anwendbar, sogar wenn das System nicht im Gleichgewicht ist, so daß die wirksamen Methoden der statistischen Mechanik des Gleichgewichtes nicht verwendet werden müssen. Weil die kinetische Theorie, kurz Kinetik genannt, aufgrund eines mikroskopischen Modells die detaillierteste Beschreibung liefert, ist sie auch die schwierigste Methode. Außerdem kann ihr detaillierter Standpunkt dazu führen, allgemeine Beziehungen, die vom speziellen mikroskopischen Modell unabhängig und von größerer Tragweite sind, zu verdecken. Die in den vier Punkten a) bis d) zusammengestellten Tatsachen, lassen sich in einer Tabelle darstellen ¹⁺⁾:

¹⁺⁾ Die Tabelle stammt aus der vom Bearbeiter gehaltenen Kursvorlesung „Thermostatik“.

Theorien makroskopischer Systeme unter Berücksichtigung ihrer Wärmebewegung

ohne Berücksichtigung ihres molekularen Aufbaus			mit Berücksichtigung ihres molekularen Aufbaus		
phänomenologische Theorien			statistische Theorien		
Thermodynamik			Kinetik	physikalische Statistik	
makroskopische Variable			Master-Gleichungen	repräsentative Ensembles	
Thermostatik	irreversible Thermodynamik	Nicht-klassische Thermodynamik	Transporttheorien	Statistik d. Gleichgewichtssysteme	Statistik irreversibler Prozesse

Historisch gesehen entstand die Thermostatik, bevor die atomare Struktur der Materie verstanden wurde. Auf die Vorstellung, daß Wärme eine Energieform sei, wurde zuerst durch die Arbeiten von Count Rumford (1798) und Davy (1799) hingewiesen. Die Idee, daß Wärme und Energie äquivalent seien, wurde ausführlich von dem deutschen Arzt R. J. Mayer (1842) dargelegt, aber sie fand erst nach den sorgfältigen experimentellen Arbeiten von Joule (1843–1849) Eingang in die Physik. Die ersten theoretischen Untersuchungen von Wärmekraftmaschinen wurden von dem französischen Ingenieur S. Carnot, 1824, durchgeführt. Die thermodynamische Theorie wurde in folgerichtiger Form von Clausius und Lord Kelvin um 1850 formuliert und von J.W. Gibbs in einigen grundlegenden Abhandlungen zur Vollendung entwickelt (1876–1878).

Die atomare Betrachtungsweise makroskopischer Systeme begann mit dem Studium der kinetischen Theorie verdünnter Gase. Diese Disziplin wurde durch die bahnbrechenden Arbeiten von Clausius, Maxwell und Boltzmann entwickelt. Maxwell entdeckte 1859 das Verteilungsgesetz für die Geschwindigkeiten der Moleküle, während Boltzmann 1872 seine fundamentale Integrodifferentialgleichung (die Boltzmannsche Gleichung) formulierte. Die kinetische Theorie der Gase erreichte ihre moderne Form, als es Chapman und Enskog (1916–1917) gelang, systematische Methoden zur Lösung dieser Gleichung zu entwickeln.

Die allgemeine Disziplin der Statistischen Mechanik entstand auch in den Arbeiten Boltzmanns, dem es ferner 1872 gelang, eine grundlegende mikroskopische Analyse der Irreversibilität und der Annäherung von Systemen an das Gleichgewicht durchzuführen. Die Theorie der Statistischen Mechanik wurde dann in voller Allgemeingültigkeit und Wirksamkeit durch die grundlegenden Beiträge von J.W. Gibbs (1902) entwickelt. Obwohl das Aufkommen der Quantenmechanik viele Änderungen gebracht hat, ist das grundlegende Gerüst der modernen Theorie immer noch das von Gibbs angegebene.

Bei der Diskussion von Vielteilchensystemen werden wir nicht danach streben, die historische Entwicklung der verschiedenen Disziplinen zu rekapitulieren: Statt dessen werden wir gleich von Anfang an einen modernen Standpunkt annehmen, der auf unseren heutigen Kenntnissen der Atomphysik und der Quantenmechanik beruht. Wir erwähnten bereits die Tatsache, daß man mit Hilfe ziemlich einfacher Methoden Vielteilchensysteme gut verstehen kann. Dies mag auf den ersten Blick ziemlich überraschend erscheinen; denn sind nicht Systeme wie Gase oder Flüssigkeiten, die eine Teilchenzahl von der Größe der Loschmidt'schen Zahl (10^{23}) besitzen, hoffnungslos kompliziert? Die Antwort ist, daß gerade die Komplexität dieser Systeme in sich den Schlüssel zu einer erfolgreichen Methode des Angriffs enthält. Da man sich nicht mit dem detaillierten Verhalten jedes einzelnen Teilchens in solchen Systemen befaßt, wird es möglich, statistische Methoden auf sie anzuwenden. Aber wie jeder Spieler, Versicherungsvertreter oder ein anderer, der sich mit Wahrscheinlichkeitsrechnung beschäftigt, weiß, werden statistische Beweise höchst zufriedenstellend, wenn sie auf große Zahlen angewandt werden können. Was für ein Vergnügen dann, sie in solchen Fällen anzuwenden, in denen die Zahlen so groß wie 10^{23} sind! In Systemen wie Gase, Flüssigkeiten oder festen Körpern, in denen man es mit sehr vielen identischen Teilchen zu tun hat, werden statistische Beweise wegen der für große Zahlen geltenden Grenzwertsätze²⁺⁾ besonders wirksam. Dies bedeutet nicht, daß alle Probleme verschwinden; die Physik der Vielteilchenprobleme läßt schon einige schwierige und faszinierende Fragen entstehen. Aber viele wichtige Probleme werden in der Tat ganz einfach, wenn man mit statistischen Hilfsmitteln an sie herangeht.

Zufallsbewegung³⁺⁾ und Binomialverteilung

1.1 Elementare statistische Begriffe und Beispiele

Die vorangehenden Bemerkungen machen offenbar, daß statistische Überlegungen in diesem Buch eine zentrale Rolle spielen werden. Wir wollen deshalb dieses erste Kapitel der Diskussion einiger elementarer Aspekte der Wahrscheinlichkeitstheorie widmen, die von großer Nützlichkeit sind.

Wir gehen davon aus, daß der Leser mit den grundlegenden Wahrscheinlichkeitsbegriffen vertraut ist. Es ist wesentlich, daran zu denken, daß immer, wenn eine Situation von einem statistischen Gesichtspunkt aus beschrieben werden soll, ein sogenanntes Ensemble betrachtet werden muß, wobei man unter Ensemble die Gesamtheit einer sehr großen Anzahl N von gleich präparierten Systemen versteht.

²⁺⁾ Diese Sätze machen Aussagen über Verteilungsfunktionen für eine Summe von Zufallsveränderlichen, wenn die Anzahl der Summanden sehr groß ist.

³⁺⁾ Engl.: random walk.

Die Wahrscheinlichkeit für das Eintreten eines bestimmten Ereignisses wird dann in bezug auf ein solches Ensemble definiert und ist durch den Bruchteil der Systeme in dem Ensemble gegeben, die durch das Eintreten dieses bestimmten Ereignisses charakterisiert sind. Beim Werfen eines Würfelpaares z.B. läßt sich eine statistische Beschreibung dadurch geben, daß man den Wurf einer sehr großen Anzahl \bar{N} (im Prinzip $\bar{N} \rightarrow \infty$) von gleichen Würfelpaaren unter gleichen Umständen betrachtet. (Alternativ dazu könnte man sich dasselbe Würfelpaar \bar{N} mal hintereinander unter gleichen Umständen geworfen denken.) Die Wahrscheinlichkeit, eine doppelte Eins zu erhalten, ist dann durch den Bruchteil derjenigen Würfe gegeben, bei denen eine doppelte Eins das Wurfresultat ist.

Man hat auch zu beachten, daß die Wahrscheinlichkeit sehr stark von der Natur des Ensembles abhängt, das bei der Definition dieser Wahrscheinlichkeit betrachtet wird. So hat es z.B. keinen Sinn, schlicht von der Wahrscheinlichkeit zu sprechen, daß aus einem einzelnen Samenkorn eine Pflanze mit roten Blüten wird. Aber man kann sinnvoll nach der Wahrscheinlichkeit fragen, daß aus einem Samenkorn, welches als Mitglied eines Ensembles gleicher, d.h. aus einer bestimmten Pflanzenmenge stammender Samenkörner betrachtet wird, eine Pflanze mit roten Blüten entsteht. Die Wahrscheinlichkeit hängt entscheidend von dem Ensemble ab, als dessen Mitglied das Samenkorn angesehen wird. So ist die Wahrscheinlichkeit, daß ein gegebenes Samenkorn rote Blüten hervorbringen wird, im allgemeinen davon abhängig, ob dieses Samenkorn betrachtet wird a) als Mitglied eines Ensembles von Samenkörnern, die von Pflanzen mit roten Blüten stammen, oder b) als Mitglied eines Ensembles von Samenkörnern, die von Pflanzen mit rosafarbenen Blüten herrühren.

Bei der folgenden Diskussion von Grundbegriffen der Wahrscheinlichkeitstheorie wird es nützlich sein, immer an ein besonders einfaches, aber wichtiges und illustratives Beispiel zu denken, nämlich an das sogenannte „Problem der Zufallsbewegung“. In seiner einfachsten Form kann das Problem in der folgenden üblichen Weise dargestellt werden: Ein Betrunkener startet von einem Laternenpfahl, der sich auf einer Straße befindet. Jeder seiner Schritte hat die gleiche Länge l . Der Mann, der sich nur längs der Bordsteinkante bewegen möge, ist jedoch so betrunken, daß die Richtung jedes einzelnen Schrittes – ob nach links oder nach rechts – vollständig unabhängig vom vorangehenden Schritt ist. Alles, was sich sagen läßt, ist, daß jedesmal, wenn der Mann einen Schritt macht, die Wahrscheinlichkeit für rechts p ist, während die Wahrscheinlichkeit für links $q = 1 - p$ ist. (Im einfachsten Fall gilt $p = q$, aber im allgemeinen wird $p \neq q$ sein; dann z.B., wenn die Straße bezüglich der Horizontalen geneigt ist, so daß ein Schritt bergab, d.h. nach rechts, wahrscheinlicher ist als ein Schritt bergauf, d.h. nach links.)

Wir legen die x -Achse in Richtung der Straße und wählen als Ursprung den Ort des Laternenpfahls. Da jeder einzelne Schritt von der Länge l ist, muß offenbar die Lokalisierung des Mannes von der Form $x = ml$ sein, wobei m eine ganze Zahl ist (positiv, negativ oder Null). Die Frage, die dann interessiert, ist die folgende: Wie groß ist die Wahrscheinlichkeit dafür, daß der Mann, nachdem er N

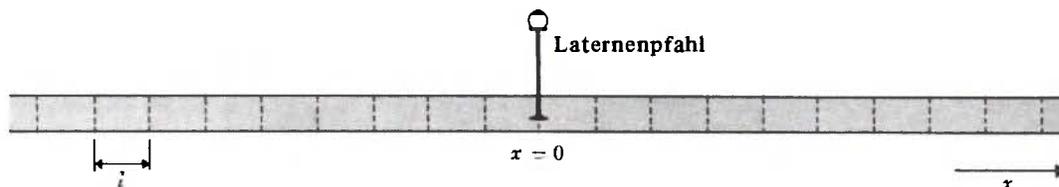


Abb. 1.1.1 Eindimensionale Zufallsbewegung eines Betrunkenen

gemacht hat, sich an der Stelle $x = ml$ befindet, wobei m eine vorgegebene ganze Zahl ist?

Die statistische Formulierung des Problems impliziert wiederum, daß man die Bewegung einer sehr großen Anzahl \bar{N} von gleich betrunkenen Männern betrachtet. (Falls die Situation sich zeitlich nicht ändert, wenn also z.B. der Mann nicht im Laufe der Zeit wieder nüchtern wird, könnte man auch alternativ dazu dasselbe Experiment \bar{N} mal mit demselben Mann wiederholen.) Bei jedem Schritt findet man, daß sich ein Bruchteil p der Männer nach rechts bewegt. Die Frage lautet dann, welcher Bruchteil der Männer sich nach N Schritten an der Stelle $x = ml$ befindet.

Dieses eindimensionale Problem läßt sich leicht auf mehrere Dimensionen, auf zwei, drei oder noch mehr, verallgemeinern. Man fragt dann wieder nach der

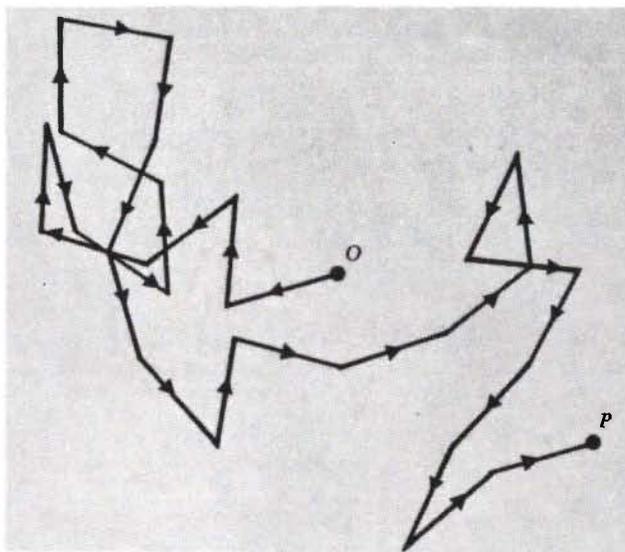


Abb. 1.1.2 Beispiel für eine zweidimensionale Zufallsbewegung

Wahrscheinlichkeit dafür, daß sich der Mann nach N Schritten in einer bestimmten Entfernung vom Ursprung befindet (obwohl nun die Entfernung nicht mehr von der Form $m \cdot l$ mit ganzzahligem m ist). Nun gilt das Hauptinteresse der Physik natürlich nicht Betrunkenen, die von irgendwelchen Laternenpfählen aus nach Hause torkeln. Aber das Problem, welches hierdurch veranschaulicht wird, ist im Grunde kein anderes als das des Addierens von N Vektoren gleicher Länge aber

mit zufälligen Richtungen (oder Richtungen, die einer bestimmten Wahrscheinlichkeitsverteilung unterworfen sind), wobei anschließend nach der Wahrscheinlichkeit dafür gefragt wird, daß der resultierende Vektor eine bestimmte Länge und eine bestimmte Richtung besitzt (siehe Abb. 1.1.2). Wir wollen ein paar physikalische Beispiele erwähnen, bei denen diese Frage von Bedeutung ist.

- a) Magnetismus: Ein Atom hat eine Spinquantenzahl $\frac{1}{2}$ und ein magnetisches Moment μ ; nach der Quantenmechanik kann sich deshalb sein Spin in bezug auf eine gegebene Richtung nur entweder „nach oben“ oder „nach unten“ einstellen. Wenn beide Möglichkeiten gleich wahrscheinlich sind, wie groß ist dann das zu erwartende Gesamtmoment von N solcher Atome?
- b) Diffusion eines Moleküls in einem Gas: Ein gegebenes Molekül legt im dreidimensionalen Raum zwischen den Zusammenstößen mit anderen Molekülen eine mittlere Entfernung l zurück. Welche Entfernung wird es nach N Zusammenstößen zurückgelegt haben?
- c) Intensität einer von N inkohärenten Lichtquellen herrührenden Strahlung: Die von jeder einzelnen Quelle ausgehende Lichtamplitude läßt sich durch einen zweidimensionalen Vektor darstellen, dessen Richtung die Phase der Erregung bestimmt. Hier sind die Phasen zufällig, so daß die resultierende Amplitude, welche die gesamte Intensität des von allen Quellen ausgehenden Lichtes bestimmt, mit statistischen Mitteln berechnet werden muß.

Der Vorgang der Zufallsbewegung hilft einige grundlegende Ergebnisse der Wahrscheinlichkeitstheorie zu verdeutlichen. Die Techniken, die beim Studium dieses Problems angewandt werden, sind nicht nur wirkungsvoll und grundlegend sondern tauchen auch in der statistischen Physik immer und immer wieder auf, so daß ein gründliches Verständnis dieses Problems außerordentlich gewinnbringend ist.

1.2 Das einfache Problem der eindimensionalen Zufallsbewegung

Der Einfachheit halber werden wir das Problem der Zufallsbewegung in einer Dimension diskutieren. Allerdings, anstatt von einem Betrunknen zu sprechen, wollen wir uns dem weniger alkoholischen Vokabular der Physik zuwenden und an ein Teilchen denken, das in einer Dimension aufeinanderfolgende Verschiebungen erfährt. Nach insgesamt N solcher Verschiebungen – jede einzelne hat die Länge l – befindet sich das Teilchen an der Stelle

$$x = ml$$

wobei m ganzzahlig und

$$-N \leq m \leq N$$

ist. Wir wollen nun die Wahrscheinlichkeit $P_N(m)$ dafür berechnen, daß sich das Teilchen nach N solcher Verschiebungen an der Stelle $x = ml$ befindet.

n_1 bezeichne die Anzahl der Verschiebungen nach rechts und n_2 die entsprechende Anzahl von Verschiebungen nach links. Die Gesamtzahl N von Verschiebungen ist dann natürlich einfach

$$N = n_1 + n_2 \quad (1.2.1)$$

Die resultierende Verschiebung (positiv gemessen nach rechts in Einheiten von l) ist gegeben durch

$$m = n_1 - n_2 \quad (1.2.2)$$

Falls bekannt ist, daß das Teilchen in einer Folge von N Einzelverschiebungen n_1 Verschiebungen nach rechts erfahren hat, dann ist damit auch die resultierende Verschiebung vom Ursprung bestimmt. In der Tat ergeben die vorangehenden Beziehungen unmittelbar

$$m = n_1 - n_2 = n_1 - (N - n_1) = 2n_1 - N \quad (1.2.3)$$

Das zeigt, daß die möglichen Werte von m ungerade sind, falls N ungerade ist, und daß sie gerade sind, falls N gerade ist.

Unsere fundamentale Annahme bestand darin, daß aufeinanderfolgende Verschiebungen statistisch unabhängig sind. Somit kann man behaupten, daß jede einzelne Verschiebung, ungeachtet dessen, was vorher geschah, durch die jeweiligen Wahrscheinlichkeiten

$$\begin{aligned} p &= \text{Wahrscheinlichkeit für „nach rechts“} \\ q &= 1 - p = \text{Wahrscheinlichkeit für „nach links“} \end{aligned}$$

charakterisiert ist.

Nun ist die Wahrscheinlichkeit für eine bestimmte Folge mit n_1 Verschiebungen nach rechts und n_2 Verschiebungen nach links einfach durch das Produkt der jeweiligen Wahrscheinlichkeiten, d.h. durch

$$\underbrace{p p \cdots p}_{n_1 \text{ Faktoren}} \underbrace{q q \cdots q}_{n_2 \text{ Faktoren}} = p^{n_1} q^{n_2} \quad (1.2.4)$$

gegeben. Aber es gibt natürlich viele verschiedene Möglichkeiten, N Verschiebungen so durchzuführen, daß n_1 von ihnen nach rechts und n_2 nach links gerichtet sind (siehe Abb. 1.2.1), und zwar ist die Anzahl der verschiedenen Möglichkeiten durch

$$\frac{N!}{n_1! n_2!} \quad (1.2.5)$$

gegeben. Folglich erhält man die Wahrscheinlichkeit $W_N(n_1)$ dafür, daß das Teilchen von insgesamt N Verschiebungen n_1 nach rechts und n_2 nach links erfährt ⁴⁺⁾, durch Multiplikation von (1.2.4) und (1.2.5), d.h. es gilt

⁴⁺⁾ in irgendeiner Reihenfolge!

►
$$W_N(n_1) = \frac{N!}{n_1!n_2!} p^{n_1}q^{n_2} \tag{1.2.6}$$

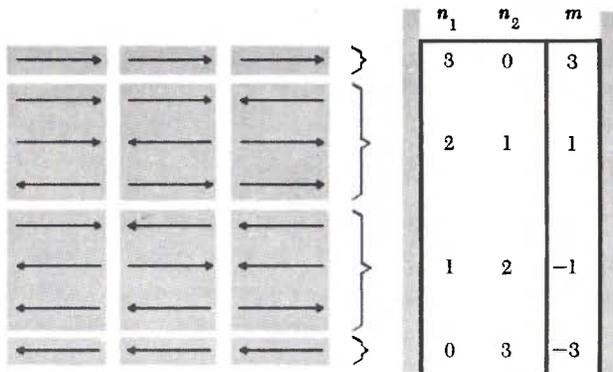


Abb. 1.2.1 Veranschaulichung der acht Folgen möglicher Verschiebungen für den Fall $N = 3$.

Einfaches Beispiel: Man betrachte das in Abb. 1.2.1 dargestellte Beispiel, das den Fall $N = 3$ behandelt. Es gibt nur eine Möglichkeit, daß alle drei aufeinanderfolgende Verschiebungen nach rechts gerichtet sind; die entsprechende Wahrscheinlichkeit $W(3)$, daß alle drei Verschiebungen nach rechts verlaufen, ist dann einfach $p \cdot p \cdot p = p^3$. Andererseits ist die Wahrscheinlichkeit für eine bestimmte Folge von Verschiebungen, bei der zwei nach rechts und eine nach links gerichtet sind, durch p^2q gegeben. Aber es gibt insgesamt drei mögliche Folgen solcher Verschiebungen, so daß die Wahrscheinlichkeit für irgendeine Folge mit zwei Verschiebungen nach rechts und einer nach links durch $3p^2q$ gegeben ist.

Begründung von Gleichung (1.2.5): Das Problem besteht darin, herauszufinden, auf wieviel verschiedene Weisen N Objekte, von denen n_1 einem Typ (1) und n_2 einem zweiten Typ (2) angehören ⁵⁺⁾, auf insgesamt $N = n_1 + n_2$ Plätze verteilt werden können. In unserem Fall kann

- der 1. Platz durch irgendeines der N Objekte
- der 2. Platz durch irgendeines der verbleibenden $(N - 1)$ Objekte
- ...
- der N . Platz durch das eine letzte Objekt besetzt werden.

Folglich können die verfügbaren Plätze auf insgesamt

$$N(N - 1)(N - 2) \cdots 1 \equiv N!$$

mögliche Weisen besetzt werden. Bei der obigen Abzählung wird aber jedes einzelne Objekt als unterscheidbar angesehen. Da aber die n_1 Objekte des ersten Types ununterscheidbar sind (z.B. sind alle n_1 Objekte Verschiebungen nach rechts), ergeben alle $n_1!$ Permutationen dieser Objekte nichts Neues. Genauso führen die $n_2!$ Permutationen der n_2 Objekte des zweiten Types stets wieder zu derselben Situation. Folglich erhält man die Anzahl der verschie-

⁵⁺⁾ Objekte, die einem Typ angehören, sollen ununterscheidbar sein.

denen Anordnungsmöglichkeiten von N Objekten, wenn n_1 einem Typ und n_2 einem zweiten Typ angehören, indem man die Gesamtzahl $N!$ der verschiedenen Permutationen der N Objekte durch die Anzahl $n_1!n_2!$ der nichts Neues liefernden Permutationen der Objekte jedes einzelnen Typs dividiert.

Beispiel: In dem vorangehenden Beispiel mit den drei Verschiebungen gibt es $N = 3$ mögliche Ereignisse (oder Plätze), die in Abb. 1.2.2 mit B_1, B_2, B_3 bezeichnet sind und die durch die drei Verschiebungen A_1, A_2, A_3 erfüllt

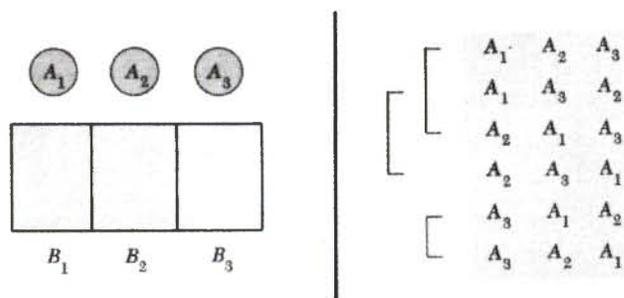


Abb. 1.2.2 Diagramm, das die Verteilung von drei Objekten A_1, A_2, A_3 auf drei Plätze B_1, B_2, B_3 verdeutlicht. Im rechten Teil sind die möglichen Anordnungen aufgeführt, wobei die eckigen Klammern diejenigen Anordnungen angeben, die identisch sind, wenn A_1 und A_2 als ununterscheidbar angesehen werden.

(realisiert) werden können. Das Ereignis B_1 kann auf irgendeine von drei Weisen, das Ereignis B_2 auf irgendeine von zwei Weisen und das Ereignis B_3 auf nur eine Weise eintreten. Somit gibt es $3 \cdot 2 \cdot 1 = 3! = 6$ mögliche Folgen dieser drei Verschiebungen. Aber nehmen wir an, daß es sich sowohl bei A_1 als auch bei A_2 um eine Verschiebung nach rechts handelt ($n_2 = 2$), während A_3 eine Verschiebung nach links darstellt ($n_2 = 1$). In diesem Fall sind Verschiebungsfolgen, die sich nur durch zwei unterschiedliche Anordnungen von A_1 und A_2 unterscheiden, in Wirklichkeit identisch. Somit verbleiben $6/2 = 3$ verschiedene Folgen mit zwei Verschiebungen nach rechts und einer nach links.

Die Wahrscheinlichkeitsfunktion (1.2.6) heißt Binomialverteilung. Der Name rührt daher, daß (1.2.6) ein typisches Glied in der als Binomialsatz bekannten Entwicklung von $(p + q)^N$ darstellt.

$$(p + q)^N = \sum_{n=0}^N \frac{N!}{n!(N-n)!} p^n q^{N-n} \quad (1.2.7)$$

Wir haben schon in (1.2.3) dargelegt, daß die resultierende Verschiebung m vom Ursprung bestimmt ist, falls man weiß, daß das Teilchen von insgesamt N Verschiebungen n_1 Verschiebungen nach rechts erfahren hat. Somit ist die Wahrscheinlichkeit $P_N(m)$ dafür, daß sich das Teilchen nach N Verschiebungen an der Stelle $x = ml$ befindet, durch $W_N(n_1)$ in (1.2.6) gegeben, also

$$P_N(m) = W_N(n_1) \quad (1.2.8)$$

Mit Hilfe von (1.2.1) und (1.2.2) findet man explizit ^{6*)}

$$n_1 = \frac{1}{2}(N + m), \quad n_2 = \frac{1}{2}(N - m) \quad (1.2.9)$$

Einsetzen dieser Relationen in (1.2.6) ergibt

$$P_N(m) = \frac{N!}{[(N + m)/2]![(N - m)/2]!} p^{(N+m)/2} (1 - p)^{(N-m)/2} \quad (1.2.10)$$

Ist speziell $p = q = \frac{1}{2}$, so nimmt (1.2.10) die symmetrische Form

$$P_N(m) = \frac{N!}{[(N + m)/2]![(N - m)/2]!} \left(\frac{1}{2}\right)^N$$

an.

Beispiele: Es sei $p = q = \frac{1}{2}$ und $N = 3$, wie in Abb. 1.2.1 dargestellt. Dann gibt es für die Anzahl der Verschiebungen nach rechts die Möglichkeiten $n_1 = 0, 1, 2, 3$; die entsprechenden Gesamtverschiebungen sind $m = -3, -1, 1, 3$; die entsprechenden Wahrscheinlichkeiten sind (wie aus Abb. 1.2.1 ersichtlich)

$$W_3(n_1) = P_3(m) = \frac{1}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{8} \quad (1.2.11)$$

Abb. 1.2.3 veranschaulicht die Binomialverteilung für den Fall, daß $p = q = \frac{1}{2}$ und die Gesamtzahl der Verschiebungen $N = 20$ ist. Die Einhüllende dieser diskreten Werte von $P_N(m)$ ist eine Glockenkurve. Was das physikalisch bedeutet, ist offensichtlich: Nach N richtungsmäßig zufälligen Verschiebungen der Länge l ist die Wahrscheinlichkeit dafür, daß sich das Teilchen in einer Entfernung $x = Nl$ vom Ursprung befindet, nur sehr gering. Dagegen ist die Wahrscheinlichkeit dafür, daß es sich in der Umgebung des Ursprungs befindet, am größten.

^{6*)} Beachte, daß nach (1.2.3) $(N + m)$ und $(N - m)$ geradzahlig sind, da sie gleich $2n_1$ bzw. gleich $2n_2$.

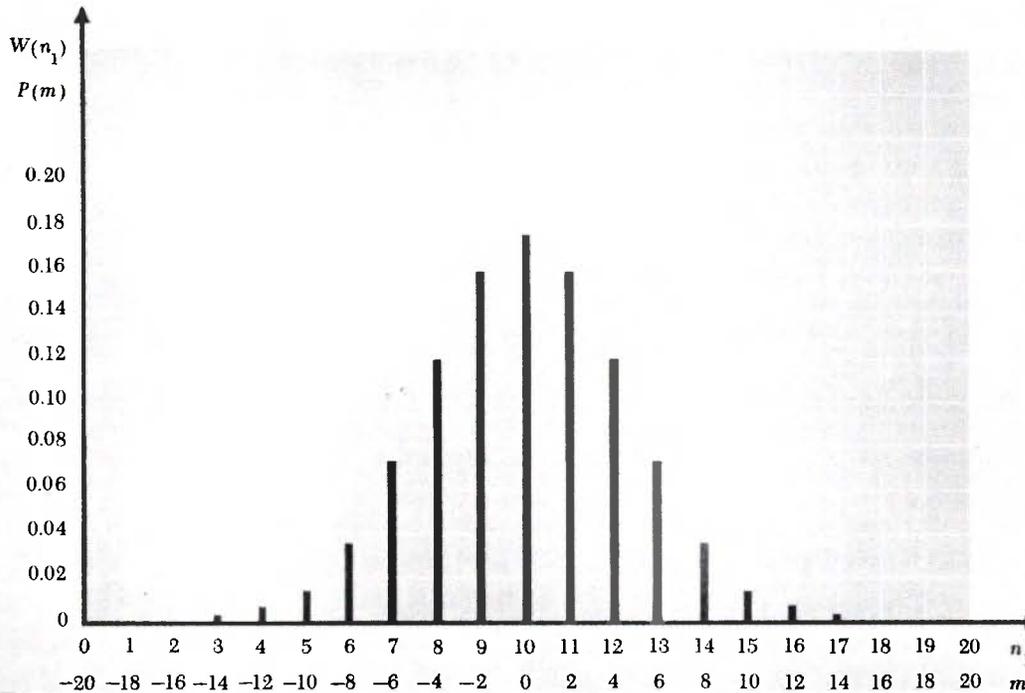


Abb. 1.2.3 Die Binomialverteilung für $p = q = \frac{1}{2}$ und $N = 20$. Das Diagramm zeigt die Wahrscheinlichkeit $W_N(n_1)$ für n_1 Rechtsverschiebungen bzw. die Wahrscheinlichkeit $P_N(m)$ für eine resultierende Verschiebung um m Einheiten nach rechts

1.3 Mittelwerte

Es sei u eine Variable, die die M diskreten Werte

$$u_1, u_2, \dots, u_M$$

mit den entsprechenden Wahrscheinlichkeiten

$$P(u_1), P(u_2), \dots, P(u_M)$$

annehmen kann. Der Mittelwert (oder Durchschnittswert) von u wird mit \bar{u} bezeichnet und ist durch

$$\bar{u} \equiv \frac{P(u_1)u_1 + P(u_2)u_2 + \dots + P(u_M)u_M}{P(u_1) + P(u_2) + \dots + P(u_M)}$$

definiert, oder kurz

$$\bar{u} \equiv \frac{\sum_{i=1}^M P(u_i)u_i}{\sum_{i=1}^M P(u_i)} \quad (1.3.1)$$

Das ist natürlich der übliche Weg, Durchschnitte zu berechnen. Ist z.B. u die Zensur eines Studenten bei einer Prüfung und $P(u)$ die Anzahl von Studenten, die diese Zensur erhalten, so berechnet sich nach (1.3.1) die Durchschnittsnote dadurch, daß man jede einzelne Zensur mit der zugehörigen Anzahl von Studenten multipliziert, das alles aufsummiert und anschließend durch die Gesamtzahl der Studenten dividiert.

Ist allgemein $f(u)$ irgendeine Funktion von u , so ist der Mittelwert von $f(u)$ durch

$$\overline{f(u)} \equiv \frac{\sum_{i=1}^M P(u_i) f(u_i)}{\sum_{i=1}^M P(u_i)} \quad (1.3.2)$$

definiert.

Dieser Ausdruck läßt sich vereinfachen. Da $P(u_i)$ als eine Wahrscheinlichkeit definiert ist, stellt die Größe

$$P(u_1) + P(u_2) + \cdots + P(u_M) \equiv \sum_{i=1}^M P(u_i)$$

die Wahrscheinlichkeit dafür dar, daß u irgendeinen seiner möglichen Werte annimmt, und die muß gleich eins sein. Folglich gilt ganz allgemein

$$\blacktriangleright \sum_{i=1}^M P(u_i) = 1 \quad (1.3.3)$$

Dies ist die sogenannte „Normierungsbedingung“, die von jeder Wahrscheinlichkeitsverteilung erfüllt wird. Damit wird aus der allgemeinen Definition (1.3.2)

$$\blacktriangleright \overline{f(u)} \equiv \sum_{i=1}^M P(u_i) f(u_i) \quad (1.3.4)$$

Man beachte die folgenden einfachen Ergebnisse. Sind $f(u)$ und $g(u)$ irgendzwei Funktionen von u , so gilt

$$\overline{f(u) + g(u)} = \sum_{i=1}^M P(u_i) [f(u_i) + g(u_i)] = \sum_{i=1}^M P(u_i) f(u_i) + \sum_{i=1}^M P(u_i) g(u_i)$$

oder

$$\blacktriangleright \overline{f(u) + g(u)} = \overline{f(u)} + \overline{g(u)} \quad (1.3.5)$$

Ist ferner c irgendeine Konstante, so gilt offensichtlich

$$\blacktriangleright \overline{cf(u)} = c \overline{f(u)} \quad (1.3.6)$$

Einige einfache Mittelwertbildungen sind für die Beschreibung von charakteristischen Merkmalen der Wahrscheinlichkeitsverteilung P besonders nützlich. Da ist zunächst der Mittelwert \bar{u} (z.B. die Durchschnittsnote in einer Gruppe von Studen-

ten). Dieser ist ein Maß für den zentralen Wert von u , um den herum die verschiedenen Werte u_i verteilt sind. Mißt man die Werte von u von ihrem Mittelwert \bar{u} aus, d.h. setzt man

$$\Delta u \equiv u - \bar{u} \quad (1.3.7)$$

so ist

$$\overline{\Delta u} = \overline{(u - \bar{u})} = \bar{u} - \bar{u} = 0 \quad (1.3.8)$$

Diese Gleichung besagt, daß der Mittelwert der Abweichung vom Mittel verschwindet.

Ein weiterer wichtiger Mittelwert ist

$$\overline{(\Delta u)^2} \equiv \sum_{i=1}^M P(u_i) (u_i - \bar{u})^2 \geq 0 \quad (1.3.9)$$

den man das „Schwankungsquadrat von u “ (oder die „Dispersion von u “) nennt. Diese Größe kann niemals negativ sein, da jeder Term in der Summe einen nicht-negativen Beitrag liefert. Nur falls $u_i = \bar{u}$ für alle Werte u_i gilt, verschwindet das Schwankungsquadrat. Je breiter die Verteilung der Werte u_i um den Wert \bar{u} herum ist, desto größer ist das Schwankungsquadrat. Das Schwankungsquadrat ist somit ein Maß für die Streuung der von u angenommenen Werte um den Mittelwert \bar{u} . Man merke sich die folgende allgemeine Beziehung, die oft bei der Berechnung des Schwankungsquadrates von Nutzen ist:

$$\overline{(u - \bar{u})^2} = \overline{(u^2 - 2u\bar{u} + \bar{u}^2)} = \overline{u^2} - 2\bar{u}\bar{u} + \bar{u}^2$$

oder

$$\blacktriangleright \quad \overline{(u - \bar{u})^2} = \overline{u^2} - \bar{u}^2 \quad (1.3.10)$$

Da die linke Seite positiv sein muß, folgt

$$\overline{u^2} \geq \bar{u}^2 \quad (1.3.11)$$

Man kann noch weitere Mittelwerte wie $\overline{(\Delta u)^n}$, $n > 2$, definieren. Diese sind jedoch wenig gebräuchlich.

Man beachte, daß die Kenntnis von $P(u)$ die vollständige Information über die tatsächliche Verteilung der Werte von u bedeutet. Die Kenntnis einiger Mittelwerte, wie \bar{u} und $\overline{(\Delta u)^2}$, bedeutet dagegen nur eine teilweise, wenn auch nützliche Information über die Charakteristik dieser Verteilung. Die Kenntnis von einigen Mittelwerten reicht nicht aus, um $P(u)$ vollständig zu bestimmen (sondern nur die Kenntnis der Werte $\overline{(\Delta u)^n}$ für alle Werte von n). Aus demselben Grund ist eine Berechnung der Wahrscheinlichkeitsverteilung $P(n)$ oft ziemlich schwierig, während sich einige einfache Mittelwerte leicht direkt ohne explizite Kenntnis von $P(u)$ berechnen lassen. Wir werden im folgenden auf einige dieser Anmerkungen noch ausführlicher eingehen.

1.4 Berechnung von Mittelwerten für das Problem der Zufallsbewegung

In (1.2.6) fanden wir, daß die Wahrscheinlichkeit, von insgesamt N Verschiebungen n_1 Verschiebungen nach rechts (und $N - n_1 \equiv n_2$ Verschiebungen nach links) zu erfahren, gegeben ist durch

$$W(n_1) = \frac{N!}{n_1!(N - n_1)!} p^{n_1} q^{N-n_1} \quad (1.4.1)$$

(Der Einfachheit halber lassen wir bei W den Index N weg, solange klar ist, was wir meinen.)

Zunächst wollen wir die Normierungsbedingung

$$\sum_{n_1=0}^N W(n_1) = 1 \quad (1.4.2)$$

verifizieren, welche besagt, daß für das Teilchen die Wahrscheinlichkeit, irgendeine Anzahl von Rechtsverschiebungen zwischen 0 und N zu erfahren, gleich eins sein muß. Durch Einsetzen von (1.4.1) in die linke Seite von (1.4.2) erhalten wir mit Hilfe des Binomialsatzes

$$\begin{aligned} \sum_{n_1=0}^N \frac{N!}{n_1!(N - n_1)!} p^{n_1} q^{N-n_1} &= (p + q)^N \\ &= 1^N = 1 \quad \text{wegen } q \equiv 1 - p \end{aligned}$$

wodurch (1.4.2) verifiziert ist.

Wie groß ist die mittlere Zahl \bar{n}_1 von Rechtsverschiebungen? Nach Definition gilt

$$\bar{n}_1 \equiv \sum_{n_1=0}^N W(n_1) n_1 = \sum_{n_1=0}^N \frac{N!}{n_1!(N - n_1)!} p^{n_1} q^{N-n_1} n_1 \quad (1.4.3)$$

Wenn nicht in jedem Term der letzten Summe der zusätzliche Faktor n_1 auftreten würde, wäre das wieder die Binomialentwicklung, und die Summation wäre kein Problem. Nun gibt es ein sehr praktisches, allgemeines Verfahren, einen solchen zusätzlichen Faktor zu behandeln, und damit die Summe in eine einfachere Form zu bringen. Dazu wollen wir uns mit dem rein mathematischen Problem der Auswertung der Summe in (1.4.3) befassen und p und q als irgendzwei *beliebige unabhängige* Parameter ansehen. Man bemerkt dann, daß sich der zusätzliche Faktor n_1 durch Differentiation erzeugen läßt, so daß

$$n_1 p^{n_1} = p \frac{\partial}{\partial p} (p^{n_1})$$

Folglich läßt sich die Summe in der Form schreiben

$$\begin{aligned}
 \sum_{n_1=0}^N \frac{N!}{n_1!(N-n_1)!} p^{n_1} q^{N-n_1} n_1 &= \sum_{n_1=0}^N \frac{N!}{n_1!(N-n_1)!} \left[p \frac{\partial}{\partial p} (p^{n_1}) \right] q^{N-n_1} \\
 &= p \frac{\partial}{\partial p} \left[\sum_{n_1=0}^N \frac{N!}{n_1!(N-n_1)!} p^{n_1} q^{N-n_1} \right] \quad \text{Durch Vertauschen von} \\
 & \quad \text{Summation und Diffe-} \\
 & \quad \text{rentation} \\
 &= p \frac{\partial}{\partial p} (p+q)^N \quad \text{mit dem Binomialsatz} \\
 &= pN(p+q)^{N-1}
 \end{aligned}$$

Da dieses Resultat für beliebige Werte von p und q gilt, muß es auch in dem uns speziell interessierenden Fall richtig sein, wo p eine bestimmte Konstante ist und $q \equiv 1-p$. Dann ist $p+q=1$, so daß (1.4.3) übergeht in

$$\blacktriangleright \quad \bar{n}_1 = Np \quad (1.4.4)$$

Wir hätten dieses Resultat erraten können. Da p die Wahrscheinlichkeit für eine Rechtsverschiebung ist, ist die mittlere Anzahl von Rechtsverschiebungen in einer Gesamtheit von N Verschiebungen einfach gegeben durch $N \cdot p$. Für die mittlere Anzahl von Linksverschiebungen ergibt sich entsprechend

$$\bar{n}_2 = Nq \quad (1.4.5)$$

wobei natürlich

$$\bar{n}_1 + \bar{n}_2 = N(p+q) = N$$

Die Gesamtverschiebung (positiv gemessen nach rechts in Einheiten von l) ist $m = n_1 - n_2$. Folglich erhalten wir für die mittlere Gesamtverschiebung

$$\blacktriangleright \quad \bar{m} = \overline{n_1 - n_2} = \bar{n}_1 - \bar{n}_2 = N(p - q) \quad (1.4.6)$$

Ist $p=q$, so gilt $\bar{m} = 0$. Dies muß auch so sein, da dann vollständige Symmetrie bezüglich links und rechts vorliegt.

Berechnung des Schwankungsquadrates $\overline{(\Delta n_1)^2}$: Nach (1.3.10) gilt

$$\overline{(\Delta n_1)^2} \equiv \overline{(n_1 - \bar{n}_1)^2} = \overline{n_1^2} - \bar{n}_1^2 \quad (1.4.7)$$

Wir kennen bereits \bar{n}_1 . Somit müssen wir noch $\overline{n_1^2}$ berechnen:

$$\begin{aligned}
 \overline{n_1^2} &\equiv \sum_{n_1=0}^N W(n_1) n_1^2 \\
 &= \sum_{n_1=0}^N \frac{N!}{n_1!(N-n_1)!} p^{n_1} q^{N-n_1} n_1^2
 \end{aligned} \quad (1.4.8)$$

Indem man p und q als unabhängige Parameter ansieht und denselben Differentiationstrick wie vorher benutzt, kann man schreiben

$$n_1^2 p^{n_1} = n_1 \left(p \frac{\partial}{\partial p} \right) (p^{n_1}) = \left(p \frac{\partial}{\partial p} \right)^2 (p^{n_1})$$

Somit läßt sich die Summe in (1.4.8) schreiben als

$$\begin{aligned} \sum_{n_1=0}^N \frac{N!}{n_1!(N-n_1)!} \left(p \frac{\partial}{\partial p} \right)^2 p^{n_1} q^{N-n_1} \\ &= \left(p \frac{\partial}{\partial p} \right)^2 \sum_{n_1=0}^N \frac{N!}{n_1!(N-n_1)!} p^{n_1} q^{N-n_1} && \text{Durch Vertauschen} \\ & && \text{von Summation und} \\ & && \text{Differentiation} \\ &= \left(p \frac{\partial}{\partial p} \right)^2 (p+q)^N \\ &= \left(p \frac{\partial}{\partial p} \right) [pN(p+q)^{N-1}] && \text{infolge des Binomi-} \\ & && \text{alsatzes} \\ &= p[N(p+q)^{N-1} + pN(N-1)(p+q)^{N-2}] \end{aligned}$$

Der in (1.4.8) interessierende Fall ist, daß $p+q=1$. Somit wird (1.4.8) einfach

$$\begin{aligned} \overline{n_1^2} &= p[N + pN(N-1)] \\ &= Np[1 + pN - p] \\ &= (Np)^2 + Npq && \text{da } 1-p=q \\ &= \bar{n}_1^2 + Npq && \text{wegen (1.4.4)} \end{aligned}$$

Folglich ergibt (1.4.7) für das Schwankungsquadrat von n_1

$$\blacktriangleright \quad \overline{(\Delta n_1)^2} = Npq \quad (1.4.9)$$

Die Größe $\overline{(\Delta n_1)^2}$ ist quadratisch in der Abweichung. Ihre Quadratwurzel, d.h. die „mittlere quadratische Abweichung $\Delta^* n_1 \equiv [(\overline{(\Delta n_1)^2})^{\frac{1}{2}}$ “, ist ein lineares Maß für die Breite des Bereichs, über den die Werte von n_1 verteilt sind. Ein gutes Maß für die relative Breite dieser Verteilung ist dann

$$\frac{\Delta^* n_1}{\bar{n}_1} = \frac{\sqrt{Npq}}{Np} = \sqrt{\frac{q}{p}} \frac{1}{\sqrt{N}}$$

Insbesondere ist für $p=q=\frac{1}{2}$

$$\frac{\Delta^* n_1}{\bar{n}_1} = \frac{1}{\sqrt{N}}$$

Man beachte, daß mit wachsendem N der Mittelwert \bar{n}_1 wie N anwächst, während die Breite $\Delta^* n_1$ nur wie $N^{\frac{1}{2}}$ ansteigt. Folglich nimmt die relative Breite $\Delta^* n_1/\bar{n}_1$ mit wachsendem N wie $N^{-\frac{1}{2}}$ ab.

Man kann auch das Schwankungsquadrat von m , der resultierenden Verschiebung nach rechts, berechnen. Nach (1.2.3) ist

$$m = n_1 - n_2 = 2n_1 - N \quad (1.4.10)$$

Somit erhält man

$$\Delta m = m - \bar{m} = (2n_1 - n) - (2n_1 - N) = 2(n_1 - \bar{n}_1) = 2\Delta n_1 \quad (1.4.11)$$

und

$$(\Delta m)^2 = 4(\Delta n_1)^2$$

Als Mittelwert erhält man dann mittels (1.4.9)

$$\blacktriangleright \overline{(\Delta m)^2} = 4\overline{(\Delta n_1)^2} = 4Npq \quad (1.4.12)$$

Ist insbesondere $p = q = \frac{1}{2}$, so gilt

$$\overline{(\Delta m)^2}$$

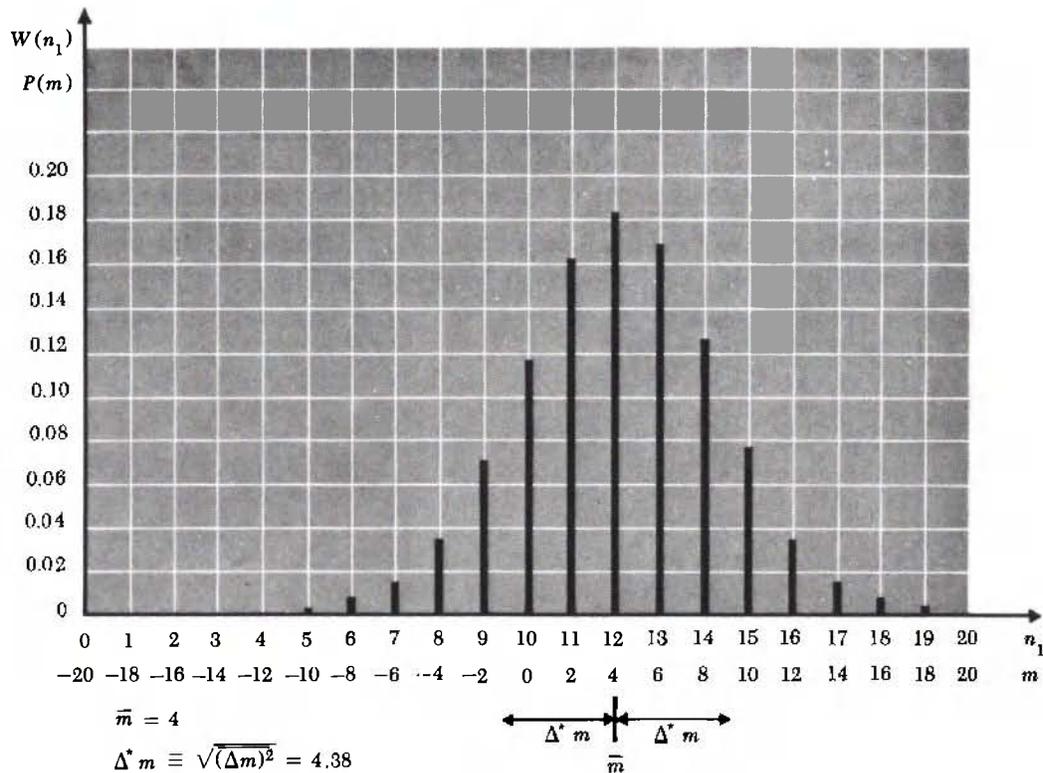


Abb. 1.4.1 Binomiale Wahrscheinlichkeitsverteilung für $p = 0.6$ und $q = 0.4$, wenn $N = 20$. Der Graph zeigt wieder die Wahrscheinlichkeit $W(n_1)$ für n_1 Rechtsverschiebungen bzw. die Wahrscheinlichkeit $P(m)$ für eine resultierende Verschiebung von m Einheiten nach rechts. Die Mittelwerte \bar{m} und $\overline{(\Delta m)^2}$ sind ebenfalls angegeben.

Beispiel: Man betrachte den Fall von $N = 100$ Einzelverschiebungen, wobei $p = q = \frac{1}{2}$ sei. Dann ist die mittlere Anzahl \bar{n}_1 von Rechtsverschiebungen (oder Linksverschiebungen) gleich 50; die mittlere Gesamtverschiebung $\bar{m} = 0$. Die mittlere quadratische Abweichung ist $[(\Delta m)^2]^{1/2} = 10$ (Verschiebungen).

1.5 Wahrscheinlichkeitsverteilung für großes N

Wenn N sehr groß wird, hat die binomiale Wahrscheinlichkeitsverteilung $W(n_1)$ von (1.4.1) die Tendenz, bei einem bestimmten Wert $n_1 = \bar{n}_1$ ein ausgeprägtes Maximum aufzuweisen und dann rasch abzufallen, sowie man sich von \bar{n}_1 entfernt (s. z.B. Abb. 1.4.1). Wir wollen diesen Umstand ausnutzen, um für $W(n_1)$ einen Näherungsausdruck herzuleiten, der für hinreichend großes N Gültigkeit besitzt.

Wenn N sehr groß ist, und wir ein Gebiet in der Umgebung der Stelle betrachten, wo W sein Maximum annimmt ⁷⁺⁾, so ist die relative Änderung von W bei einer Änderung von n_1 um eins dort sehr klein:

$$|W(n_1 + 1) - W(n_1)| \ll W(n_1) \quad (1.5.1)$$

Somit kann W in guter Näherung als eine stetige Funktion der kontinuierlichen Variablen n_1 angesehen werden, obwohl natürlich nur ganzzahlige Werte von n_1 physikalisch relevant sind. Die Stelle $n_1 = \bar{n}_1$ des Maximums von W ist dann näherungsweise bestimmt durch die Bedingung

$$\frac{dW}{dn_1} = 0 \text{ bzw. } \frac{d \ln W}{dn_1} = 0 \quad (1.5.2)$$

wobei die Ableitungen an der Stelle $n_1 = \bar{n}_1$ zu nehmen sind. Um das Verhalten von $W(n_1)$ in der Umgebung von \bar{n}_1 zu untersuchen, setzen wir

$$n_1 \equiv \bar{n}_1 + \eta \quad (1.5.3)$$

und entwickeln $\ln W(n_1)$ um die Stelle \bar{n}_1 in eine Taylorreihe. Der Grund dafür, daß wir $\ln W$ entwickeln anstatt die Funktion W selbst, besteht darin, daß $\ln W$ eine viel langsamer veränderliche Funktion von n_1 ist als W selbst, so daß die Potenzreihenentwicklung für $\ln W$ viel schneller konvergieren wird als die für W .

Ein Beispiel möge dies klarer machen. Angenommen, man möchte für $y \ll 1$ einen Näherungsausdruck für die Funktion

$$f \equiv (1 + y)^{-N}$$

finden, wobei N sehr groß ist. Direkte Entwicklung in eine Taylorreihe (oder mit Hilfe des Binomialsatzes) würde

$$f = 1 - Ny + \frac{1}{2}N(N+1)y^2 \dots$$

⁷⁺⁾ Wenn, wie wir voraussetzen wollen, p nicht sehr klein ist, so ist in diesem Gebiet mit N auch n_1 sehr groß.

ergeben. Da N sehr groß ist, gilt sogar für sehr kleine Werte von y die Beziehung $Ny \gtrsim 1$, so daß die obige Entwicklung keine brauchbare Näherung liefert. Man kann um diese Schwierigkeit herumkommen, indem man zuerst den Logarithmus bildet:

$$\ln f = -N \ln(1 + y)$$

Entwickelt man diesen in eine Taylorreihe, so erhält man

$$\ln f = -N(y - \frac{1}{2}y^2 \dots)$$

oder $f = e^{-N(y - \frac{1}{2}y^2 \dots)}$

eine Reihe mit betragsmäßig fallenden Gliedern, so lange nur $y < 1$ gilt.

Entwickelt man $\ln W$ in eine Taylorreihe, so erhält man

$$\ln W(n_1) = \ln W(\bar{n}_1) + B_1 \eta + \frac{1}{2} B_2 \eta^2 + \frac{1}{6} B_3 \eta^3 + \dots \quad (1.5.4)$$

wobei $B_k \equiv \frac{d^k \ln W}{dn_1^k}$ (1.5.5)

die k -te Ableitung von $\ln W$ an der Stelle $n_1 = \bar{n}_1$ ist. Da es sich um die Entwicklung um eine Extremstelle handelt, gilt nach (1.5.2) $B_1 = 0$. Außerdem folgt aus der Tatsache, daß W an der Extremstelle ein Maximum hat, daß B_2 bzw. das Glied $\frac{1}{2}B_2\eta^2$ negativ sein muß. Um das klar zum Ausdruck zu bringen, wollen wir schreiben $B_2 = -|B_2|$. Folglich ergibt (1.5.4), wenn wir $\bar{W} = W(\bar{n}_1)$ setzen,

$$W(n_1) = \bar{W} e^{\frac{1}{2}B_2\eta^2 + \frac{1}{6}B_3\eta^3 \dots} = \bar{W} e^{-\frac{1}{2}|B_2|\eta^2} e^{\frac{1}{6}B_3\eta^3 \dots} \quad (1.5.6)$$

Für das Gebiet, in dem η hinreichend klein ist, können in der Entwicklung Glieder höherer Ordnung vernachlässigt werden, so daß man in erster Näherung einen Ausdruck der einfachen Form

$$W(n_1) = \bar{W} e^{-\frac{1}{2}|B_2|\eta^2} \quad (1.5.7)$$

erhält.

Wir wollen nun die Entwicklung (1.5.4) genauer untersuchen. Nach (1.4.1) gilt

$$\ln W(n_1) = \ln N! - \ln n_1! - \ln(N - n_1)! + n_1 \ln p + (N - n_1) \ln q \quad (1.5.8)$$

Aber wenn n irgendeine sehr große ganze Zahl ist, so daß $n \gg 1$, dann kann $\ln n!$ als eine nahezu stetige Funktion von n angesehen werden, da die relative Änderung von $\ln n!$ wieder sehr klein ist, wenn sich n um eine kleine ganze Zahl ändert. Folglich ist

$$\frac{d \ln n!}{dn} \approx \frac{\ln(n+1)! - \ln n!}{1} = \ln \frac{(n+1)!}{n!} = \ln(n+1)$$

und für $n \gg 1$ gilt somit

$$\frac{d \ln n!}{dn} \approx \ln n \quad (1.5.9)$$

Damit wird aus (1.5.8)

$$\frac{d \ln W}{dn_1} = -\ln n_1 + \ln(N - n_1) + \ln p - \ln q \quad (1.5.10)$$

Durch Nullsetzen dieser ersten Ableitung erhält man den Wert $n_1 = \tilde{n}_1$, für den W sein Maximum annimmt:

$$\ln \left[\frac{(N - \tilde{n}_1) p}{\tilde{n}_1 q} \right] = 0$$

oder $(N - \tilde{n}_1)p = \tilde{n}_1 q$

so daß

► $\tilde{n}_1 = Np \quad (1.5.11)$

da $p + q = 1$ ist.

Weitere Differentiation von (1.5.10) ergibt

$$\frac{d^2 \ln W}{dn_1^2} = -\frac{1}{n_1} - \frac{1}{N - n_1} \quad (1.5.12)$$

Berechnet man diesen Ausdruck für die durch (1.5.11) gegebene Stelle \tilde{n}_1 , so erhält man

$$B_2 = -\frac{1}{Np} - \frac{1}{N - Np} = -\frac{1}{N} \left(\frac{1}{p} + \frac{1}{q} \right)$$

oder

► $B_2 = -\frac{1}{Npq} \quad (1.5.13)$

da $p + q = 1$ ist. Somit ist B_2 tatsächlich negativ, wie es im Falle eines Maximums von W ja auch sein muß.

Wenn man noch weitere Ableitungen bildet, kann man auch die Glieder höherer Ordnung in der Entwicklung (1.5.4) untersuchen. Somit erhält man durch Differentiation von (1.5.12)

$$B_3 = \frac{1}{\tilde{n}_1^2} - \frac{1}{(N - \tilde{n}_1)^2} = \frac{1}{N^2 p^2} - \frac{1}{N^2 q^2}$$

oder $|B_3| = \frac{|q^2 - p^2|}{N^2 p^2 q^2} < \frac{1}{N^2 p^2 q^2}$

Man sieht, daß in (1.5.4) der k -te Term betragsmäßig kleiner ist als $\eta^k/(Npq)^{k-1}$. Die Vernachlässigung von Termen größerer als zweiter Ordnung, die zu (1.5.7) führt, ist somit gerechtfertigt, falls η hinreichend klein ist, so daß ^{8*)}:

$$\eta \ll Npq \quad (1.5.14)$$

Andererseits bewirkt der Faktor $\exp(-\frac{1}{2}|B_2|\eta^2)$ in (1.5.6), daß W mit wachsenden Werten von $|\eta|$ sehr rasch abfällt, falls $|B_2|$ sehr groß ist. Falls nämlich

$$|B_2|\eta^2 = \frac{\eta^2}{Npq} \gg 1 \quad (1.5.15)$$

wird die Wahrscheinlichkeit $W(n_1)$ verglichen mit $W(\bar{n}_1)$ vernachlässigbar klein. Somit folgt, daß für Werte von η , die (1.5.14) und (1.5.15) erfüllen, den Ausdruck (1.5.7) eine ausgezeichnete Näherung für W in dem Gebiet liefert, wo W eine spürbare Größe besitzt. Diese Bedingung der gleichzeitigen Gültigkeit von (1.5.14) und (1.5.15) erfordert, daß

$$\sqrt{Npq} \ll \eta \ll Npq$$

$$\text{bzw. } Npq \gg 1 \quad (1.5.16)$$

ist. Dies zeigt, daß in dem ganzen Gebiet, in dem die Wahrscheinlichkeit W nicht vernachlässigbar klein ist, der Ausdruck (1.5.7) eine sehr gute Näherung liefert, solange nur N sehr groß ist und weder p noch q zu klein sind ^{9*)}.

Der Wert der Konstanten \tilde{W} in (1.5.7) läßt sich aus der Normierungsbedingung (1.4.2) bestimmen. Da W und n_1 als quasikontinuierliche Variable behandelt werden können, kann die Summe über alle ganzzahligen Werte von n_1 durch ein Integral ersetzt werden. Somit läßt sich die Normierungsbedingung schreiben

$$\sum_{n_1=0}^N W(n_1) \approx \int W(n_1) dn_1 = \int_{-\infty}^{\infty} W(\bar{n}_1 + \eta) d\eta = 1 \quad (1.5.17)$$

Hier kann das Integral über η in guter Näherung von $-\infty$ bis $+\infty$ erstreckt werden, da der Integrand überall dort einen vernachlässigbaren Beitrag zum Integral liefert, wo $|\eta|$ so groß ist, daß W weit von seinem ausgeprägten Maximum entfernt ist. Einsetzen von (1.5.7) in (1.5.17) ergibt dann unter Benutzung von (A.4.2)

^{8*)} Beachte, daß die Bedingung (1.5.1) der Bedingung $|\partial W/\partial n_1| \ll W$ und somit wegen (1.5.7) und (1.5.13) $|B_2\eta| = (Npq)^{-1}|\eta| \ll 1$ äquivalent ist. Somit ist sie auch in dem Gebiet (1.5.14) erfüllt, wo W nicht zu klein ist.

^{9*)} Wenn $p \ll 1$ oder $q \ll 1$, läßt sich eine andere Näherung für die Binomialverteilung gewinnen, nämlich die sogenannte „Poisson-Verteilung“ (siehe Aufgabe 1.9).

$$\bar{W} \int_{-\infty}^{\infty} e^{-\frac{1}{2}|B_2|\eta^2} d\eta = \bar{W} \sqrt{\frac{2\pi}{|B_2|}} = 1$$

Somit wird aus (1.5.7)

$$\blacktriangleright \quad W(n_1) = \sqrt{\frac{|B_2|}{2\pi}} e^{-\frac{1}{2}|B_2|(n_1 - \bar{n}_1)^2} \quad (1.5.18)$$

Die Überlegungen, die zu der funktionalen Gestalt (1.5.7) oder auch (1.5.18), der sogenannten „Gauss-Verteilung“, führten, waren sehr allgemeiner Natur. Deshalb ist es nicht überraschend, daß Gauss-Verteilungen sehr häufig in der Statistik auftreten, wenn man es mit großen Zahlen zu tun hat. In unserem Fall der Binomialverteilung geht der Ausdruck (1.5.18) mit (1.5.11) und (1.5.13) in

$$\blacktriangleright \quad W(n_1) = (2\pi Npq)^{-\frac{1}{2}} \exp \left[-\frac{(n_1 - Np)^2}{2Npq} \right] \quad (1.5.19)$$

über. Man beachte, daß (1.5.19) viel einfacher ist als (1.4.1), da hier die Berechnung der binomischen Koeffizienten wegfällt. Man beachte ferner, daß sich (1.5.19) mit Hilfe von (1.4.4) und (1.4.9) durch die Mittelwerte \bar{n}_1 und $(\Delta n_1)^2$ ausdrücken läßt:

$$W(n_1) = [2\pi(\Delta n_1)^2]^{-\frac{1}{2}} \exp \left[-\frac{(n_1 - \bar{n}_1)^2}{2(\Delta n_1)^2} \right]$$

1.6 Gaußsche Wahrscheinlichkeitsverteilungen

Der Gaußsche Näherungsausdruck (1.5.19) liefert unmittelbar die Wahrscheinlichkeit $P(m)$ dafür, daß bei einer großen Anzahl N von Einzelverschiebungen die effektive Verschiebung gleich m ist. Die entsprechende Anzahl von Rechtsverschiebungen ist nach (1.2.9) $n_1 = \frac{1}{2}(N + m)$. Folglich liefert (1.5.19)

$$P(m) = W \left(\frac{N + m}{2} \right) = [2\pi Npq]^{-\frac{1}{2}} \exp \left\{ -\frac{[m - N(p - q)]^2}{8Npq} \right\} \quad (1.6.1)$$

da $n_1 - Np = \frac{1}{2}[N + m - 2Np] = \frac{1}{2}[m - N(p - q)]$.

Nach (1.2.3) gilt $m = 2n_1 - N$, so daß m nur ganzzahlige Werte annimmt, die einen gegenseitigen Abstand von $\Delta m = 2$ haben.

Wir können dieses Ergebnis auch mit Hilfe der eigentlichen Verschiebungsvariablen x ausdrücken:

$$x = ml \quad (1.6.2)$$

wobei l die Schrittlänge ist. Wenn l klein ist verglichen mit der kleinsten Länge, die bei dem betrachteten physikalischen Problem von Belang ist ^{10*)}, so wird die Tatsache, daß x eigentlich nur in Abständen von $2l$ liegende, diskrete Werte annehmen kann, bedeutungslos. Ferner, wenn N sehr groß ist, ändert sich die Wahrscheinlichkeit $P(m)$ für das Auftreten einer resultierenden Verschiebung m nicht wesentlich beim Übergang zu einem benachbarten Wert von m , d.h.

$$|P(m+2) - P(m)| \ll P(m).$$

Deshalb kann $P(m)$ als eine stetige Funktion von x angesehen werden. Ein Graph

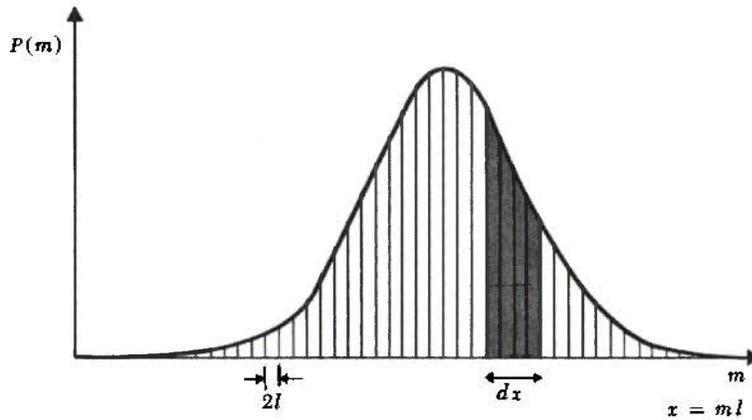


Abb. 1.6.1 Die Wahrscheinlichkeit $P(m)$ für eine resultierende Verschiebung nach rechts um m Einheiten, wenn die Gesamtzahl N der Einzelverschiebungen sehr groß und die Verschiebungslänge l sehr klein ist.

vom Typ, wie in Abb. 1.4.1 gezeigt, nimmt dann die in Abb. 1.6.1 dargestellte Form an, in der die senkrechten Linien sehr dicht liegen und die Einhüllende eine glatte Kurve bildet.

Unter diesen Umständen kann man x als eine kontinuierliche Variable in einem makroskopischen Maßstab ansehen und nach der Wahrscheinlichkeit dafür fragen, daß das Teilchen nach N Verschiebungen in einem Gebiet zwischen x und $x + dx$ angetroffen wird ^{11*)}. Da m nur ganzzahlige Werte mit einem Abstand von $\Delta m = 2$ annimmt, enthält das Gebiet der Länge dx insgesamt $dx/2l$ mögliche Werte von m , die alle mit ungefähr derselben Wahrscheinlichkeit $P(m)$ auftreten. Somit erhält man die Wahrscheinlichkeit dafür, daß sich das Teilchen irgendwo in dem Gebiet zwischen x und $x + dx$ befindet, einfach dadurch, daß man $P(m)$ aufsummiert über alle Werte von m , die in dx liegen, d.h. daß man $P(m)$ mit $dx/2l$ multipliziert.

^{10*)} Wenn man z.B. die Zufallsbewegung (Diffusion) eines Atoms in einem Festkörper betrachtet, ist l von der Ordnung der Gitterkonstanten, d.h. ungefähr 10^{-8} cm. Aber bei dem makroskopischen Maßstab der experimentellen Messung liegt die kleinste Länge L von physikalischem Belang bei 10^{-4} cm.

^{11*)} Hier ist dx als ein Differential im makroskopischen Sinn zu verstehen, d.h. $dx \ll L$, wobei L die kleinste makroskopisch relevante Länge ist, aber $dx \gg l$. (In anderen Worten: dx ist makroskopisch klein, aber mikroskopisch groß.)

Diese Wahrscheinlichkeit ist somit proportional zu dx (wie zu erwarten war) und läßt sich als

$$\mathcal{P}(x) dx = P(m) \frac{dx}{2l} \quad (1.6.3)$$

schreiben, wobei die Größe $\mathcal{P}(x)$, die unabhängig von der Intervalllänge dx ist, „Wahrscheinlichkeitsdichte“ heißt. Man beachte, daß sie mit einem differentiellen Wegelement der Länge dx multipliziert werden muß, um die Wahrscheinlichkeit selbst zu ergeben.

Unter Benutzung von (1.6.1) erhält man dann

$$\blacktriangleright \quad \mathcal{P}(x) dx = \frac{1}{\sqrt{2\pi} \sigma} e^{-(x-\mu)^2/2\sigma^2} dx \quad (1.6.4)$$

wobei wir die Abkürzungen

$$\mu \equiv (p - q)Nl \quad (1.6.5)$$

$$\text{und} \quad \sigma \equiv 2\sqrt{Npql} \quad (1.6.6)$$

eingeführt haben.

Der Ausdruck (1.6.4) ist die Standardform der Gaußschen Wahrscheinlichkeitsverteilung. Die große Allgemeinheit der zu (1.5.19) führenden Überlegungen, legt die Vermutung nahe, daß solche Gaußverteilungen sehr häufig in der Wahrscheinlichkeitstheorie auftreten, wenn man es mit großen Zahlen zu tun hat.

Ausgehend von (1.6.4) kann man ganz allgemein die Mittelwerte \bar{x} und $\overline{(x - \bar{x})^2}$ berechnen. Bei der Berechnung dieser Mittelwerte gehen Summen über alle möglichen Intervalle dx natürlich in Integrale über. (Die Grenzen von x können zu $-\infty < x < +\infty$ gewählt werden, da $\mathcal{P}(x)$ vernachlässigbar klein wird, wenn $|x|$ so groß ist, daß sich die entsprechende Gesamtverschiebung mit N Einzelverschiebungen gar nicht erreichen läßt.)

Zunächst wollen wir verifizieren, daß $\mathcal{P}(x)$ geeignet normiert ist, d.h. daß die Wahrscheinlichkeit dafür, daß Teilchen irgendwo vorzufinden, gleich eins ist:

$$\begin{aligned} \int_{-\infty}^{\infty} \mathcal{P}(x) dx &= \frac{1}{\sqrt{2\pi} \sigma} \int_{-\infty}^{\infty} e^{-(x-\mu)^2/2\sigma^2} dx \\ &= \frac{1}{\sqrt{2\pi} \sigma} \int_{-\infty}^{\infty} e^{-y^2/2\sigma^2} dy \\ &= \frac{1}{\sqrt{2\pi} \sigma} \sqrt{\pi 2\sigma^2} \\ &= 1 \end{aligned} \quad (1.6.7)$$

Hier haben wir $y \equiv x - \mu$ gesetzt und das Integral nach (A.4.2) ausgewertet.

Als nächstes berechnen wir den Mittelwert

$$\begin{aligned} \bar{x} &= \int_{-\infty}^{\infty} x \varphi(x) dx \\ &= \frac{1}{\sqrt{2\pi} \sigma} \int_{-\infty}^{\infty} x e^{-(x-\mu)^2/2\sigma^2} dx \\ &= \frac{1}{\sqrt{2\pi} \sigma} \left[\int_{-\infty}^{\infty} y e^{-y^2/2\sigma^2} dy + \mu \int_{-\infty}^{\infty} e^{-y^2/2\sigma^2} dy \right] \end{aligned}$$

Da der Integrand in dem ersten Integral eine ungerade Funktion von y ist, verschwindet das erste Integral aus Symmetriegründen. Das zweite Integral ist dasselbe wie in (1.6.7), so daß man

$$\bar{x} = \mu \tag{1.6.8}$$

erhält. Dies ist eine einfache Folge der Tatsache, daß $\varphi(x)$ eine Funktion von

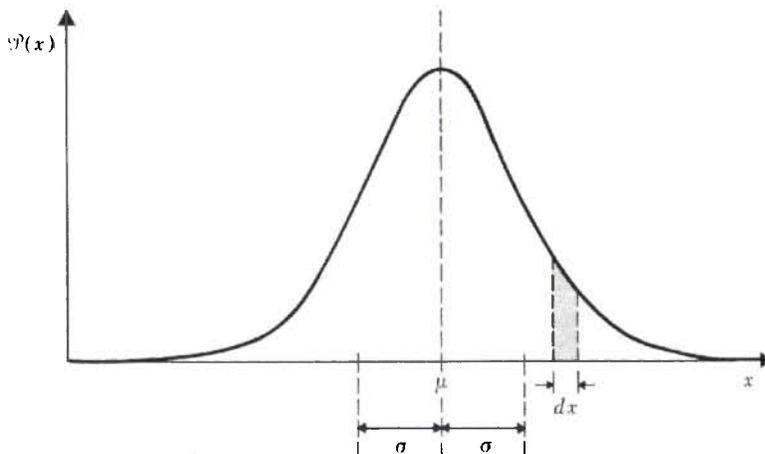


Abb. 1.6.2 Die Gauß-Verteilung. Hier ist $P(x)dx$ die Fläche unter der Kurve im Intervall zwischen x und $x+dx$ und ist somit die Wahrscheinlichkeit dafür, daß die Variable x in diesem Gebiet liegt.

$|x - \mu|$ ist und somit symmetrisch um die Stelle $x = \mu$ ihres Maximum ist. Folglich entspricht dieser Punkt auch dem Mittelwert von x .

Für das Schwankungsquadrat ergibt sich

$$\begin{aligned} \overline{(x - \mu)^2} &\equiv \int_{-\infty}^{\infty} (x - \mu)^2 \varphi(x) dx \\ &= \frac{1}{\sqrt{2\pi} \sigma} \int_{-\infty}^{\infty} y^2 e^{-y^2/2\sigma^2} dy \\ &= \frac{1}{\sqrt{2\pi} \sigma} \left[\frac{\sqrt{\pi}}{2} (2\sigma^2)^{\frac{3}{2}} \right] \\ &= \sigma^2 \end{aligned}$$

wobei wir die Integralformeln (A.4.6) benutzt haben. Somit haben wir

$$\blacktriangleright \quad \overline{(\Delta x)^2} = \overline{(x - \mu)^2} = \sigma^2 \quad (1.6.9)$$

Folglich ist σ einfach die mittlere quadratische Abweichung von x vom Mittelwert der Gauß-Verteilung.

Mit Hilfe von (1.6.5) und (1.6.6) erhält man dann für das Problem der Zufallsbewegung die Beziehungen

$$\bar{x} = (p - q)Nl \quad (1.6.10)$$

$$\overline{(\Delta x)^2} = 4Npql^2 \quad (1.6.11)$$

Diese Ergebnisse (hier abgeleitet für großes N) stehen natürlich im Einklang mit den Mittelwerten $\bar{x} = \bar{m}l$ und $\overline{(\Delta x)^2} = \overline{(\Delta m)^2}l^2$, die wir in (1.4.6) und (1.4.12) bereits für den allgemeinen Fall eines beliebigen N berechnet haben.

Allgemeine Diskussion der Zufallsbewegung

Unsere Diskussion des Problems der Zufallsbewegung hat eine große Anzahl von wichtigen Ergebnissen geliefert und viele grundlegende Begriffe der Wahrscheinlichkeitstheorie eingeführt. Der im wesentlichen auf eine kombinatorische Analyse fußende Zugang, den wir hier zur Berechnung der Wahrscheinlichkeitsverteilung benutzt haben, hat jedoch strenge Grenzen. Es ist insbesondere schwierig, diesen Weg auf andere Fälle zu verallgemeinern, z.B. auf Situationen, in denen die Länge der Einzelverschiebungen nicht mehr konstant ist oder in denen die betrachtete Bewegung nicht mehr eindimensional ist. Wir wenden uns deshalb jetzt der Diskussion von schlagkräftigeren Methoden zu, die sich leicht verallgemeinern lassen und die dennoch einen einfachen und direkten Zugang ermöglichen.

1.7 Wahrscheinlichkeitsverteilungen mit mehreren Variablen

Die statistische Beschreibung einer Situation, bei der mehrere Variable auftreten, erfordert lediglich die direkte Verallgemeinerung der wahrscheinlichkeitstheoretischen Überlegungen, die auf eine einzige Variable anwendbar waren. Wir wollen der Einfachheit halber den Fall von nur zwei Variablen u und v betrachten, die die möglichen Werte

$$u_i \quad i = 1, 2, \dots, M$$

und

$$v_j \quad j = 1, 2, \dots, N$$

annehmen können. Es sei $P(u_i, v_j)$ die Wahrscheinlichkeit dafür, daß u den Wert u_i und v den Wert v_j annimmt.

Die Wahrscheinlichkeit dafür, daß die Variablen u und v irgendeinen ihrer möglichen Werte annehmen, muß eins sein; d.h. man hat die Normierungsbedingung

$$\sum_{i=1}^M \sum_{j=1}^N P(u_i, v_j) = 1 \quad (1.7.1)$$

wobei sich die Summation über alle möglichen Werte von u und alle möglichen Werte von v erstreckt.

Die Wahrscheinlichkeit $P_u(u_i)$, daß u den Wert u_i annimmt, ohne Rücksicht darauf, welchen Wert die Variable v annimmt, ist die Summe der Wahrscheinlichkeiten für alle möglichen Situationen, die mit dem gegebenen Wert u_i verträglich sind; d.h.

$$P_u(u_i) = \sum_{j=1}^N P(u_i, v_j) \quad (1.7.2)$$

wobei sich die Summation über alle möglichen Werte v_j erstreckt. Genauso ist die Wahrscheinlichkeit $P_v(v_j)$, daß v den Wert v_j annimmt, ohne Rücksicht darauf, welchen Wert u annimmt, durch

$$P_v(v_j) = \sum_{i=1}^M P(u_i, v_j) \quad (1.7.3)$$

gegeben.

Jede der Wahrscheinlichkeiten P_u und P_v ist natürlich geeignet normiert. Z.B. hat man aufgrund von (1.7.2) und (1.7.1)

$$\sum_{i=1}^M P_u(u_i) = \sum_{i=1}^M \left[\sum_{j=1}^N P(u_i, v_j) \right] = 1 \quad (1.7.4)$$

Ein wichtiger Spezialfall liegt vor, wenn die Wahrscheinlichkeit dafür, daß die eine Variable einen gewissen Wert annimmt, nicht von dem Wert abhängt, der von der anderen Variablen angenommen wird. Die Variablen heißen dann „statistisch unabhängig“ oder auch „unkorreliert“. Die Wahrscheinlichkeit $P(u_i, v_j)$ läßt sich dann sehr einfach durch die Wahrscheinlichkeiten $P_u(u_i)$ und $P_v(v_j)$ ausdrücken. In diesem Fall erhält man nämlich [die Anzahl der Fälle in den Ensemble, bei denen $u = u_i$ und gleichzeitig $v = v_j$ ist] einfach dadurch, daß man [die Anzahl der Fälle, bei denen $u = u_i$ ist] mit [der Anzahl der Fälle, bei denen $v = v_j$ ist,] multipliziert; somit gilt

$$P(u_i, v_j) = P_u(u_i)P_v(v_j) \quad (1.7.5)$$

falls u und v statistisch unabhängig sind.

Wir wollen nun einige Eigenschaften der Mittelwerte erwähnen. Ist $F(u, v)$ irgendeine Funktion von u und v , so ist ihr Mittelwert durch

$$\overline{F(u,v)} \equiv \sum_{i=1}^M \sum_{j=1}^N P(u_i, v_j) F(u_i, v_j) \quad (1.7.6)$$

definiert. Man beachte, daß, falls $f(u)$ eine nur von u abhängige Funktion ist, aus (1.7.2)

$$\overline{f(u)} = \sum_i \sum_j P(u_i, v_j) f(u_i) = \sum_i P_u(u_i) f(u_i) \quad (1.7.7)$$

folgt. Sind F und G irgendwelche Funktionen von u und v , so hat man das allgemeine Ergebnis

$$\begin{aligned} \overline{F+G} &\equiv \sum_i \sum_j P(u_i, v_j) [F(u_i, v_j) + G(u_i, v_j)] \\ \blacktriangleright &= \sum_i \sum_j P(u_i, v_j) F(u_i, v_j) + \sum_i \sum_j P(u_i, v_j) G(u_i, v_j) \end{aligned} \quad (1.7.8)$$

$$\overline{F+G} = \overline{F} + \overline{G}$$

d.h. der Mittelwert einer Summe ist gleich der Summe der Mittelwerte.

Sind irgendzwei Funktionen $f(u)$ und $g(v)$ gegeben, so kann man ebenfalls eine allgemeine Aussage über den Mittelwert ihres Produktes machen, vorausgesetzt, daß u und v statistisch unabhängige Variable sind. Man findet nämlich

$$\begin{aligned} \overline{f(u)g(v)} &\equiv \sum_i \sum_j P(u_i, v_j) f(u_i) g(v_j) \\ &= \sum_i \sum_j P_u(u_i) P_v(v_j) f(u_i) g(v_j) \quad \text{nach (1.7.5)} \\ &= \left[\sum_i P_u(u_i) f(u_i) \right] \left[\sum_j P_v(v_j) g(v_j) \right] \end{aligned}$$

$$\blacktriangleright \quad \overline{f(u)g(v)} = \overline{f(u)} \overline{g(v)} \quad (1.7.9)$$

d.h. der Mittelwert eines Produktes ist gleich dem Produkt der Mittelwerte, falls u und v statistisch unabhängig sind. Falls u und v statistisch nicht unabhängig sind, ist die Aussage (1.7.9) im allgemeinen nicht richtig.

Die Verallgemeinerung der Definitionen und Resultate dieses Abschnittes auf mehr als zwei Variable ist evident.

1.8 Bemerkungen zu kontinuierlichen Wahrscheinlichkeitsverteilungen

Wir betrachten zunächst den Fall einer einzigen Variablen u , die irgendeinen Wert im Intervall $a_1 < u < a_2$ annehmen kann. Um für eine solche Situation eine sta-

tistische Beschreibung zu geben, kann man ein infinitesimales Intervall zwischen u und $u + du$ betrachten und nach der Wahrscheinlichkeit dafür fragen, daß die Variable einen Wert aus diesem Intervall annimmt. Man kann erwarten, daß diese Wahrscheinlichkeit proportional zu du ist, falls das Intervall hinreichend klein ist; d.h. man kann erwarten, daß diese Wahrscheinlichkeit in der Form $\mathcal{P}(u) du$ geschrieben werden kann, wo $\mathcal{P}(u)$ unabhängig von der Größe von du ist^{12*)}. Die Größe $\mathcal{P}(u)$ nennt man eine „Wahrscheinlichkeitsdichte“. Man beachte, daß diese mit du multipliziert werden muß, um eine tatsächliche Wahrscheinlichkeit zu ergeben.

Es ist ohne weiteres möglich, ein kontinuierliches Problem auf ein äquivalentes diskretes Problem zurückzuführen, wobei die möglichen Werte der Variablen abzählbar werden. Dazu braucht man nur das erlaubte Intervall $a_1 < u < a_2$ in (beliebig kleine) Teilintervalle gleicher Größe δu einzuteilen und diese durchnummerieren. Der Wert von u im Intervall wird dann einfach mit u_i bezeichnet und die Wahrscheinlichkeit dafür, daß u diesen Wert annimmt, mit $P(u_i)$. Man hat es dann wieder mit einer abzählbaren Menge von Werten der Variablen u zu tun (wobei jeder dieser Werte einem der fest gewählten infinitesimalen Intervalle entspricht). Damit ist auch klar, daß wahrscheinlichkeitstheoretische Beziehungen für diskrete Variable in gleicher Weise für kontinuierliche Variable Gültigkeit besitzen. So sind z.B. die einfachen Mittelwerteigenschaften (1.3.5) und (1.3.6) auch anwendbar, wenn u eine kontinuierliche Variable ist.



Abb. 1.8.1 Unterteilung des Gebietes $a_1 < u < a_2$ einer kontinuierlichen Variablen u in eine abzählbare Anzahl von infinitesimalen Intervallen der festen Größe δu .

Um den Zusammenhang zwischen dem kontinuierlichen und dem diskreten Standpunkt ganz deutlich zu machen, beachte man, daß aufgrund der ursprünglichen infinitesimalen Einteilung, die auf die Wahrscheinlichkeitsdichte $\mathcal{P}(u)$ führte, für das Intervall δu

$$P(u) = \mathcal{P}(u) \delta u$$

gilt. Wenn man nun irgendein Intervall zwischen u und $u + du$ betrachtet, das zwar makroskopisch klein ist, für das aber $du \gg \delta u$ gilt, so enthält dieses Intervall $du/\delta u$ mögliche Werte u_i , wobei die zugehörigen Wahrscheinlichkeiten $P(u_i)$ im wesentlichen denselben Wert haben, den wir einfach $P(u)$

^{12*)} Die Wahrscheinlichkeit muß nämlich als eine Potenzreihe in du darstellbar sein und muß verschwinden, wenn $du \rightarrow 0$. Somit muß der führende Term von der Form $\mathcal{P}(du)$ sein, während Glieder mit höheren Potenzen von du vernachlässigbar sind, falls du hinreichend klein ist.

nennen wollen. Die Wahrscheinlichkeit $P(u) du$ dafür, daß die Variable einen Wert zwischen u und $u + du$ annimmt, ergibt sich dann offenbar dadurch, daß man die Wahrscheinlichkeit $P(u_i)$ für irgendeinen in diesem Intervall liegenden diskreten Wert mit der Anzahl $du/\delta u$ der darin vorhandenen diskreten Werte multipliziert; d.h. man hat die Beziehung

$$\varphi(u) du = P(u_i) \frac{du}{\delta u} = \frac{P(u)}{\delta u} du \quad (1.8.1)$$

Man beachte, daß die Summen, die bei der Berechnung von Normierungsbedingungen oder Mittelwerten auftreten, als Integrale geschrieben werden können, wenn die Variable kontinuierlich ist. Die Normierungsbedingung besteht z.B. in der Aussage, daß die Summe der Wahrscheinlichkeiten über alle möglichen Werte der Variablen gleich eins sein muß:

$$\sum_i P(u_i) = 1 \quad (1.8.2)$$

Aber wenn die Variable kontinuierlich ist, kann man zunächst über alle Werte der Variablen in einem Gebiet zwischen u und $u + du$ summieren, womit man die Wahrscheinlichkeit $\varphi(u) du$ erhält, daß die Variable in diesem Gebiet liegt. Anschließend kann man dann die Summe (1.8.2) zu Ende führen, indem man über alle möglichen Gebiete du summiert (d.h. integriert). Somit ist (1.8.2) gleichbedeutend mit

$$\int_{a_1}^{a_2} \varphi(u) du = 1 \quad (1.8.3)$$

womit die Normierungsbedingung durch die Wahrscheinlichkeitsdichte $\varphi(u)$ ausgedrückt wird. Genauso lassen sich die Mittelwerte über $\varphi(u)$ berechnen. Die allgemeine Definition des Mittelwertes einer Funktion f war im Falle einer diskreten Variablen durch (1.3.4) gegeben, d.h.

$$\overline{f(u)} = \sum_i P(u_i) f(u_i) \quad (1.8.4)$$

Bei einer kontinuierlichen Variablen kann man zunächst wieder über alle Werte zwischen u und $u + du$ summieren [das liefert zur Summe (1.8.3) den Beitrag $\varphi(u) du f(u)$] und anschließend dann über alle Gebiete du integrieren. Somit ist (1.8.4) gleichbedeutend mit

$$\overline{f(u)} = \int_{a_1}^{a_2} \varphi(u) f(u) du \quad (1.8.5)$$

Anmerkung: Man beachte, daß in einigen Fällen die Wahrscheinlichkeitsdichte $\varphi(u)$ für gewisse Werte von u unendlich werden kann. Dies führt aber zu

keinerlei Schwierigkeiten, solange das Integral $\int_{c_1}^{c_2} \mathcal{P}(u) du$, welches ja die Wahrscheinlichkeit mißt, daß u irgendeinen Wert im Intervall zwischen c_1 und c_2 annimmt, für beliebiges c_1 und c_2 immer endlich bleibt.

Diese Bemerkungen lassen sich unmittelbar auf Wahrscheinlichkeitsverteilungen mit mehreren Variablen ausdehnen. Betrachten wir z.B. den Fall von zwei Variablen u und v , die jeweils kontinuierlich alle Werte in den Gebieten $a_1 < u < a_2$ und $b_1 < v < b_2$ annehmen können. Man kann dann von der Wahrscheinlichkeit

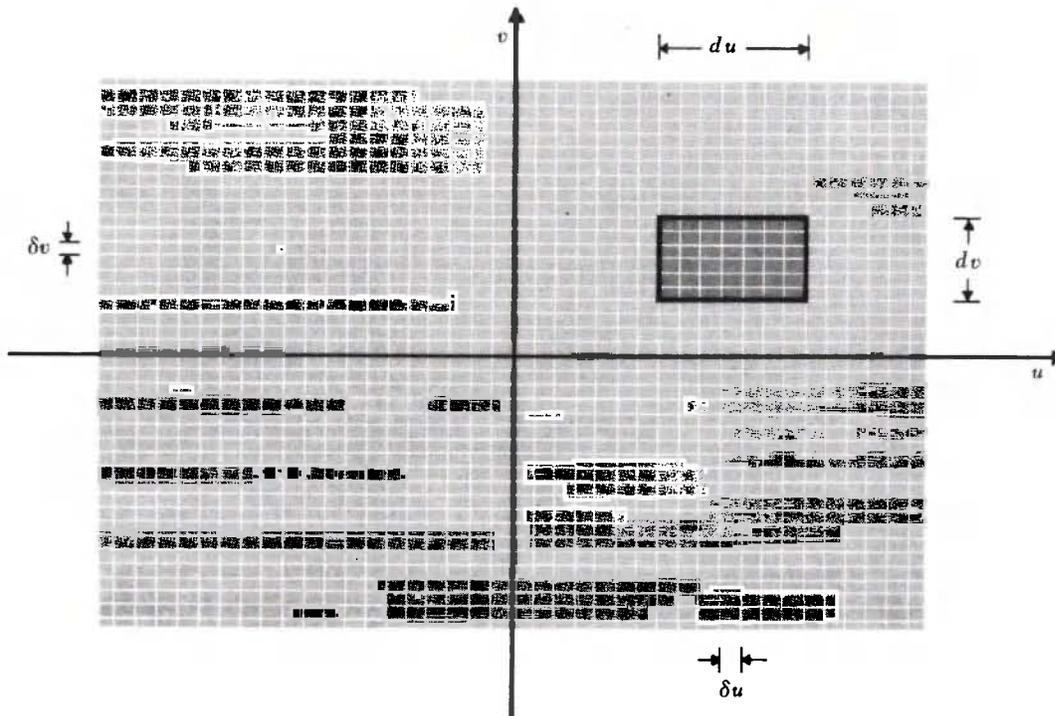


Abb. 1.8.2 Einteilung der kontinuierlichen Variablen u und v in kleine Intervalle der Größe δu und δv .

$\mathcal{P}(u, v) du dv$ sprechen, daß die Variablen in den Gebieten zwischen u und $u + du$ bzw. v und $v + dv$ liegen, wobei $\mathcal{P}(u, v)$ eine von der Größe von du und dv unabhängige Wahrscheinlichkeitsdichte ist. Man kann dann das kontinuierliche Problem wieder auf ein äquivalentes diskretes Problem mit abzählbar vielen Werten der Variablen zurückführen. Man braucht dazu nur die Definitionsbereiche von u und v jeweils in feste infinitesimale Intervalle der Größe δu bzw. δv zu unterteilen und erstere etwa durch einen Index i und letztere durch einen Index j zu kennzeichnen. Dann kann man von der Wahrscheinlichkeit $P(u_i, v_j)$ sprechen, daß $u = u_i$ und daß gleichzeitig $v = v_j$. Analog zu (1.8.1) erhält man dann die Beziehung

$$\mathcal{P}(u,v) du dv = P(u,v) \frac{du}{\delta u} \frac{dv}{\delta v}$$

wobei der Faktor hinter $P(u, v)$ einfach die Anzahl der infinitesimalen Zellen mit der Größe $\delta u \delta v$ angibt, die in dem Gebiet zwischen u und $u + du$ und v und $v + dv$ liegen.

Die Normierungsbedingung (1.7.2) läßt sich dann durch die Wahrscheinlichkeitsdichte $\mathcal{P}(u, v)$ ausdrücken gemäß

$$\int_{a_1}^{a_2} \int_{b_1}^{b_2} du dv \mathcal{P}(u,v) = 1 \quad (1.8.6)$$

Analog zu (1.7.7) kann man dann auch schreiben

$$\overline{F(u,v)} = \int_{a_1}^{a_2} \int_{b_1}^{b_2} du dv \mathcal{P}(u,v) F(u,v) \quad (1.8.7)$$

Da sich das Problem sowohl in diskreter als auch in kontinuierlicher Form behandeln läßt, bleiben die allgemeinen Eigenschaften (1.7.8) und (1.7.9) natürlich auch im kontinuierlichen Fall erhalten.

Funktionen von Zufallsvariablen. Wir betrachten den Fall einer einzigen Variablen u und nehmen an, daß $\mathcal{P}(u)$ eine stetige Funktion von u ist. Es stellt sich dann häufig die folgende Frage: Ist $\mathcal{P}(u)du$ die Wahrscheinlichkeit dafür, daß u im Gebiet zwischen u und $u + du$ liegt, wie groß ist dann die entsprechende Wahrscheinlichkeit $W(\varphi)d\varphi$ dafür, daß φ im Gebiet zwischen φ und $\varphi + d\varphi$ liegt? Man erhält diese Wahrscheinlichkeit offenbar dadurch, daß man die Wahrscheinlichkeiten für alle Werte von u aufsummiert, die ein im Gebiet zwischen φ und $\varphi + d\varphi$ liegendes φ ergeben, d.h.

$$W(\varphi) d\varphi = \int_{d\varphi} \mathcal{P}(u) du \quad (1.8.8)$$

Hier kann u als eine Funktion von φ angesehen werden, wobei sich das Integral über alle Werte von u erstreckt, die im Gebiet zwischen $u(\varphi)$ und $u(\varphi + d\varphi)$ liegen. Somit wird aus (1.8.8) einfach

$$W(\varphi) d\varphi = \int_{\varphi}^{\varphi+d\varphi} \mathcal{P}(u) \left| \frac{du}{d\varphi} \right| d\varphi = \mathcal{P}(u) \left| \frac{du}{d\varphi} \right| d\varphi \quad (1.8.9)$$

Im letzten Schritt wird angenommen, daß u eine eindeutige Funktion von φ ist, woraus dann das Ergebnis folgt, da sich das Integral nur über ein infinitesimales Gebiet $d\varphi$ erstreckt. Da $u = u(\varphi)$ ist, kann die rechte Seite von (1.8.9) natürlich vollständig durch φ ausgedrückt werden. Falls $u(\varphi)$ keine eindeutige Funktion von φ ist, kann es passieren, daß das Integral (1.8.8) aus mehreren Beiträgen der Form (1.8.9) besteht (siehe Abb. 1.8.3). Ähnliche Überlegungen können angestellt werden, wenn es darum geht, die Wahrscheinlichkeiten für Funktionen von mehreren

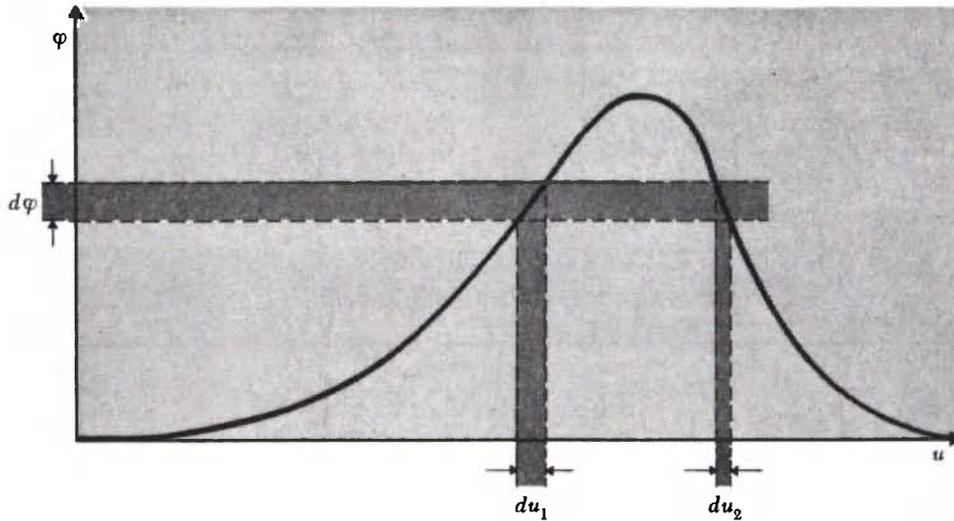


Abb. 1.8.3 Illustration einer Funktion $\varphi(u)$, deren Umkehrfunktion $u(\varphi)$ zweideutig¹³⁺⁾ ist. Hier entspricht das Gebiet $d\varphi$ einem u , das entweder im Gebiet du_1 oder im Gebiet du_2 liegt.

Variablen zu finden, wenn die Wahrscheinlichkeiten für die Variablen selbst bekannt sind.

Beispiel: Angenommen, ein zweidimensionaler Vektor B mit der konstanten Länge $B = |B|$ weist mit gleicher Wahrscheinlichkeit in jede durch den Winkel θ charakterisierte Richtung (siehe Abb. 1.8.4). Die Wahrscheinlichkeit $\varphi(\theta)d\theta$ dafür, daß dieser Winkel im Gebiet zwischen θ und $\theta + d\theta$ liegt, ist

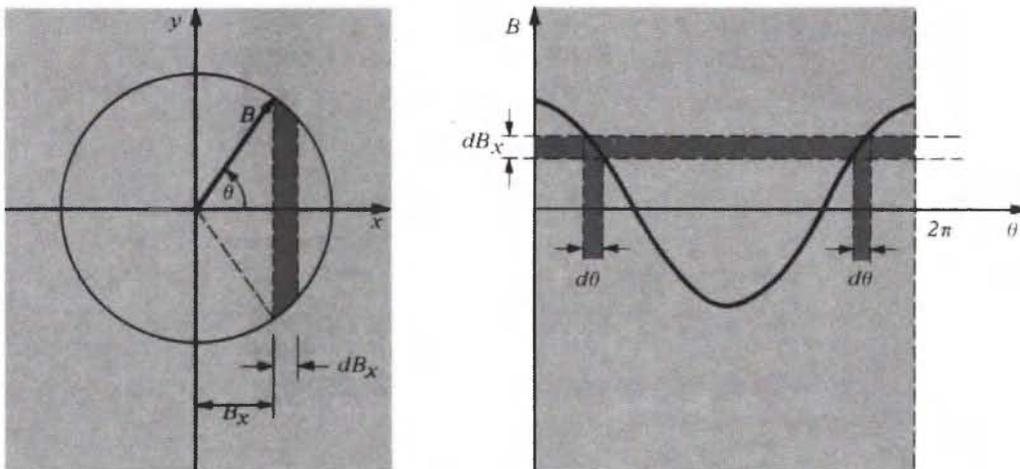


Abb. 1.8.4 Abhängigkeit der x -Komponente $B_x = B \cos \theta$ eines zweidimensionalen Vektors B von seinem Polarwinkel θ .

dann gegeben durch das Verhältnis des Winkelbereichs $d\theta$ zu dem Gesamtwinkelbereich 2π eines Vollkreises, d.h.

¹³⁺⁾ Genauer: deren Umkehrfunktion nicht existiert.

$$w(\theta) d\theta = \frac{d\theta}{2\pi} \quad (1.8.10)$$

Wenn der Vektor mit der x -Achse einen Winkel θ einschließt, ist seine x -Komponente gegeben durch

$$B_x = B \cos \theta \quad (1.8.11)$$

Wie groß ist die Wahrscheinlichkeit $W(B_x) dB_x$ dafür, daß die x -Komponente dieses Vektors zwischen B_x und $B_x + dB_x$ liegt? Offenbar gilt für B_x immer $-B \leq B_x \leq B$. In diesem Intervall entspricht ein infinitesimales Gebiet zwischen B_x und $B_x + dB_x$ zwei möglichen infinitesimalen Gebieten von θ (siehe Abb. 1.8.4), wobei jedes dieser beiden Gebiete die Größe $d\theta$ besitzt und mit dB_x durch die Beziehung (1.8.11) verknüpft ist, so daß $dB_x = |B \sin \theta| d\theta$. Aufgrund von (1.8.10) ist die Wahrscheinlichkeit $W(B_x) dB_x$ dann durch

$$W(B_x) dB_x = 2 \left[\frac{1}{2\pi} \frac{dB_x}{|B \sin \theta|} \right] = \frac{1}{\pi B} \frac{dB_x}{|\sin \theta|}$$

gegeben. Aber nach (1.8.11) ist

$$|\sin \theta| = (1 - \cos^2 \theta)^{1/2} = \left[1 - \left(\frac{B_x}{B} \right)^2 \right]^{1/2}$$

$$W(B_x) dB_x = \begin{cases} \frac{1}{\pi \sqrt{B^2 - B_x^2}} dB_x & \text{für } -B \leq B_x \leq B \\ 0 & \text{sonst} \end{cases} \quad (1.8.12)$$

Die Wahrscheinlichkeitsdichte ist maximal (und zwar unendlich), wenn $|B_x| \rightarrow B$, und sie ist minimal, wenn $B_x = 0$. Dieses Ergebnis läßt sich aus der Geometrie der Abb. 1.8.4 direkt ersehen, da ein gegebenes schmales Gebiet dB_x einem relativ großen Gebiet des Winkels θ entspricht, wenn $B_x \approx B$. Dagegen ist das zu dB_x gehörige Gebiet von θ sehr viel kleiner, wenn $B_x \approx 0$.

1.9 Allgemeine Berechnung von Mittelwerten für die Zufallsbewegung

Die Ausführungen des Abschnitts 1.7 erlauben eine äußerst einfache und durchsichtige Berechnung der Mittelwerte für sehr allgemeine Situationen. Wir wollen nun eine ganz allgemeine Form der eindimensionalen Zufallsbewegung betrachten. Wir bezeichnen mit s_i die (positive oder negative) Länge der i -ten Verschiebung und mit

$w(s_i) ds_i$ die Wahrscheinlichkeit dafür, daß der Wert von s_i im Intervall zwischen s_i und $s_i + ds_i$ liegt.

Wir wollen wieder annehmen, daß diese Wahrscheinlichkeit unabhängig von den anderen Verschiebungen ist, und gehen ferner der Einfachheit halber davon aus, daß die Wahrscheinlichkeitsverteilung w für jede einzelne Verschiebung i dieselbe ist. Es ist klar, daß die hier betrachtete Situation wesentlich allgemeiner ist als vorher, da wir nicht mehr für jede einzelne Verschiebung eine (bis aufs Vorzeichen bestimmte) konstante Länge l voraussetzen, sondern eine durch die Funktion w bestimmte Wahrscheinlichkeitsverteilung für einen ganzen Längenbereich.

Wir interessieren uns für die Gesamtverschiebung x ¹⁴⁺⁾ nach N Einzelverschiebungen. Wir können dazu nach der Wahrscheinlichkeit $\mathcal{P}(x) dx$ fragen, daß x im Gebiet zwischen x und $x + dx$ liegt. Ebenso können wir nach Mittelwerten von x fragen. Es wird sich in diesem Abschnitt zeigen, daß diese Mittelwerte sehr einfach auch ohne vorherige Kenntnis von $\mathcal{P}(x)$ berechnet werden können.

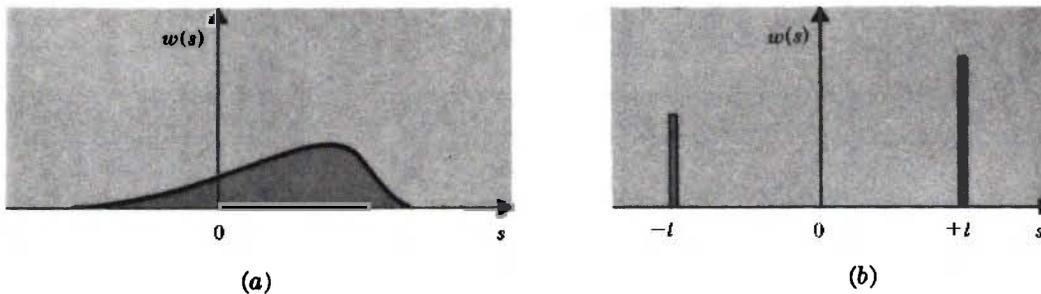


Abb. 1.9.1 Beispiele für Wahrscheinlichkeitsverteilungen, die für jede *einzelne* Verschiebung die Wahrscheinlichkeit $w(s) ds$ dafür angeben, daß die Verschiebung zwischen s und $s + ds$ liegt.
 a) Ein ziemlich allgemeiner Fall, bei dem Verschiebungen nach rechts wahrscheinlicher sind als solche nach links.
 b) Der in Abschnitt 1.2 diskutierte Spezialfall. Hier sind die „peaks“ bei $+l$ und $-l$ sehr schmal; die Fläche unter dem rechten „peak“ ist gleich p , die unter dem linken gleich q .
 (Die Kurven a) und b) sind nicht im gleichen Maßstab gezeichnet. Die von ihnen eingeschlossenen Gesamtflächen müssen in beiden Fällen gleich eins sein.)

Die Gesamtverschiebung x ist

$$x = s_1 + s_2 + \cdots + s_N = \sum_{i=1}^N s_i \quad (1.9.1)$$

Bilden wir auf beiden Seiten den Mittelwert, so ergibt sich

$$\bar{x} = \overline{\sum_{i=1}^N s_i} = \sum_{i=1}^N \bar{s}_i \quad (1.9.2)$$

¹⁴⁺⁾ Man sollte an sich entsprechend der vorangehenden Bezeichnungsweise von der „Länge x der Gesamtverschiebung“ sprechen.

wobei wir von der Eigenschaft (1.7.8) Gebrauch gemacht haben. Da nun die Funktion $w(s_i)$ unabhängig von i und damit für jede Verschiebung dieselbe ist, sind alle Mittelwerte \bar{s}_i gleich. Somit ist (1.9.2) einfach die Summe von N gleichen Termen und man erhält

$$\blacktriangleright \quad \bar{x} = N\bar{s} \quad (1.9.3)$$

$$\text{wobei} \quad \bar{s} \equiv \bar{s}_i = \int ds w(s)/s \quad (1.9.4)$$

die mittlere Einzelverschiebung ist.

Als nächstes berechnen wir das Schwankungsquadrat

$$\overline{(\Delta x)^2} \equiv \overline{(x - \bar{x})^2} \quad (1.9.5)$$

Nach (1.9.1) und (1.9.2) gilt

$$\begin{aligned} \Delta x &:= x - \bar{x} = \sum_i (s_i - \bar{s}) \\ \Delta x &= \sum_{i=1}^N \Delta s_i \end{aligned} \quad (1.9.6)$$

$$\text{wobei} \quad \Delta s_i = s_i - \bar{s} \quad (1.9.7)$$

Durch Quadrieren von (1.9.6) erhält man

$$(\Delta x)^2 = \left(\sum_{i=1}^N \Delta s_i \right) \left(\sum_{j=1}^N \Delta s_j \right) = \sum_i (\Delta s_i)^2 + \sum_i \sum_{i \neq j} (\Delta s_i)(\Delta s_j) \quad (1.9.8)$$

Hier enthält der erste Term auf der rechten Seite alle quadratischen und der zweite Term alle gemischten Glieder, die sich durch die Multiplikation der Summe mit sich selbst ergeben. Bildet man den Mittelwert von (1.9.8), so erhält man wegen (1.7.8)

$$\overline{(\Delta x)^2} = \sum_i \overline{(\Delta s_i)^2} + \sum_i \sum_{i \neq j} \overline{\Delta s_i \Delta s_j} \quad (1.9.9)$$

Bei den gemischten Gliedern machen wir von der Tatsache Gebrauch, daß verschiedene Verschiebungen statistisch unabhängig sind, so daß wir die Beziehung (1.7.9) anwenden und für $i \neq j$ schreiben können

$$\overline{(\Delta s_i)(\Delta s_j)} = \overline{(\Delta s_i)} \overline{(\Delta s_j)} = 0 \quad (1.9.10)$$

$$\text{da} \quad \overline{\Delta s_i} = \bar{s}_i - \bar{s} = 0$$

Jeder einzelne gemischte Term verschwindet also im Mittel, so daß sich (1.9.9) einfach auf die Summe der quadratischen Glieder reduziert

$$\overline{(\Delta x)^2} = \sum_{i=1}^N \overline{(\Delta s_i)^2} \quad (1.9.11)$$

Von diesen quadratischen Gliedern kann natürlich keines negativ sein. Da die Wahrscheinlichkeitsverteilung $w(s_i)$ unabhängig von i für jede Verschiebung dieselbe ist, folgt wieder, daß $(\Delta s_i)^2$ für alle Verschiebungen gleich ist. Somit besteht die Summe in (1.9.11) lediglich aus N gleichen Termen, so daß wir

$$\blacktriangleright \quad \overline{(\Delta x)^2} = N \overline{(\Delta s)^2} \quad (1.9.12)$$

erhalten, wobei

$$\overline{(\Delta s)^2} \equiv \overline{(\Delta s_i)^2} = \int ds w(s) (\Delta s)^2 \quad (1.9.13)$$

das Schwankungsquadrat für die einzelne Verschiebung ist.

Trotz ihrer großen Einfachheit stellen die Beziehungen (1.9.3) und (1.9.12) sehr allgemeine und wichtige Resultate dar. Das Schwankungsquadrat $\overline{(\Delta x)^2} = (x - \bar{x})^2$ ist ein Maß für das *Quadrat* der Breite des Gebietes, über das die Gesamtverschiebung x um ihren Mittelwert \bar{x} herum verstreut ist. Die Quadratwurzel $\Delta^*x \equiv [(\Delta x)^2]^{1/2}$, d.h. „die mittlere quadratische Abweichung vom Mittel“, liefert deshalb ein direktes Maß für die Breite dieses Gebietes. Mit Hilfe der Ergebnisse (1.9.3) und (1.9.12) lassen sich nun die folgenden interessanten Aussagen über die Summe (1.9.1) statistisch unabhängiger Variabler machen. Falls $\bar{x} \neq 0$ und die Anzahl N der Variablen (d.h. der Verschiebungen) anwächst, so wächst der Mittelwert \bar{x} ihrer Summe proportional zu N , während die Breite Δ^*x der Verteilung um das Mittel nur proportional zu $N^{1/2}$ anwächst. Somit haben wir für die *relative* Größe der Breite Δ^*x , bezogen auf das Mittel \bar{x} selbst, eine *Abnahme* proportional zu $N^{-1/2}$, d.h. nach (1.9.3) und (1.9.12) für $\bar{x} \neq 0$

$$\frac{\Delta^*x}{\bar{x}} = \frac{\Delta^*s}{\bar{s}} \frac{1}{\sqrt{N}}$$

wobei $\Delta^*s \equiv [(\Delta s)^2]^{1/2}$. Das bedeutet, daß die prozentuale Abweichung der x -Werte von ihrem Mittelwert \bar{x} zunehmend vernachlässigbar wird, wenn die Anzahl N sehr stark anwächst. Dies ist ein charakteristisches Merkmal statistischer Verteilungen.

Beispiel: Wir wollen die allgemeinen Resultate (1.9.3) und (1.9.12) dieses Abschnittes auf den Spezialfall der Zufallsbewegung mit fester Verschiebungslänge l anwenden, den wir ja schon im Abschn. 1.2 gesondert diskutiert haben. Die Wahrscheinlichkeit für einen Schritt nach rechts ist dort p und die für einen Schritt nach links $q = 1 - p$. Die mittlere Länge der Einzelverschiebung ist dann durch

$$\bar{s} = pl + q(-l) = (p - q)l = (2p - 1)l \quad (1.9.14)$$

gegeben. Als Kontrolle beachte man, daß $\bar{x} = 0$, falls $p = q$, wie es die Symmetrie erfordert.

Ebenso gilt

$$\begin{aligned}\overline{s^2} &= pl^2 + q(-l)^2 = (p+q)l^2 = l^2 \\ \overline{(\Delta s)^2} &= \overline{s^2} - \overline{s}^2 = l^2[1 - (2p-1)^2] \\ &= l^2[1 - 4p^2 + 4p - 1] = 4l^2 p(1-p)\end{aligned}$$

$$\text{Somit } \overline{(\Delta s)^2} = 4pql^2 \quad (1.9.15)$$

Folglich ergeben die Beziehungen (1.9.3) und (1.9.12)

$$\left. \begin{aligned}\bar{x} &= (p-q)Nl \\ \overline{(\Delta x)^2} &= 4pql^2\end{aligned}\right\} \quad (1.9.16)$$

Da $x = ml$, stimmen diese Beziehungen mit den früher berechneten Ergebnissen (1.4.6) und (1.4.12) genau überein.

* 1.10 Berechnung der Wahrscheinlichkeitsverteilung

Bei dem im letzten Abschnitt diskutierten Problem ist die Gesamtverschiebung x bei N Einzelverschiebungen durch

$$x = \sum_{i=1}^N s_i \quad (1.10.1)$$

gegeben. Wir wollen nun die Wahrscheinlichkeit $\mathcal{P}(x) dx$ dafür ermitteln, daß sich x im Gebiet zwischen x und $x + dx$ befindet. Da die einzelnen Verschiebungen voneinander statistisch unabhängig sind, ist die Wahrscheinlichkeit für eine *spezielle* Folge von Verschiebungen, bei der

die 1. Verschiebung im Gebiet zwischen s_1 und $s_1 + ds_1$

die 2. Verschiebung im Gebiet zwischen s_2 und $s_2 + ds_2$

...

die N . Verschiebung im Gebiet zwischen s_N und $s_N + ds_N$

liegt, einfach durch das Produkt der entsprechenden Wahrscheinlichkeiten gegeben, d.h. durch

$$w(s_1) ds_1 \cdot w(s_2) ds_2 \cdots w(s_N) ds_N$$

Wenn wir diese Wahrscheinlichkeit über alle möglichen Einzelverschiebungen aufsummieren, die mit der Bedingung verträglich sind, daß die Gesamtverschiebung x in (1.10.1) immer im Gebiet zwischen x und $x + dx$ liegt, so erhalten wir die gesuchte Wahrscheinlichkeit $\mathcal{P}(x)dx$, die also unabhängig von der Folge der Einzelverschiebungen ist, die diese Gesamtverschiebung hervorbringt. Wir haben somit

$$\mathcal{P}(x) dx = \int \int \cdots \int_{\substack{\infty \\ (dx) \\ -\infty}} w(s_1)w(s_2) \cdots w(s_N) ds_1 ds_2 \cdots ds_N \quad (1.10.2)$$

wobei sich die Integration über alle möglichen Werte der Variablen s_i erstreckt, die der Einschränkung

$$x < \sum_{i=1}^N s_i < x + dx \quad (1.10.3)$$

genügen. Offensichtlich ist mit der Auswertung des Integrals (1.10.2) die gesuchte Funktion $\mathcal{P}(x)$ gefunden.

Was die Praxis anbelangt, so ist das Integral sehr schwer zu berechnen, da die Bedingung (1.10.3) die Integrationsgrenzen sehr unangenehm macht; d.h. man hat es hier mit dem komplizierten geometrischen Problem zu tun, den mit (1.10.3) konsistenten Unterraum zu bestimmen, über den zu integrieren ist. Eine sehr wirkungsvolle Methode, Probleme dieser Art zu behandeln, besteht darin, daß man das geometrische Problem eliminiert, indem man über *alle* Werte der Variablen s_i *ohne* Einschränkung integriert, und dabei die durch (1.10.3) eingeführte Komplikation auf den *Integranden* abwälzt. Dies läßt sich ohne Weiteres dadurch bewerkstelligen, daß man den Integranden in (1.10.2) mit einem Faktor multipliziert, der gleich eins ist, wenn die s_i der Bedingung (1.10.3) genügen, und sonst immer Null. Die Diracsche δ -Funktion $\delta(x - x_0)$, die im Anhang A.7 besprochen wird, hat nun die Eigenschaft, daß man für (1.10.2) auch

$$\mathcal{P}(x) dx = \iint \cdots \int_{-\infty}^{\infty} w(s_1)w(s_2) \cdots w(s_N) \left[\delta \left(x - \sum_{i=1}^N s_i \right) dx \right] ds_1 ds_2 \cdots ds_N \quad (1.10.4)$$

schreiben kann, wobei nun das Integrationsgebiet *keiner* Einschränkung mehr unterworfen ist. (Man überlege sich die Äquivalenz von (1.10.2) mit (1.10.4) durch Betrachtung der Fälle $N = 1, 2$ und 3 .) An dieser Stelle können wir die übliche Integraldarstellung (A.7.14) der δ -Funktion benutzen: d.h. wir können

$$\delta(x - \Sigma s_i) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ik[\Sigma s_i - x]} \quad (1.10.5)$$

schreiben. Einsetzen von (1.10.5) in (1.10.4) ergibt

$$\begin{aligned} \mathcal{P}(x) &= \iint \cdots \int w(s_1)w(s_2) \cdots w(s_N) \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ik(s_1 + \cdots + s_N - x)} ds_1 ds_2 \cdots ds_N \\ \mathcal{P}(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{-ikx} \int_{-\infty}^{\infty} ds_1 w(s_1) e^{iks_1} \cdots \int_{-\infty}^{\infty} ds_N w(s_N) e^{iks_N} \end{aligned} \quad (1.10.6)$$

wobei wir die Reihenfolge der Integrationen vertauscht und von der multiplikativen Eigenschaft der Exponentialfunktion Gebrauch gemacht haben. Abgesehen von dem belanglosen Symbol für die Integrationsvariable sind die letzten N Integrale alle gleich und wir setzen

$$\blacktriangleright \quad Q(k) \equiv \int_{-\infty}^{\infty} ds e^{iks} w(s) \quad (1.10.7)$$

Somit wird aus (1.10.6)

$$\blacktriangleright \quad \mathcal{P}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{-ikx} Q^N(k) \quad (1.10.8)$$

Folglich ist das gestellte Problem mit der Berechnung von zwei einfachen (Fourier-)Integralen vollständig gelöst.

Beispiel: Wir wollen diese letzten Ergebnisse wieder auf den in Abschnitt 1.2 bereits besprochenen Fall konstanter Verschiebungslänge l anwenden. Die Wahrscheinlichkeit für eine Verschiebung $+l$ ist dort gleich p , die für eine Verschiebungslänge $-l$ ist gleich $q = 1 - p$; d.h. die entsprechende Wahrscheinlichkeitsdichte w ist gegeben durch

$$w(s) = p\delta(s - l) + q\delta(s + l)$$

Die Größe (1.10.7) wird

$$Q(k) \equiv \overline{e^{iks}} = p e^{ikl} + q e^{-ikl}$$

Unter Benutzung der Binomialentwicklung erhält man

$$\begin{aligned} Q^N(k) &= (p e^{ikl} + q e^{-ikl})^N \\ &= \sum_{n=0}^N \frac{N!}{n!(N-n)!} (p e^{ikl})^n (q e^{-ikl})^{N-n} \\ &= \sum_{n=0}^N \frac{N!}{n!(N-n)!} p^n q^{N-n} e^{ikl(2n-N)} \end{aligned}$$

Somit ergibt (1.10.8)

$$\begin{aligned} \mathcal{P}(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{-ikx} Q^N(k) \\ &= \sum_{n=0}^N \frac{N!}{n!(N-n)!} p^n q^{N-n} \left\{ \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ik[(2n-N)l-x]} \right\} \\ \mathcal{P}(x) &= \sum_{n=0}^N \frac{N!}{n!(N-n)!} p^n q^{N-n} \delta[x - (2n - N)l] \end{aligned} \quad (1.10.9)$$

Dies besagt, daß die Wahrscheinlichkeitsdichte $\mathcal{P}(x)$ verschwindet, sofern nicht

$$x = (2n - N)l, \quad \text{wobei } n = 0, 1, 2, \dots, N$$

Die Wahrscheinlichkeit $P(2n - N)$, ein Teilchen an einer solchen Stelle zu finden, ist dann gegeben durch

$$P(2n - N) = \int_{(2n-N)l-\epsilon}^{(2n-N)l+\epsilon} \mathcal{P}(x) dx = \frac{N!}{n!(N-n)!} p^n q^{N-n}$$

wobei ϵ irgendeine hinreichend kleine Größe ist; d.h. P ist gegeben durch den Koeffizienten der entsprechenden δ -Funktion in (1.10.9). Somit erhalten wir wieder das Ergebnis des Abschnitts 1.2, und zwar jetzt als Spezialfall einer allgemeineren und *ohne* kombinatorische Überlegungen operierenden Formulierung.

1.11 Wahrscheinlichkeitsverteilung für großes N

Wir betrachten die Integrale (1.10.7) und (1.10.8), die die Berechnung von $P(x)$ ermöglichen, und fragen nach passenden Näherungen für den Fall, daß N sehr groß wird. Die Überlegungen, auf die wir uns dabei stützen werden, ähneln weitgehend denen, die im Anhang A.6 zur Herleitung der Stirling-Formel benutzt werden.

Der Integrand in (1.10.7) enthält den Faktor e^{iks} , der eine oszillierende Funktion von s ist und umso schneller oszilliert, je größer k ist. Somit wird die durch das Integral (1.10.7) gegebene Größe $Q(k)$ im allgemeinen immer kleiner, wenn k groß wird. (Siehe unten.) Daraus folgt für großes N , daß die Potenz $Q^N(k)$ mit wachsendem k sehr rasch abnimmt. Es genügt deshalb zur Berechnung von $P(x)$ aus (1.10.8), wenn man $Q^N(k)$ für *kleine* Werte von k kennt, da für große Werte von k der Beitrag von $Q^N(k)$ zum Integral vernachlässigbar klein ist. Aber für kleine Werte von k läßt sich $Q^N(k)$ durch eine geeignete Potenzreihenentwicklung in k approximieren. Da $Q^N(k)$ selbst eine sehr schnell veränderliche Funktion von k ist, ist es (wie im Abschnitt 1.5) vorteilhaft, die schneller konvergierende Potenzreihenentwicklung des langsam veränderlichen Logarithmus $\ln Q^N(k)$ aufzusuchen.

Anmerkung: Solange $w(s)$ über eine Schwingungsperiode nur langsam variiert, gilt $Q(k) = \int ds e^{iks} w(s) \approx 0$. In jedem Intervall $a < s < b$, nämlich in dem w so langsam variiert, daß $|dw/ds|(b-a) \ll w$, das aber immer noch so viele Schwingungen enthält, daß $(b-a)k \gg 1$, gilt für das Integral

$$\int_a^b ds e^{iks} w(s) \approx w(a) \int_a^b ds e^{iks} \approx 0$$

Man kann deshalb sagen, daß

$$\int_{-\infty}^{\infty} ds e^{iks} w(s) \approx 0$$

solange k groß genug ist, so daß überall

$$\left| \frac{dw}{ds} \right| \frac{1}{k} \ll w$$

Die eigentliche Berechnung geht nun ganz einfach vonstatten. Wir beginnen mit $Q(k)$ für kleine Werte von k . Entwickeln wir e^{iks} in eine Taylorreihe, so wird aus (1.10.7)

$$\begin{aligned}
 Q(k) &\equiv \int_{-\infty}^{\infty} ds w(s) e^{iks} = \int_{-\infty}^{\infty} ds w(s) (1 + iks - \frac{1}{2}k^2s^2 + \dots) \\
 Q(k) &= 1 + i\bar{s}k - \frac{1}{2}\bar{s}^2k^2 \dots
 \end{aligned}
 \tag{1.11.1}$$

$$\text{wobei } \bar{s}^n \equiv \int_{-\infty}^{\infty} ds w(s) s^n \tag{1.11.2}$$

eine Konstante ist, die nach unserer früheren Definition einfach das n -te Moment von s darstellt. Wir nehmen hier an, daß schnell genug $|w(s)| \rightarrow 0$, wenn $|s| \rightarrow \infty$, so daß diese Momente endlich bleiben. Damit ergibt (1.11.1)

$$\ln Q^N(k) = N \ln Q(k) = N \ln [1 + i\bar{s}k - \frac{1}{2}\bar{s}^2k^2 \dots] \tag{1.11.3}$$

Benutzen wir die für $y \ll 1$ gültige Potenzreihenentwicklung

$$\ln(1 + y) = y - \frac{1}{2}y^2 \dots$$

so wird aus (1.11.3), wenn wir nur bis zu quadratischen Gliedern in k gehen

$$\begin{aligned}
 \ln Q^N &= N[i\bar{s}k - \frac{1}{2}\bar{s}^2k^2 - \frac{1}{2}(i\bar{s}k)^2 \dots] \\
 &= N[i\bar{s}k - \frac{1}{2}(\bar{s}^2 - \bar{s}^2)k^2 \dots] \\
 &= N[i\bar{s}k - \frac{1}{2}(\overline{\Delta s})^2k^2 \dots]
 \end{aligned}$$

$$\text{wobei } \overline{(\Delta s)^2} \equiv \bar{s}^2 - \bar{s}^2 \tag{1.11.4}$$

Somit erhalten wir

$$Q^N(k) = e^{iN\bar{s}k - \frac{1}{2}N(\overline{\Delta s})^2k^2} \tag{1.11.5}$$

und aus (1.10.8) wird

$$\mathcal{P}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{i(N\bar{s}-x)k - \frac{1}{2}N(\overline{\Delta s})^2k^2} \tag{1.11.6}$$

Hier ist das Integral von der folgenden Form,

$$\begin{aligned}
 \int_{-\infty}^{\infty} du e^{-au^2+bu} &= \int_{-\infty}^{\infty} du e^{-a[u^2-(b/a)u]} \\
 &= \int_{-\infty}^{\infty} du e^{-a(u-b/2a)^2+b^2/4a} \quad (\text{quadratische Ergänzung}) \\
 &= e^{b^2/4a} \int_{-\infty}^{\infty} dy e^{-ay^2} \quad (y = u - \frac{b}{2a}) \\
 &= e^{b^2/4a} \sqrt{\frac{\pi}{a}} \quad (\text{A.4.2})
 \end{aligned}$$

$$\text{somit } \int_{-\infty}^{\infty} du e^{-au^2+bu} = \sqrt{\frac{\pi}{a}} e^{b^2/4a} \tag{1.11.7}$$

wobei a reel und positiv ist. Wenden wir diese Integralformel auf (1.11.6) an, so erhalten wir mit $b = i(N\bar{s} - x)$ und $a = \frac{1}{2}N(\overline{\Delta s})^2$ das Ergebnis

$$\mathcal{P}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} \quad (1.11.8)$$

$$\left. \begin{array}{l} \mu \equiv N\bar{s} \\ \sigma^2 \equiv N(\Delta s)^2 \end{array} \right\} \quad (1.11.9)$$

Somit hat die Verteilung die Gaußsche Form, der wir ja schon früher in Abschnitt 1.6 begegnet sind. Man beachte jedoch die außerordentliche Allgemeinheit dieses Ergebnisses. *Ganz gleich wie* die Wahrscheinlichkeitsverteilung $w(s)$ für die Einzelverschiebung aussieht, solange die Verschiebungen *statistisch unabhängig* sind und $w(s)$ mit $|s| \rightarrow \infty$ schnell genug abfällt, wird die Gesamtverschiebung x immer nach dem Gaußschen Gesetz verteilt sein, *falls N hinreichend groß ist*. Dieses sehr wichtige Resultat ist der Inhalt des sogenannten „zentralen Grenzwertsatzes“, der, wie der Name schon sagt, im Mittelpunkt der mathematischen Wahrscheinlichkeitstheorie steht^{15*)}. Die Allgemeinheit des Resultats erklärt auch die Tatsache, daß so viele Naturphänomene (z.B. Meßfehler) angenähert einer Gauß-Verteilung gehorchen.

Wir haben bereits gezeigt, daß für die Gauß-Verteilung (1.6.4) gilt

$$\left. \begin{array}{l} \bar{x} = \mu \\ (\Delta x)^2 = \sigma^2 \end{array} \right\}$$

Somit besagt (1.11.9)

$$\left. \begin{array}{l} \bar{x} = N\bar{s} \\ (\Delta x)^2 = N(\Delta s)^2 \end{array} \right\} \quad (1.11.10)$$

was mit den Ergebnissen übereinstimmt, die wir aus unseren allgemeinen Berechnungen (1.9.3) und (1.9.12) gewonnen haben.

Ergänzende Literatur

Wahrscheinlichkeitstheorie

Mosteller, F.; Rourke, R.E.K.; Thomas, G.B.: Probability with statistical applications. 2. Aufl., Addison-Wesley Reading, Mass. (1970).

(Eine elementare Einführung).

^{15*)} Ein Beweis dieses Satzes mit besonderer Rücksicht auf mathematische Strenge findet sich bei Chintschin, A.J.: Mathematische Grundlagen der Statistischen Mechanik (B-I-Hochschultaschenbücher Nr. 58/58 a.) S. 164 ff, Bibliograph. Institut Mannheim (1964).

Feller, W.: An introduction to probability theory and its applications. 3. Aufl. Wiley, New York (1968).

Cramér, H.: The elements of probability theory. Wiley, New York (1955).

Zufallsbewegung

Chandrasekhar, S.: Stochastic problems in physics and astronomy. Rev. mod. Phys. 15 (1943), 1–89. Dieser Artikel ist auch zu finden bei:

Wax, M.: Selected papers on noise and stochastic processes. Dover Publ. New York (1954).

Lindsay, R.B.: Introduction to physical statistics. Kap. 2. Wiley, New York (1941). (Eine elementare Diskussion der Zufallsbewegung und verwandter physikalische Probleme.)

Aufgaben

- 1.1 Wie groß ist die Wahrscheinlichkeit dafür, mit drei Würfeln insgesamt höchstens 6 Augen zu werfen?
- 1.2 Bei einem Spiel werden sechs ideale Würfel geworfen. Man bestimme die Wahrscheinlichkeit dafür, daß man dabei
 - a) genau eine Eins,
 - b) mindestens eine Eins
 - c) genau zwei Einsenwirft.
- 1.3 Es wird eine Zahl zwischen 0 und 1 aufs Geratewohl gewählt. Wie groß ist die Wahrscheinlichkeit dafür, daß genau 5 ihrer ersten 10 Dezimalstellen kleiner als 5 sind?
- 1.4 Ein Betrunkener startet von einem Laternenpfahl aus, der sich genau in der Mitte zwischen den Enden einer Straße befindet. Er macht dabei mit gleicher Wahrscheinlichkeit gleich lange Schritte entweder nach links oder nach rechts. Wie groß ist die Wahrscheinlichkeit dafür, daß der Mann nach N Schritten wieder am Laternenpfahl angekommen ist,
 - a) falls N gerade ist?
 - b) falls N ungerade ist?
- 1.5 Beim Russischen Roulette (ein vom Autor nicht empfohlenes Spiel) steckt man in die Trommel eines Revolvers eine einzige Patrone und läßt die übrigen fünf Kammern der Trommel leer. Dann versetzt man die Trommel in eine rasche Drehbewegung und drückt schließlich, nachdem die Trommel wieder zum Stillstand gekommen ist, auf den eigenen Kopf ab.
 - a) Wie groß ist die Wahrscheinlichkeit, nach N Runden noch am Leben zu sein?

- b) Wie groß ist die Wahrscheinlichkeit, $(N - 1)$ Runden zu überleben und dann beim N -ten Mal erschossen zu werden?
- c) Wie oft hat im Mittel ein Spieler bei diesem makabren Spiel die Gelegenheit, am Abzug zu ziehen?
- 1.6 Man betrachte das Problem der Zufallsbewegung mit $p = q$ und bezeichne mit $m = n_1 - n_2$ die Gesamtverschiebung nach rechts. Wie groß sind bei insgesamt N Einzelverschiebungen die Mittelwerte \bar{m} , \bar{m}^2 , \bar{m}^3 und \bar{m}^4 ?
- 1.7 Man leite die Binomialverteilung auf die folgende algebraische Art her, die keinerlei explizite kombinatorische Untersuchungen enthält. Gesucht ist wieder die Wahrscheinlichkeit $W(n)$ für n positive Ausgänge aus einer Gesamtheit von N unabhängigen Versuchen. Es sei $w_1 \equiv p$ die Wahrscheinlichkeit für einen positiven Ausgang und $w_2 = 1 - p = q$ die entsprechende Wahrscheinlichkeit für einen negativen Ausgang. Dann ist $W(n)$ offenbar durch diejenige Teilsumme von

$$W(n) = \sum_{i=1}^2 \sum_{j=1}^2 \sum_{k=1}^2 \cdots \sum_{m=1}^2 w_i w_j w_k \cdots w_m \quad (1)$$

gegeben, in der bei jedem Summanden w_1 n -mal als Faktor auftritt (denn jeder einzelne Summand in (1) gibt die Wahrscheinlichkeit für eine bestimmte Kombination von positiven und negativen Ausgängen an). Man berechne diese Teilsumme, indem man unter Benutzung elementarer Eigenschaften von Mehrfachsummen (1) so umformt, daß sich das Binomialtheorem anwenden, d.h. eine Entwicklung nach Potenzen von w_1 durchführen läßt.

- 1.8 Zwei Betrunkene starten gemeinsam vom Ursprung der x -Achse aus und bei beiden ist die Wahrscheinlichkeit für einen Schritt nach rechts genauso groß wie die für einen Schritt nach links. Man bestimme die Wahrscheinlichkeit dafür, daß sie sich nach N Schritten wieder treffen. Dabei wird natürlich vorausgesetzt, daß die Männer ihre Schritte gleichzeitig machen. (Die Betrachtung der Relativbewegung kann hier hilfreich sein.)
- 1.9 Die Wahrscheinlichkeit $W(n)$ dafür, daß ein durch die Wahrscheinlichkeit p charakterisiertes Ereignis bei N Versuchen n -mal eintritt, ist wie gezeigt wurde, durch die Binomialverteilung

$$W(n) = \frac{N!}{n!(N-n)!} p^n (1-p)^{N-n} \quad (1)$$

gegeben. Man betrachte eine Situation, für die die Wahrscheinlichkeit p sehr klein ist ($p \ll 1$) und bei der man an dem Fall $n \ll N$ interessiert ist. [Man beachte, daß, falls N sehr groß ist, $W(n)$ mit $n \rightarrow N$ sehr klein wird, da der Faktor p^n sehr klein für $p \ll 1$ ist.] Es lassen sich dann verschiedene Näherungen durchführen, um (1) in eine einfachere Form zu bringen.

- a) Unter Benutzung von $\ln(1-p) \approx -p$ zeige man, daß $(1-p)^{N-n} \approx e^{-Np}$.

- b) Man zeige, daß $N!/(N-n)! \approx N^n$.
 c) Damit zeige man, daß (1) sich auf

$$W(n) = \frac{\lambda^n}{n!} e^{-\lambda} \quad (2)$$

reduziert, wobei $\lambda \equiv Np$ die mittlere Anzahl der Ereignisse ist. Die Verteilung (2) heißt „Poisson-Verteilung“.

1.10 Man betrachte die Poisson-Verteilung der vorangehenden Aufgabe:

- a) Man zeige, daß die Normierung lautet: $\sum_{n=0}^N W_n = 1$.

(Dabei kann die Summation über n in guter Näherung auf „unendlich“ ausgedehnt werden, da W_n vernachlässigbar klein ist, wenn $n \gtrsim N$.)

- b) Man benutze die Poisson-Verteilung, um \bar{n} zu berechnen.

- c) Man benutze die Poisson-Verteilung, um $(\Delta n)^2 \equiv (n - \bar{n})^2$ zu berechnen.

1.11 Angenommen, einem Schriftsetzer unterlaufen beim Druck eines Buches von 600 Seiten vollkommen zufällig 600 Druckfehler. Man benutze die Poisson-Verteilung, um die Wahrscheinlichkeit dafür zu berechnen,

- a) daß eine Seite keinen Fehler enthält,
 b) daß eine Seite mindestens drei Fehler enthält.

1.12 Man betrachte die α -Teilchen, die von einer radioaktiven Quelle während eines Zeitintervalls t emittiert werden. Man kann sich dieses Zeitintervall in viele kleine Intervalle der Länge Δt unterteilt denken. Da die Zeitpunkte, zu denen α -Teilchen emittiert werden, zufällig sind, ist die Wahrscheinlichkeit für einen radioaktiven Zerfall während eines solchen Zeitintervalls Δt vollständig unabhängig von irgendwelchen Zerfällen zu anderen Zeiten. Man kann sich nun Δt so klein gewählt denken, daß die Wahrscheinlichkeit für mehr als einen Zerfall während der Zeit Δt vernachlässigbar klein ist. Das bedeutet, daß es eine Wahrscheinlichkeit p dafür gibt, daß während eines Zeitintervalls Δt (genau) ein Zerfall stattfindet (mit $p \ll 1$, da Δt klein genug gewählt wurde), und eine Wahrscheinlichkeit $1-p$ dafür, daß während dieser Zeit kein Zerfall stattfindet. Jedes solche Zeitintervall kann dann als ein unabhängiger Versuch angesehen werden, wobei es während der Zeit t insgesamt $N = t/\Delta t$ solcher Versuche gibt.

- a) Man zeige, daß die Wahrscheinlichkeit $W(n)$ für n Zerfälle während der Zeit t durch eine Poisson-Verteilung gegeben ist.
 b) Angenommen die radioaktive Quelle ist so stark, daß die mittlere Anzahl von Zerfällen pro Minute gleich 24 ist. Wie groß ist dann die Wahrscheinlichkeit für n Zerfälle in einem Zeitintervall von 10 Sekunden? Man gebe speziell die numerischen Werte von $W(n)$ für alle Werte von n von 0 bis 8 an.

1.13 Ein Metall verdampft im Vakuum von einem heißen Glühfaden aus. Die Metallatome fallen auf eine Quarzplatte, die sich in einiger Entfernung davon befindet, und bilden dort eine dünne metallische Schicht. Diese Quarzplatte wird auf einer niedrigen Temperatur gehalten, so daß jedes auffallende Me-

tallatom an der Stelle seines Auftreffens verbleibt, ohne sich weiter fortzubewegen. Von den Metallatomen kann angenommen werden, daß sie auf jedes Flächenelement der Platte mit gleicher Wahrscheinlichkeit auftreffen. Man zeige, daß die Anzahl der Metallatome, die sich auf einem Flächenelement der Größe b^2 anhäufen (wo b der Durchmesser der Metallatome ist), näherungsweise nach einer Poisson-Verteilung verteilt ist. Angenommen, man verdampft genügend Metall, damit sich ein Film mit einer mittleren Dicke von 6 Atomschichten bilden kann. Welcher Bruchteil der Untergrundfläche ist dann überhaupt nicht mit Metall bedeckt? Welcher Bruchteil ist mit Metall einer Dicke von 3 Atomschichten und welcher Bruchteil mit Metall einer Dicke von 6 Atomschichten bedeckt?

- 1.14 Eine Münze wird 400-mal geworfen. Wie groß ist die Wahrscheinlichkeit dafür, daß 215-mal als Ergebnis die „Zahl“ erscheint?
- 1.15 Ein Satz von Telefonleitungen soll installiert werden, um die Stadt A mit der Stadt B zu verbinden. Die Stadt A hat 2000 Telephone. Wenn jeder Teilnehmer von A seine eigene Direktverbindung nach B beanspruchen würde, wären 2000 Telefonleitungen notwendig. Dies wäre ziemlich verschwenderisch. Angenommen, daß während der Hauptgeschäftszeit jeder Teilnehmer in A im Durchschnitt eine Telefonverbindung nach B für zwei Minuten benötigt und daß diese Telefongespräche ganz und gar zufällig geführt werden. Wie groß ist dann die Minimalzahl M von Telefonleitungen nach B , die notwendig ist, damit höchstens 1 Prozent der Anrufer aus A bei ihrer Wahl nach B nicht sofort durchkommen? (Hilfestellung: Man approximiere die Verteilung durch eine Gauß-Verteilung, um die Rechnung zu erleichtern.)
- 1.16 In einem Behälter mit dem Volumen V_0 befindet sich ein Gas von N_0 nicht-wechselwirkenden Molekülen. Man betrachte in diesem Behälter irgendein Teilvolumen V und bezeichne mit N die Anzahl der darin befindlichen Moleküle. Jedes Molekül befindet sich an jeder Stelle des Behälters mit gleicher Wahrscheinlichkeit; somit ist die Wahrscheinlichkeit, daß sich ein gegebenes Molekül innerhalb des Teilvolumens V befindet, einfach gegeben durch V/V_0 .
- Wie groß ist die mittlere Anzahl N von Molekülen in V ? Man drücke das Ergebnis durch N_0 , V_0 und V aus.
 - Man bestimme das relative Schwankungsquadrat $\overline{(N - \bar{N})^2}/\bar{N}^2$ der in V befindliche Anzahl von Molekülen. Man drücke das Ergebnis durch \bar{N} , V und V_0 aus.
 - Was wird aus dem Ergebnis von b), wenn $V \ll V_0$?
 - Welchen Wert sollte das Schwankungsquadrat $(N - \bar{N})^2$ annehmen, wenn $V \rightarrow V_0$? Stimmt das Ergebnis von b) mit dieser Erwartung überein?
- 1.17 Angenommen, daß für das in der vorangehenden Aufgabe betrachtete Volumen gilt: $0 \ll V/V_0 \ll 1$. Wie groß ist die Wahrscheinlichkeit dafür, daß die Anzahl von Molekülen in diesem Volumen zwischen N und $N + dN$ liegt?
- 1.18 Ein Molekül legt in einem Gas – in jede Richtung mit gleicher Wahrscheinlichkeit – gleiche Strecken l zwischen seinen Zusammenstößen mit anderen Molekülen zurück. Wie groß ist nach insgesamt N Einzelverschiebungen die

mittlere quadratische Gesamtverschiebung $\overline{R^2}$ des Moleküls von seinem Ausgangspunkt?

- 1.19 Eine Batterie mit einer EMK V ist mit einem Widerstand R verbunden. Als Ergebnis wird in diesem Widerstand die Leistung $P = V^2/R$ verbraucht. Die Batterie selbst besteht aus N in Reihe geschalteten Einzelzellen, so daß V gerade die Summe der EMKs dieser Einzelzellen ist. Die Batterie ist jedoch alt, so daß nicht alle Zellen ganz in Ordnung sind. Es besteht deshalb nur eine Wahrscheinlichkeit p , daß die EMK irgendeiner Einzelzelle ihren Normalwert v hat und eine Wahrscheinlichkeit $1 - p$, daß sie aufgrund eines inneren Kurzschlusses den Wert Null hat. Die Einzelzellen sind alle voneinander statistisch unabhängig. Man berechne unter diesen Bedingungen die mittlere, im Widerstand verbrauchte Leistung \overline{P} und drücke das Ergebnis durch N , v und p aus.
- 1.20 Man betrachte N gleiche Antennen, die linear polarisierte elektromagnetische Strahlung der Wellenlänge λ und der Geschwindigkeit c aussenden. Die Antennen sind entlang der x -Achse mit einem gegenseitigen Abstand λ angeordnet. Ein Beobachter befindet sich auf der x -Achse in großer Entfernung von den Antennen. Wenn eine *einzelne* Antenne strahlt, mißt der Beobachter eine Intensität I (d.h. die mittlere quadratische E -Feldstärkenamplitude).
- a) Wie groß ist die vom Beobachter gemessene Gesamtintensität, wenn alle Antennen durch denselben Generator mit der Frequenz $\nu = c/\lambda$ in Phase gespeist werden?
- b) Wie groß ist die vom Beobachter gemessene mittlere Intensität, wenn die Antennen alle mit derselben Frequenz $\nu = c/\lambda$, aber mit vollständig zufälligen Phasen strahlen?
(Hinweis: Man stelle die Amplituden durch Vektoren dar und leite die beobachtete Intensität von der resultierenden Amplitude ab.)
- 1.21 Neuerdings sind Radarsignale von dem Planeten Venus reflektiert worden. Angenommen, daß bei einem solchen Experiment ein elektromagnetisches Signal der Dauer τ von der Erde zur Venus gesendet wird. Eine Zeitspanne t später (die das Signal benötigt, um von der Erde zur Venus und zurück zu gelangen) wird die Empfangsantenne auf der Erde für eine Zeit τ eingeschaltet. Das zurückkehrende Echo sollte dann von dem Meßinstrument, das am Ausgang der elektronischen Anlage hinter der Empfangsantenne angebracht ist, als äußerst schwaches Signal mit einer bestimmten Amplitude a_s verzeichnet werden. Außerdem wird aber (herrührend von den unvermeidlichen Schwankungen im Strahlungsfeld außerhalb der Erde und vorallem den Stromschwankungen in dem äußerst empfindlichen Empfangssystem selbst) von dem Meßinstrument ein Zufallssignal mit einer Amplitude a_n registriert. Das Meßinstrument registriert somit eine Gesamtamplitude $a = a_s + a_n$. Obwohl im Mittel $\bar{a}_n = 0$ ist, da a_n mit gleicher Wahrscheinlichkeit positiv oder negativ ist, besteht eine beträchtliche Wahrscheinlichkeit dafür, daß a_n Werte annimmt, die weit höher liegen als die von a_s ; d.h. die Größe $(\overline{a_n^2})^{1/2}$ kann weitaus größer sein als das interessierende Signal a_s . Angenommen, daß

$(\overline{a_n^2})^{1/2} = 1000 a_s$. Dann ruft das schwankende Signal a_n ein Rauschen hervor, welches eine Beobachtung des erwünschten Echsignals praktisch unmöglich macht.

Angenommen, daß N solche Radarsignale hintereinander ausgesandt werden und daß die Gesamtamplituden a , die von der Meßapparatur nach dem Signal aufgefangen werden, alle aufaddiert sind, bevor sie von dem Meßgerät angezeigt werden. Die resultierende Amplitude muß dann die Form $A = A_s + A_n$ haben, wo A_n die resultierende Rauschamplitude (mit $\overline{A_n} = 0$) und $\overline{A} = A_s$ die resultierende Echo-Signalamplitude bedeuten. Wieviele Signale müssen ausgesandt werden, bevor $(\overline{A_n^2})^{1/2} = A_s$ ist, so daß das Echsignal feststellbar wird?

- 1.22 Man betrachte das eindimensionale Problem der Zufallsbewegung mit der Wahrscheinlichkeitsverteilung

$$w(s) ds = (2\pi\sigma^2)^{-1/2} e^{-(s-l)^2/2\sigma^2} ds$$

Wie groß ist nach N Schritten (Einzelverschiebungen)

- a) die mittlere Gesamtverschiebung \bar{x} vom Ursprung?
 - b) das Schwankungsquadrat $(x - \bar{x})^2$?
- 1.23 Man betrachte das Problem der eindimensionalen Zufallsbewegung für ein Teilchen. Angenommen, daß jede einzelne Verschiebung stets positiv ist und mit gleicher Wahrscheinlichkeit irgendwo im Bereich zwischen $l - b$ und $l + b$ liegt, wobei $b < l$. Wie groß ist nach N Einzelverschiebungen
- a) die mittlere Gesamtverschiebung \bar{x} ?
 - b) das Schwankungsquadrat $(x - \bar{x})^2$?
- 1.24 a) Ein Teilchen befindet sich an jeder Stelle auf dem Umfang eines Kreises mit gleicher Wahrscheinlichkeit. Die z -Achse sei irgendeine Gerade in der Ebene des Kreises, die durch seinen Mittelpunkt hindurchgeht und θ sei der Winkel zwischen der z -Achse und der Geraden, die den Mittelpunkt des Kreises und das Teilchen auf seinem Umfang verbindet. Wie groß ist die Wahrscheinlichkeit dafür, daß dieser Winkel zwischen θ und $\theta + d\theta$ liegt?
- b) Ein Teilchen befindet sich an jeder Stelle auf der Oberfläche einer Kugel mit gleicher Wahrscheinlichkeit. Die z -Achse sei irgendeine Gerade durch den Kugelmittelpunkt und θ der Winkel zwischen der z -Achse und der den Kugelmittelpunkt und das Teilchen verbindenden Geraden. Wie groß ist die Wahrscheinlichkeit dafür, daß dieser Winkel zwischen θ und $\theta + d\theta$ liegt?
- 1.25 Man betrachte eine polykristalline Probe von $\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$ in einem äußeren Magnetfeld B in z -Richtung. Das innere Magnetfeld (in z -Richtung), hervorgerufen am Ort eines gegebenen Protons im H_2O Molekül durch das Nachbarproton, ist gegeben durch $(\mu/a^3)(3 \cos^2 \theta - 1)$, falls der Spin dieses Nachbarprotons in Richtung des angelegten Feldes zeigt; es ist gegeben durch $-(\mu/a^3)(3 \cos^2 \theta - 1)$, falls der Spin des benachbarten Protons in die entgegengesetzte Richtung zeigt. Hier ist μ das magnetische Moment des Protons und a die Entfernung zwischen den beiden Protonen, während θ den Winkel zwischen der Verbindungslinie der beiden Protonen und der z -Achse bezeich-

net. In dieser Probe von zufällig orientierten Kristallen befindet sich das Nachbarproton mit gleicher Wahrscheinlichkeit an jeder Stelle auf der Kugel mit dem Radius a , die das gegebene Proton umgibt.

a) Wie groß ist die Wahrscheinlichkeit $W(b) db$ dafür, daß das innere Feld b zwischen b und $b + db$ liegt, wenn der Spin des Nachbarprotons parallel zu B ist?

b) Wie groß ist die Wahrscheinlichkeit $W(b) db$, wenn der Spin des Nachbarprotons mit gleicher Wahrscheinlichkeit entweder parallel oder antiparallel zu B ist? Man skizziere $W(b)$ als Funktion von b .

(Bei einem kernmagnetischen Resonanzexperiment ist die Frequenz, bei der Energie von einem radiofrequenten Magnetfeld absorbiert wird, proportional zur lokalen magnetischen Feldstärke, die am Ort eines Protons herrscht. Die Antwort von Teil b) gibt deshalb die Gestalt der Absorptionslinie, die man bei diesem Experiment beobachtet.)

- *1.26 Man betrachte das Problem der eindimensionalen Zufallsbewegung und nehme an, daß die Wahrscheinlichkeit für eine Einzelverschiebung zwischen s und $s + ds$ gegeben ist durch

$$w(s) ds = \frac{1}{\pi} \frac{b}{s^2 + b^2} ds$$

Man berechne die Wahrscheinlichkeit $\mathcal{P}(x) dx$ dafür, daß die Gesamtverschiebung nach N Schritten zwischen x und $x + dx$ liegt.

- 1.27 Man betrachte das sehr allgemeine Problem einer eindimensionalen Zufallsbewegung, bei dem die Wahrscheinlichkeit dafür, daß die i -te Verschiebung zwischen s_i und $s_i + ds_i$ liegt, gegeben ist durch $w_i(s_i) ds_i$. Hier kann also die Wahrscheinlichkeitsdichte w_i von Schritt zu Schritt verschieden und somit von i abhängig sein. Es gilt jedoch immer noch, daß verschiedene Verschiebungen statistisch unabhängig sind, d.h. bei einer beliebigen Verschiebung hängt w_i nicht von irgendwelchen anderen Verschiebungen ab, die das Teilchen erfahren hat. Man führe ähnliche Überlegungen an, wie im Abschn. 1.11, um zu zeigen, daß die Wahrscheinlichkeit $\mathcal{P}(x) dx$ sich nach wie vor der Gaussischen Form nähert mit einem Mittelwert $\bar{x} = \Sigma \bar{s}_i$ und einem Schwankungsquadrat $(\Delta x)^2 = \Sigma (\Delta s_i)^2$, falls die Anzahl N der Einzelverschiebungen sehr groß wird. Dieses Ergebnis bildet eine sehr allgemeine Form des zentralen Grenzwertsatzes.

- *1.28 Man betrachte die dreidimensionale Zufallsbewegung eines Teilchens und bezeichne mit $w(s) d^3s$ die Wahrscheinlichkeit dafür, daß seine Verschiebung s im Gebiet zwischen s und $s + ds$ liegt (d.h. daß s_x zwischen s_x und $s_x + ds_x$ liegt, daß s_y zwischen s_y und $s_y + ds_y$ liegt, und daß s_z zwischen s_z und $s_z + ds_z$ liegt). $\mathcal{P}(r) d^3r$ bezeichne die Wahrscheinlichkeit dafür, daß die Gesamtverschiebung r des Teilchens nach N Einzelverschiebungen im Gebiet zwischen r und $r + dr$ liegt. Man zeige durch Verallgemeinerung der Überlegungen von Abschnitt 1.10 auf drei Dimensionen, daß

$$\mathcal{P}(\mathbf{r}) = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} d^3\mathbf{k} e^{-i\mathbf{k}\cdot\mathbf{r}} Q^N(\mathbf{k})$$

wobei $Q(\mathbf{k}) = \int_{-\infty}^{\infty} d^3\mathbf{s} e^{i\mathbf{k}\cdot\mathbf{s}} w(\mathbf{s})$

- *1.29 a) Unter Benützung einer geeigneten Diracschen Deltafunktion bestimme man die Wahrscheinlichkeitsdichte $w(\mathbf{s})$ für Verschiebungen mit einheitlicher Länge aber zufälliger Richtung im Raum.
[Hinweis: Man denke daran, daß $w(\mathbf{s})$ so beschaffen sein muß, daß $\iiint w(\mathbf{s}) d\mathbf{s} = 1$, wenn über den ganzen Raum integriert wird.]
- b) Man benutze das Resultat von Teil a), um $Q(\mathbf{k})$ zu berechnen. (Man führe die Integration in Kugelkoordinaten aus.)
- c) Man benutze diesen Wert von $Q(\mathbf{k})$ und berechne damit für $N = 3$ die Größe $\mathcal{P}(\mathbf{r})$.

2. Statistische Beschreibung von Vielteilchensystemen

Systemzustände klassischer Systeme werden durch $2f$ Koordinaten beschrieben, nämlich durch f generalisierte Koordinaten q_k und durch weitere f generalisierte Impulse p_k , wobei f die Anzahl der Freiheitsgrade des Systems ist. Werden die q_k und die p_k als kartesische Koordinaten aufgefaßt, so spannen sie einen $2f$ -dimensionalen Raum auf, der Γ -Raum oder (großer) Phasenraum genannt wird. Systemzustände quantenmechanischer Systeme werden durch die Werte (n_1, n_2, n_3, \dots) , n_j ganzzahlig, ihrer Quantenzahlen n_j charakterisiert. Ein statistisches Ensemble besteht aus einer Menge von identischen Systemen, die den gleichen makroskopischen Nebenbedingungen unterworfen sind und die sich in verschiedenen mikroskopischen Zuständen befinden. Diese Zustände sind mit den makroskopischen Nebenbedingungen verträglich und werden „die dem System zugänglichen Zustände“ genannt. Das grundlegende Postulat der Statistischen Physik lautet: „Im Gleichgewicht wird ein isoliertes System durch ein Ensemble charakterisiert, das über die dem System zugänglichen Zustände im Γ -Raum gleichverteilt ist.“ Daher werden Nichtgleichgewichtssysteme und nicht isolierte Systeme durch nichtgleichverteilte Ensembles dargestellt. Der in einem Nichtgleichgewichtssystem ablaufende irreversible Prozeß wird damit durch den Übergang des nichtgleichverteilten Ensembles in ein gleichverteiltes beschrieben. Die Zustandsdichte eines Systems, d.h. die Anzahl seiner Zustände in einer Energieschale bezogen auf ihre Dicke, ist eine überaus schnell wachsende Funktion der Energie ($\sim E^f$, f Anzahl der Freiheitsgrade des Systems). Makroskopische Systeme können in thermischer und mechanischer Wechselwirkung miteinander stehen. Der Energieaustausch bei thermischer Wechselwirkung verändert die Besetzungen der Energieniveaus, die bei dieser Wechselwirkung unverändert bleiben. Dagegen ändert der Energieaustausch bei mechanischer Wechselwirkung die Lage der Energieniveaus ab. Geschieht dies hinreichend langsam, so wird dieser Prozeß „quasi-statisch“ genannt. Die Energieänderung bei einem infinitesimalen Prozeß ist stets ein vollständiges Differential. Die infinitesimale Arbeit und der infinitesimale Wärmeübergang sind im allgemeinen keine vollständige Differentiale, es sei denn bei rein mechanischer bzw. rein thermischer Wechselwirkung.

Nachdem wir uns mit einigen elementaren statistischen Ideen vertraut gemacht haben, sind wir bereit, uns dem Hauptanliegen dieses Buches zuzuwenden: Die Besprechung der Systeme, die aus sehr vielen Teilchen bestehen. Bei der Untersuchung solcher Systeme werden wir einige statistische Ideen mit unserer Kenntnis von den Gesetzen der Mechanik kombinieren, denen diese Teilchen unterliegen. Diese Art der Betrachtung bildet die Basis der Disziplin „Statistische Mechanik“. Die Methode ist jener ähnlich, die wir bei der Besprechung des Glücksspiels benutzt haben. Um die Analogie klar zu machen, betrachtet man z.B. ein System, das aus 10 Würfeln besteht, die in einem Experiment (Wurf) aus einem Becher auf einen Tisch geworfen werden. Die wesentlichen Fakten, die für eine Beschreibung dieser Situation notwendig sind, sind die folgenden:

1. *Beschreibung des Systemzustandes:* Man benötigt eine umfassende Methode, das Ergebnis eines jeden Experimentes eindeutig zu beschreiben; z.B. in diesem Falle erfordert die Beschreibung des Systemzustandes nach dem Wurf die Angabe, welche Zahl jeder der 10 Würfel zeigt.

2. *Statistisches Ensemble:* Im Prinzip könnte das Problem im folgenden Sinne vorherbestimmt sein: Falls wir wirklich alle Anfangslagen und Orientierungen und die zugehörigen Geschwindigkeiten aller Würfel im Becher am Anfang des Versuchs kennen würden, und wenn wir weiter die Behandlung des Bechers während des Wurfes wüßten, dann könnten wir tatsächlich den Ausgang des Experimentes durch die Anwendung der Gesetze der klassischen Mechanik und die Lösung der zugehörigen Differentialgleichungen vorhersagen. Aber solche genauen Informationen über den Anfangszustand des Systems sind uns nicht zugänglich. Deshalb werden wir so vorgehen, daß das Experiment mit wahrscheinlichkeitstheoretischen Methoden beschrieben werden soll. Das bedeutet, daß wir anstelle eines einzelnen Experimentes ein Ensemble betrachten, das aus vielen solcher Experimente besteht, die alle unter vergleichbaren Bedingungen ausgeführt werden (nach einigem Schütteln werden 10 Würfel aus dem Becher geworfen). Die Ergebnisse eines jeden Experimentes werden im allgemeinen unterschiedlich sein. Aber wir können nach der *Wahrscheinlichkeit* fragen, mit der ein bestimmtes Ergebnis auftritt, d.h. wir können den Bruchteil der Fälle aller Experimente bestimmen, der durch einen speziellen Endzustand der Würfel gekennzeichnet ist. Dieses Verfahren zeigt, wie die Wahrscheinlichkeit experimentell bestimmt wird ¹⁺⁾. Unser theoretisches Ziel ist es, auf der Grundlage einiger Postulate diese Wahrscheinlichkeit vorauszusagen.

¹⁺⁾ Unter gleichen Bedingungen werden hinreichend viele Einzelversuche durchgeführt, deren Ergebnis trotz gleicher Ausführungsbedingungen verschieden ausfallen. Sind M Versuchsausgänge möglich und werden N Versuche durchgeführt, so heißt $h_j = N_j/N$ die relative Häufigkeit, mit der der j -te Versuchsausgang auftritt, wenn N_j die Anzahl der Versuche darstellt, für die der j -te Versuchsausgang auftritt. Es gilt $\sum_{j=1}^M N_j = N$ und $\sum_{j=1}^M h_j = 1$. h_j ist im allgemeinen eine Funktion der Versuchsanzahl: $h_j = h_j(N)$. Wird die Versuchsanzahl unter Beibehaltung aller Versuchsbedingungen erhöht, so heißt der Grenzwert $\lim_{N \rightarrow \infty} h_j(N) = W_j$ die Wahrscheinlichkeit dafür, daß der j -te Versuchsausgang eintritt.

3. *Grundlegendes Postulat über die a-priori-Wahrscheinlichkeiten:* Um theoretisch voranzukommen, müssen wir einige fundamentale Postulate einführen. Unsere Kenntnis der physikalischen Situation läßt uns aufgrund der Gesetze der Mechanik erwarten, daß für einen regulären Würfel einheitlicher Dichte nichts für das bevorzugte Erscheinen einer bestimmten Zahl des Würfels gegenüber allen anderen Zahlen spricht. Deshalb können wir das Postulat einführen, daß *a priori* (d.h. aufgrund unserer vorhergehenden Ansicht, die bis jetzt durch Beobachtungen noch nicht bestätigt ist) die Wahrscheinlichkeiten dafür, daß irgendeine der sechs Oberflächen eines Würfels nach dem Wurf oben liegt, gleich sind. Dieses Postulat ist außerordentlich einleuchtend und widerspricht keinem Gesetz der Mechanik. Ob das Postulat wirklich zutrifft, kann nur dadurch entschieden werden, daß theoretische Voraussagen, die mit seiner Hilfe gemacht wurden, durch experimentelle Beobachtungen geprüft und bestätigt werden. Werden solche Voraussagen wiederholt bestätigt, so kann die Gültigkeit dieses Postulates mit wachsendem Vertrauen angenommen werden.

4. *Wahrscheinlichkeitsrechnung:* Wenn das grundlegende Postulat bestätigt worden ist, erlaubt die Wahrscheinlichkeitstheorie die Berechnung der Wahrscheinlichkeit für den Ausgang eines jeden Experimentes mit diesen Würfeln.

Beim Studium der Vielteilchensysteme werden unsere Betrachtungen gleich jenen sein, die wir im vorstehenden Problem der 10 Würfel benutzt haben.

Statistische Formulierung des mechanischen Problems

2.1 Beschreibung des Systemzustandes

Es wird irgendein System von Teilchen betrachtet, wie kompliziert auch immer (z.B. eine Menge schwach gekoppelter, harmonischer Oszillatoren, ein Gas, eine Flüssigkeit, ein Automobil). Wir wissen, daß diese Teilchen in einem solchen System (z.B. die Elektronen, Atome oder Moleküle, die das System bilden) durch die Gesetze der Quantenmechanik beschrieben werden können. Insbesondere kann das System durch eine Wellenfunktion $\psi(q_1, \dots, q_f)$ beschrieben werden, die eine Funktion von f Koordinaten (einschließlich möglicher Spinvariablen) ist, die zur Charakterisierung des Systems erforderlich sind. Die Zahl f ist die „Anzahl der Freiheitsgrade“ des Systems. Ein spezieller Quantenzustand des Systems ist durch einen Satz von f Quantenzahlen gegeben. Die Beschreibung ist vollständig, wenn durch die Vorgabe von ψ zu irgendeiner Zeit t die Bewegungsgleichung der Quantenmechanik die Vorhersage von ψ zu irgendeiner anderen Zeit gestattet.

Beispiel 1: Es wird ein System betrachtet, das aus einem einzigen Teilchen an einem festen Ort besteht und das den Spin $\frac{1}{2}$ besitzt (d.h. das Spin-Drehmoment ist $\frac{1}{2}\hbar$). Der Zustand dieses Teilchens ist in quantenmechanischer Beschreibung durch die Projektion m des Spins auf eine beliebige, feste Achse gegeben, die zur z-Achse gewählt werden kann. Die Quantenzahl m kann zwei Werte $m = \frac{1}{2}$ oder $m = -\frac{1}{2}$ annehmen, d.h. grob gesagt, der Spin kann entweder nach „oben“ oder nach „unten“ bezüglich der z-Achse orientiert sein.

Beispiel 2: Es wird ein System aus N Teilchen betrachtet, die sich alle an festen Orten befinden. Jedes Teilchen besitzt den Spin $\frac{1}{2}$. N sei groß, etwa von der Größenordnung der Loschmidtschen Zahl $N_L = 6 \cdot 10^{23} \text{ mol}^{-1}$. Die Quantenzahl m eines jeden Teilchens kann zwei Werte $m = \pm\frac{1}{2}$ annehmen. Der Zustand des Gesamtsystems ist dann durch die Angabe der N Quantenzahlen m_1, \dots, m_N festgelegt, die die Spinorientierung jedes einzelnen Teilchens beschreiben.

Beispiel 3: Es wird ein System betrachtet, das aus einem eindimensionalen harmonischen Oszillator in fester Position x besteht. Die möglichen Quantenzustände dieses Oszillators können durch eine Quantenzahl n numeriert werden, so daß die Energie des Oszillators

$$E_n = (n + \frac{1}{2})\hbar\omega$$

ist. ω ist die klassische Kreisfrequenz der Schwingung. Die Quantenzahl n ist ganzzahlig $n = 0, 1, 2, \dots$

Beispiel 4: Es wird ein System aus N untereinander schwach wechselwirkenden, eindimensionalen harmonischen Oszillatoren betrachtet. Der Quantenzustand dieses Systems kann durch die Menge der Quantenzahlen n_1, \dots, n_N beschrieben werden, wobei n_i zum i -ten Oszillator gehört und jede Quantenzahl ganzzahlig ist.

Beispiel 5: Es wird ein System betrachtet, das aus einem spin- und kräftefreien Teilchen in einem rechtwinkligen Kasten besteht (die Teilchenkoordinaten liegen somit innerhalb des Gebietes $0 \leq x \leq L_x, 0 \leq y \leq L_y, 0 \leq z \leq L_z$). Die Wellenfunktion ψ dieses Teilchens der Masse m muß die Schrödinger-Gleichung

$$-\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \psi = E\psi \quad (2.1.1)$$

erfüllen. Weil das Teilchen den Kasten nicht verlassen kann, muß ψ an den Wänden verschwinden. Die Wellenfunktion mit dieser Randbedingung, die (2.1.1) erfüllt, hat die Form

$$\psi = \sin \left(\pi \frac{n_x x}{L_x} \right) \sin \left(\pi \frac{n_y y}{L_y} \right) \sin \left(\pi \frac{n_z z}{L_z} \right) \quad (2.1.2)$$

Die Energie E des Teilchens ergibt sich zu

$$E = \frac{\hbar^2}{2m} \pi^2 \left(\frac{n_x^2}{L_x^2} + \frac{n_y^2}{L_y^2} + \frac{n_z^2}{L_z^2} \right) \tag{2.1.3}$$

Die Randbedingung $\psi = 0$, wenn $x = 0$ oder $x = L_x$, $y = 0$ oder $y = L_y$, $z = 0$ oder $z = L_z$, erzwingt, daß die drei Zahlen n_x, n_y, n_z ganzzahlig sind. Der Zustand des Teilchens ist durch die Angabe dieser drei ganzzahligen Quantenzahlen gegeben. Die quantisierte Energie eines Zustandes ergibt sich aus (2.1.3).

Bemerkungen zur klassischen Beschreibung. Atome und Moleküle werden in Begriffen der Quantenmechanik beschrieben. Deshalb wird im weiteren bei der Erörterung der Vielteilchensysteme durchgängig die Quantenmechanik als Grundlage benutzt. Eine Beschreibung in Begriffen der klassischen Mechanik kann – obgleich sie im allgemeinen nicht zutreffend ist – manchmal eine nützliche Näherung darstellen. Deshalb ist es angebracht, einige Bemerkungen über die Beschreibung des Systemzustandes im Rahmen der klassischen Mechanik zu machen.

Wir beginnen mit einem sehr einfachen Fall – ein einziges Teilchen in einer Dimension. Dieses System kann durch die Angabe seiner Ortskoordinate q und seines zugehörigen Impulses p vollständig beschrieben werden. (Die Beschreibung ist deshalb vollständig, weil es die Gesetze der klassischen Mechanik gestatten, aus der Kenntnis von q und p zu irgendeiner Zeit die Werte von q und p zu irgendeiner anderen Zeit vorherzusagen.) Diese Situation läßt sich geometrisch dadurch veranschaulichen, indem in einem kartesischen p - q -Achsenkreuz (Abb. 2.1.1) für den Zustand des betrachteten Einteilchensystem ein Punkt gemacht wird. Der zweidimensionale p - q -Raum wird als „Phasenraum“ dieses Systems bezeichnet. Wenn Ortskoordinate und Impuls des Teilchens zeitabhängig sind, bewegt sich der den Teilchenzustand beschreibende Punkt (repräsentativer Punkt) durch diesen Phasenraum.

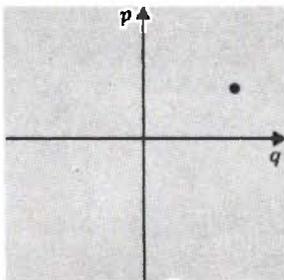


Abb. 2.1.1 Klassischer Phasenraum eines Teilchens in einer Dimension.

Um die Situation so zu beschreiben, daß die möglichen Zustände des Teilchens abzählbar sind, ist es angebracht, das Gebiet der Variablen q und p beliebig in kleine Intervalle einzuteilen; z.B. kann man für die Einteilung von q feste Intervalle der Länge δq wählen, für die von p solche der Länge δp . Somit ist in diesem

Falle der Phasenraum in kleine Zellen gleichen Volumens (hier wegen der Dimension 2 gleicher Flächen)

$$\delta q \delta p = h_0$$

eingeteilt. Dabei ist h_0 eine kleine Konstante der Dimension des Drehimpulses (Wirkung). Der Systemzustand kann nunmehr dadurch beschrieben werden, daß seine Ortskoordinate in einem Intervall zwischen q und $q + \delta q$ und sein Impuls zwischen p und $p + \delta p$ liegt, d.h. durch die Angabe, daß der repräsentative Punkt (q, p) in einer bestimmten Zelle des Phasenraums liegt. Die Beschreibung des Systemzustandes wird selbstverständlich umso genauer je kleiner die Zellengröße der Zellen gewählt wird, in die der Phasenraum eingeteilt wurde, d.h. je kleiner die Größe von h_0 gewählt wird. Natürlich kann in dieser klassischen Beschreibung h_0 beliebig klein gewählt werden.

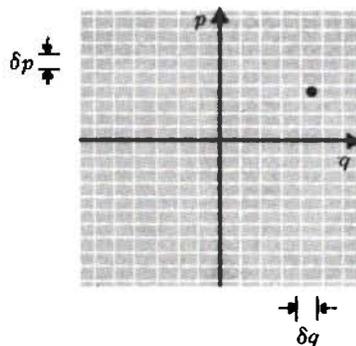


Abb. 2.1.2 Dies ist der in Abb. 2.1.1 gezeigte Phasenraum, der in Zellen gleichen Volumens $\delta q \delta p = h_0$ eingeteilt ist.

Es muß angemerkt werden, daß die Quantentheorie der Genauigkeit, mit der die gleichzeitige Bestimmung der Ortskoordinate q und des Impulses p möglich ist, eine Begrenzung auferlegt. Diese Begrenzung wird durch das Heisenbergsche Unbestimmtheitsprinzip ausgedrückt, das aussagt, daß die Unbestimmtheiten δp und δq derart sind, daß $\delta q \delta p \geq \hbar$ gilt, wobei \hbar die durch 2π dividierte Plancksche Konstante ist ²⁺⁾. Somit ist eine Einteilung des Phasenraums in Zellen, deren Volumen kleiner als \hbar ist, physikalisch bedeutungslos, d.h. eine Wahl von $h_0 < \hbar$ würde zu einer genaueren Beschreibung des Systems führen als es die Quantentheorie erlaubte.

Die Verallgemeinerung der eingangs gemachten Bemerkungen auf ein beliebig komplexes System ist naheliegend. Ein solches System kann durch eine Menge von f Ortskoordinaten q_1, \dots, q_f und f zugehörigen Impulsen p_1, \dots, p_f , d.h. durch $2f$ Parameter (eindeutig) beschrieben werden. Die Anzahl f der voneinander unabhängigen Ortskoordinaten, die zur Beschreibung des Systems benötigt wird, heißt „Anzahl der Freiheitsgrade“ des Systems. (Wenn wir es z.B. mit einem System aus N punktförmigen Teilchen zu tun haben, so ist jedes Teilchen durch seine drei Ortskoor-

²⁺⁾ δq und δp sind die mittleren quadratischen Abweichungen (nach 1.4.9) des Orts- und Impulsoperators.

dinaten charakterisiert, so daß $f = 3N$ ist.) Die Menge der Zahlen $\{q_1, \dots, q_f, p_1, \dots, p_f\}$ kann wieder als ein „Punkt“ in einem „Phasenraum“ von $2f$ Dimensionen angesehen werden, in dem jede kartesische Koordinatenachse durch eine der Orts- oder Impulskoordinaten gekennzeichnet ist. (Obgleich dieser Raum unanschaulich ist, entspricht er völlig dem 2-dimensionalen Raum in Abb. 2.1.1). Dieser Raum kann wieder in kleine Zellen eingeteilt werden. (Wählt man z.B. die Einteilung so, daß wie im 2-dimensionalen Beispiel $\delta q_k \delta p_k = h_0$ für alle k gilt, so ist das Zellenvolumen im $2f$ -dimensionalen Phasenraum $\delta q_1 \dots \delta q_f \delta p_1 \dots \delta p_f = h_0^f$.) Der Systemzustand kann dann wieder durch die Angabe beschrieben werden, in welchem Gebiet oder in welcher Zelle des Phasenraums die Koordinaten $q_1, \dots, q_f, p_1, \dots, p_f$ des Systems zu finden sind.

Zusammenfassung: Der mikroskopische Zustand oder kurz „Mikrozustand“ eines Systems von Teilchen kann einfach auf folgende Art beschrieben werden:

Man nummeriere und indiziere in einer zweckmäßigen (also beliebigen) Reihenfolge alle möglichen Quantenzustände des Systems ($r = 1, 2, 3, \dots$). Der Systemzustand wird dann durch die Angabe der speziellen Zustandsnummer r des Zustandes beschrieben, in dem das System vorgefunden wird.

Falls die Näherung der klassischen Mechanik benutzt werden soll, ist das Vorgehen völlig analog. Nachdem der Phasenraum des Systems in geeignete Zellen gleicher Größe eingeteilt worden ist, kann man diese Zellen in beliebiger Weise mit einem Index r nummerieren ($r = 1, 2, 3, \dots$). Der Systemzustand wird dann durch die Angabe des Index r der Zelle beschrieben, in der sich der repräsentative Punkt des Systems befindet.

Somit sind die quantenmechanische und klassische Beschreibung sehr ähnlich: Eine Zelle im Phasenraum ist das klassische Analogon zum Quantenzustand.

2.2 Statistisches Ensemble

Im Prinzip ist der Zustand eines Vielteilchensystems in dem Sinne völlig determinierend, daß die Angabe der Wellenfunktion ψ des Systems zu irgendeiner Zeit die Berechnung aller physikalischen Eigenschaften des Systems sowie die Vorhersage von ψ zu allen anderen Zeiten gestattet. (Analog ist es in der klassischen Mechanik, in der die Beschreibung des Systemzustands durch die Angabe aller Ortskoordinaten und aller Momente es gestattet, alle physikalischen Eigenschaften des Systems zu berechnen und alle Ortskoordinaten und Momente für alle anderen Zeiten vorauszusagen.) Aber im allgemeinen ist uns weder eine solche vollständige Beschreibung des Systems zugänglich noch sind wir an einer solchen interessiert. Deshalb werden wir das System wahrscheinlichkeitstheoretisch beschreiben. Zu diesem Zweck betrachten wir *nicht* nur ein einziges System, sondern wir denken uns ein Ensemble aus einer sehr großen Anzahl identischer Systeme, die alle irgendwelchen, als bekannt anzunehmenden Bedingungen unterworfen worden wa-

ren³⁺⁾. Im allgemeinen sind die Systeme des Ensembles in verschiedenen Zuständen und werden deshalb auch durch verschiedene makroskopische Parameter charakterisiert (z.B. durch verschiedene Werte des Druckes oder der Magnetisierung). Aber wir können danach fragen, mit welcher Wahrscheinlichkeit ein bestimmter Wert eines solchen Parameters auftritt, d.h. wir können den Anteil der Fälle im Ensemble bestimmen, für den der Parameter diesen bestimmten Wert annimmt. Das Ziel der Theorie besteht darin, die Wahrscheinlichkeit vorauszusagen, mit der im Ensemble verschiedene Werte eines solchen Parameters auftreten, und zwar auf der Grundlage einiger fundamentaler Postulate.

Beispiel: Es wird ein System aus drei unbeweglichen Teilchen betrachtet, von denen jedes den Spin $\frac{1}{2}$ hat, so daß jeder Spin nach oben oder nach unten zeigen kann. (d.h. in Richtung oder entgegengesetzt einer zur z -Achse gewählten Richtung.) Jedes Teilchen hat ein magnetisches Moment μ in Richtung der z -Achse, wenn der Spin nach oben zeigt, ein solches $-\mu$, wenn der Spin nach unten gerichtet ist. Das System wird in ein äußeres Magnetfeld H gebracht, das in die positive Richtung der z -Achse zeigt.

Der Zustand eines jeden Teilchens i kann durch seine magnetische Quantenzahl m_i , die die beiden Werte $m_i = \pm \frac{1}{2}$ annehmen kann, beschrieben werden. Der Zustand des Gesamtsystems wird durch die Werte der drei Quantenzahlen m_1, m_2 und m_3 beschrieben. Ein Teilchen hat die Energie $-\mu H$, wenn sein Spin nach oben zeigt und die Energie μH , wenn der Spin abwärts gerichtet ist.

In der nachfolgenden Tafel sind alle möglichen Zustände des Systems zusammengestellt. Desgleichen einige Parameter, wie das resultierende magnetische Moment und die Gesamtenergie, die das System als Ganzes charakterisieren. (Zur Abkürzung ist $m = \frac{1}{2}$ durch $+$, $m = -\frac{1}{2}$ durch $-$ gekennzeichnet.)

Zustands-Nr. r	Quantenzahlen m_1, m_2, m_3	resultierendes magn. Moment	Gesamtenergie
1	+++	3μ	$-3\mu H$
2	++-	μ	$-\mu H$
3	+ - +	μ	$-\mu H$
4	- + +	μ	$-\mu H$
5	+ - -	$-\mu$	μH
6	- + -	$-\mu$	μH
7	- - +	$-\mu$	μH
8	- - -	-3μ	$3\mu H$

³⁺⁾ Diese makroskopischen Nebenbedingungen sind für die Konstruktion des Ensembles wesentlich. Veränderte Nebenbedingungen ergeben im allgemeinen andere Ensembles

Gewöhnlich besitzt man von dem zu betrachtenden System nur einige Teilkenntnisse (z.B. könnten die Gesamtenergie und das Volumen eines Gases bekannt sein). Das System kann dann nur in einem seiner Zustände sein, die mit der vom System bekannten Information verträglich ist. Diese Zustände werden als die „dem System zugänglichen Zustände“ bezeichnet. In einer statistischen Beschreibung enthält somit das repräsentative Ensemble (die Menge der repräsentativen Punkte im Phasenraum) nur solche Systeme, die alle mit dem speziellen, vorhandenen Wissen über das betrachtete System verträglich sind, d.h. die Systeme im Ensemble müssen alle allein auf die verschiedenen zugänglichen Zustände verteilt sein.

Beispiel: Es wird angenommen, daß in dem vorhergehenden Beispiel des Systems, das aus den drei Spins besteht, die Gesamtenergie bekannt und gleich $-\mu H$ sei. Falls dies die einzige zugängliche Information ist, kann das System nur in einem der drei folgenden Zustände sein (vgl. Tabelle):

(+ + -) (+ - +) (- + +).

Natürlich wissen wir nicht, in welchem dieser Zustände sich das System wirklich befindet, noch kennen wir die unentbehrliche relative Wahrscheinlichkeit, das System in einem jeden dieser Zustände zu finden.

2.3 Grundlegende Postulate

Um in den theoretischen Betrachtungen voranzukommen, ist es nötig, einige Postulate darüber einzuführen, wie groß die relativen Wahrscheinlichkeiten sind, ein System in irgendeinem seiner zugänglichen Zustände vorzufinden. Es wird zunächst angenommen, daß das betrachtete System *isoliert* sei und daher keine Energie mit seiner Umgebung austauschen kann. Gemäß den Gesetzen der Mechanik ist dann die Gesamtenergie des Systems eine Erhaltungsgröße. Daraus folgt, daß das System stets durch diesen Wert der Energie charakterisiert sein muß und daß die zugänglichen Zustände des Systems alle diese Energie besitzen müssen. Im allgemeinen gibt es eine große Menge solcher Zustände, und das System kann in irgendeinem von ihnen sein. Was kann man nun über die relative Wahrscheinlichkeit aussagen, das System in irgendeinem dieser zugänglichen Zustände vorzufinden? ⁴⁺⁾

Man kann hoffen, einige allgemeine Aussagen für den einfachen Fall eines isolierten Systems im *Gleichgewicht* zu machen. Eine solche Gleichgewichtssituation ist dadurch ausgezeichnet, daß die Wahrscheinlichkeit dafür, das System in irgendeinem zugänglichen Zustand vorzufinden, zeitunabhängig ist. (d.h. das repräsentative Ensemble ist unabhängig von der Zeit stets dasselbe.) Alle makroskopischen

⁴⁺⁾ Das ist wie folgt zu verstehen: Eine große Menge gleichartiger isolierter Systeme gleicher Energie wird betrachtet. Dann wird der Zustand jedes dieser Systeme durch Messung ermittelt. Wie viele dieser Systeme waren in einem bestimmten, vorgegebenen Zustand?

Parameter, die zum isolierten System gehören, sind dann ebenfalls zeitunabhängig. Wenn man ein solches isoliertes System im Gleichgewicht betrachtet, so besteht die einzige Information, die über das System vorhanden ist darin, daß es sich in einem seiner zugänglichen Zustände befindet, die mit dem Wert seiner Energie verträglich sind. Es läßt sich nichts in den Gesetzen der Mechanik finden, daß Anlaß zu der Erwartung geben könnte, das System befinde sich häufiger in einem bestimmten seiner zugänglichen Zustände als in irgendeinem anderen. Daher erscheint es außerordentlich einleuchtend anzunehmen, daß das System in jedem seiner zugänglichen Zustände gleich wahrscheinlich vorgefunden wird ⁵⁺⁾. Tatsächlich kann man mit den Gesetzen der Mechanik explizit zeigen, daß aus der Annahme, das repräsentative Ensemble eines solchen isolierten Systems sei zu *irgendeiner* Anfangszeit gleichmäßig (mit gleicher Wahrscheinlichkeit) über seine zugänglichen Zustände verteilt, die gleichmäßige Verteilung zu *allen Zeiten* folgt ^{6*)}. Dies zeigt, daß eine solche gleichmäßige Verteilung der Systeme des Ensembles über ihre zugänglichen Zustände in der Tat einer möglichen Gleichgewichtssituation entspricht, die sich zeitlich nicht ändert. Es liegt also nahe, daß es nichts in den Gesetzen der Mechanik gibt, das einige Zustände auf Kosten anderer vorzieht, weil (nach dem Liouvilleschen Satz) keine Tendenz besteht, die gleichmäßige Verteilung dadurch zu zerstören, daß einige Zustände bevorzugt besetzt und andere entleert werden.

Die vorstehenden Betrachtungen legen es nahe, daß alle zugänglichen Zustände eines isolierten Systems die gleiche Besetzungswahrscheinlichkeit besitzen. Somit wird man zur Einführung des folgenden grundlegenden Postulates über die gleichen a priori-Wahrscheinlichkeiten geführt:

► Ein isoliertes System im Gleichgewicht ist gleichwahrscheinlich in jedem seiner zugänglichen Zustände ⁷⁺⁾.

Dasselbe Postulat wird im Fall der klassischen Mechanik aufgestellt, in dem die Zustände einer Zelle im Phasenraum entsprechen. Das bedeutet bei einer Einteilung des Phasenraums in kleine Zellen gleicher Größe, daß ein isoliertes System im Gleichgewicht gleichwahrscheinlich in jeder seiner zugänglichen Phasenzellen ist ^{8*)}.

Dieses grundlegende Postulat ist außerordentlich einsichtig und widerspricht zweifellos keinem der Gesetze der Mechanik. Ob das Postulat wirklich gilt, kann na-

⁵⁺⁾ D.h., die in der Fußnote auf Seite 56 eingeführten Wahrscheinlichkeiten W_j , $j = 1, 2, \dots, M$, hängen *nicht* vom Index j ab, wenn dem System M Zustände zugänglich sind.

^{6*)} Dies ist eine Folgerung aus einem Theorem, das als „Liouvillescher Satz“ bezeichnet wird. Ein Beweis dieses Theorem für die klassische Mechanik findet sich im Anhang A.13. Eine ausführlichere Diskussion steht in Tolman, R.C.: *The Principles of Statistical Mechanics*, Kap. 3, Oxford University Press, Oxford, 1938. Eine Erörterung dieses Theorems für die Quantenmechanik kann im Kap. 9 des gleichen Buches gefunden werden.

⁷⁺⁾ Dies ist ein Ausdruck dafür, daß das Ensemble, das das System repräsentiert, gleichmäßig über die dem System zugänglichen Zustände verteilt ist.

^{8*)} Weitere Erläuterungen zu diesem Postulat finden sich am Ende dieses Abschnittes.