
Statistische Physik

Hendrik van Hees
mailto: Hendrik.vanHees@theo.physik.uni-giessen.de
Institut für Theoretische Physik
Justus-Liebig-Universität Gießen
Heinrich-Buff-Ring 16
D-35392 Gießen

19. November 2008

Inhaltsverzeichnis

1	Statistik und Informationstheorie	7
1.1	Ereignisse und Wahrscheinlichkeiten	7
1.2	Zufallsvariable	11
1.3	Vektorwertige Zufallsvariable	13
1.4	Stochastische Prozesse	16
1.5	Information	17
1.6	Erinnerung an die Quantentheorie	21
1.7	Wahl des Bildes der Zeitentwicklung	24
1.8	Formale Lösung der Bewegungsgleichungen	26
1.9	Ein Beispiel: Das freie Teilchen	28
1.10	Quantenstatistik	31
1.11	Gemische und statistische Operatoren	31
1.12	Die Entropie und das Jaynessche Prinzip	34
1.13	Beispiel: Ort und Impuls	36
1.14	Anhang: Beweis zweier wichtiger Theoreme	38
2	Thermodynamisches Gleichgewicht	43
2.1	Adiabatische Änderungen im thermodynamischen Gleichgewicht	43
2.2	Phänomenologische Thermodynamik	45
3	Ideale Gase im großkanonischen Ensemble	49
3.1	Nichtrelativistische ideale Gase	49
3.1.1	Das ideale Fermigas	55
3.1.2	Der klassische Limes	55
3.1.3	Der Quantenlimes	56
3.1.4	Das ideale Bosegas	58
	Literaturverzeichnis	63

Inhaltsverzeichnis

Einführung

Mit diesem Skript versuche ich eine in sich geschlossene Behandlung der statistischen Physik. Es wird vorausgesetzt, daß der Leser mit den Grundlagen der Quantentheorie vertraut ist. Eine darauf aufbauende Erklärung der Vielteilchentheorie im Feldoperatorformalismus ist allerdings enthalten.

Die Grundlegende Idee, systematisch statistische Konzepte in der Physik anzuwenden geht vor allem auf *Boltzmann* im 19. Jh. zurück. Die wesentliche Bestätigung der atomistischen Struktur der Materie (und auch der Felder) ist allerdings eine Errungenschaft des 20. Jahrhunderts, vor allem die berühmten Arbeiten Einsteins über Schwankungserscheinungen (Brownsche Molekularbewegung, kritische Opaleszenz).

Einen wesentlichen Fortschritt in der kohärenten und in sich konsistenten Darstellung der statistischen Physik hat meiner Meinung nach die Anwendung informationstheoretischer Methoden, die auf Shannon (1948) und vor allem Jaynes zurückgehen. Die Idee ist, daß es neben den grundlegenden Axiomen der mathematischen Statistik auch eines Konzeptes bedarf, wie eine Wahrscheinlichkeitsverteilung einem System unter der Voraussetzung, daß bestimmte restringierende Informationen über den Zustand des Systems vorliegen, möglichst objektiv zu wählen ist. Hier hat sich das Jaynesche Prinzip des geringsten Vorurteils als äußerst schlagkräftig erwiesen: Den Begriff der *Entropie* als Maß für die fehlende Information bei gegebener Wahrscheinlichkeitsverteilung für bestimmte Ereignisse von Shannon nutzend wird eine solche Verteilung gesucht, die bei Einschränkung auf die vorliegende Information, in der Physik meist durch Erwartungswerte von Observablen gegeben, die Entropie maximal macht. Wie wir sehen werden, ist dieses Konzept *nicht* auf Gleichgewichtszustände beschränkt, sondern ermöglicht vielmehr eine konsistente Herleitung von *Transportgleichungen*, die (Quanten-) Transportvorgänge zu beschreiben imstande sind.

Dieses Skript ist wie folgt aufgebaut: Im ersten Abschnitt ist eine in sich geschlossene Einführung in die mathematische Statistik und die Informationstheorie enthalten. Im zweiten Kapitel werden diese Konzepte auf die Quantenstatistik angewendet. Im dritten Kapitel werden zunächst die allgemeinen Aussagen der *phänomenologischen Gleichgewichtsthermodynamik hergeleitet* und an den einfachsten Beispielen idealer Quantengase exemplifiziert.

Wir schließen unsere Betrachtungen mit einer Einführung in die Konzepte der Nichtgleichgewichtsthermodynamik, insbesondere in der Nähe des thermodynamischen Gleichgewichts nebst einigen klassischen Anwendungen zur Theorie der Schwankungserscheinungen wie die Brownsche Bewegung.

Inhaltsverzeichnis

Kapitel 1

Statistik und Informationstheorie

In diesem Kapitel ist eine in sich geschlossene, wenngleich kurze, Einführung in die mathematische Statistik und die Informationstheorie enthalten, auf deren Grundlage im nächsten Kapitel die statistische Physik entwickelt wird.

1.1 Ereignisse und Wahrscheinlichkeiten

Die Statistik versucht, allgemein gesprochen, Ereignisse von Experimenten zu beschreiben, deren Ausgang unvorhersehbar ist. Die Unkenntnis über den Ausgang des Experiments kann dabei verschiedene Ursachen haben, von denen wir aber zunächst abstrahieren wollen. Die grundlegende intuitive Idee ist es, eine Beschreibung der „*mittleren Häufigkeit*“ des Auftretens bestimmter Resultate des Experiments zu erreichen, wenn das Experiment nur hinreichend oft ausgeführt wird.

Diese sehr vagen Ideen haben schließlich im Laufe der Zeit zur Entwicklung der *mathematischen Statistik* auf axiomatischer Grundlage durch Kolmogorov geführt, und wir verfolgen nun das Ziel, diese Grundlage im Hinblick auf konkrete Anwendungen zu verstehen. Dazu bedienen wir uns des klassischen Beispiels eines Zufallsexperiments, nämlich dem Werfen eines Würfels.

Zunächst müssen wir angeben, welche möglichen Resultate ein einzelnes Experiment haben kann. Wir nennen die Menge Ω der möglichen Resultate eines einzelnen Experiments die *Menge der Elementarereignisse*. In unserem Beispiel des Würfelspiels ist die Menge der Elementarereignisse durch die mögliche Augenzahl gegeben, die bei einem Wurf des Würfels herauskommt $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Manchmal kommt es jedoch bei einem Spiel z.B. nur darauf an, ob eine „6“ gewürfelt wurde oder nicht. Es kann auch sein, daß es uns nur interessiert, ob die Zahl gerade ist oder nicht. Solche „Versuchsergebnisse“ bezeichnen wir als *Ereignis*. Jedes Ereignis kann dabei als eine Teilmenge von Ω angesehen werden. So ist etwa das Ereignis, daß eine „6 und nur eine 6“ gefallen sei, durch $E = \{6\}$ gegeben. Das Ereignis „eine gerade Zahl ist gefallen“ wird hingegen durch $E = \{2, 4, 6\}$ charakterisiert. Das jeweilige Ereignis wird durch ein bestimmtes Zufallsexperiment realisiert oder nicht, je nachdem ob das dabei herauskommende Elementarereignis in der Ereignismenge E enthalten ist oder nicht. Die Menge aller möglichen Ereignisse ist demzufolge durch die *Potenzmenge* $\mathcal{P}\Omega$ der Elementarereignisse gegeben.

Wir können nun die elementaren Operationen der Mengenlehre in dem eben definierten Sinne

auf Ereignisse anwenden. Nehmen wir zwei Ereignisse $E_1, E_2 \in \mathcal{P}\Omega$. Dann ist klar, daß das Resultat eines konkreten Zufallsexperiment zu beiden Ereignissen gehört, wenn es in der Schnittmenge $E_1 \cap E_2$ enthalten ist. Gibt man einander widersprechende Ereignisse vor, etwa im Falle unseres Würfels $E_1 = \{1, 3, 5\}$ und $E_2 = \{2, 4, 6\}$, so gelangen wir zu $E_1 \cap E_2 = \emptyset \in \mathcal{P}\Omega$. Das bedeutet, daß die leere Menge das „unmögliche Ereignis“ repräsentiert, d.h. ein Ereignis, das nie eintreten kann.

Es ist klar, daß die Vereinigungsmenge $E_1 \cup E_2$ zweier Ereignisse das Ereignis bedeutet, daß E_1 oder E_2 eingetreten sind.

Weiter kann man auch fragen, ob ein Ereignis E nicht eingetreten sei. Das ist wieder ein Ereignis, welches durch die Komplementmenge $\bar{E} := \Omega \setminus E$ gegeben ist. Insbesondere ergibt die Komplementbildung der leeren Menge selbstverständlich Ω selbst, das „sichere Ereignis“, denn per definitionem ist jedes mögliche Resultat eines Zufallsexperiments in Ω enthalten, d.h. bei jedem Zufallsexperiment tritt sicher das Ereignis Ω ein.

Für den formalen Aufbau der Statistik ist es nun wichtig, daß die elementaren Operationen \cap, \cup und $\bar{}$, allesamt definiert als Abbildungen der Potenzmenge $\mathcal{P}\Omega$ über der Elementarereignismenge in sich eine *Algebra* bilden, d.h. diese sämtlichen Abbildungen vom Typ $\mathcal{P}\Omega \rightarrow \mathcal{P}\Omega$ erfüllen die Gesetze:

$$A \cup B = B \cup A, \quad A \cap B = B \cap A, \quad (1.1)$$

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C), \quad A \cup (B \cap C) = (A \cup B) \cap (A \cup C), \quad (1.2)$$

$$A \cup \emptyset = A, \quad A \cap \Omega = A, \quad (1.3)$$

$$A \cup \bar{A} = \Omega, \quad A \cap \bar{A} = \emptyset \quad (1.4)$$

Wir werden nicht alle grundlegenden Rechenregeln für Rechnungen mit dieser Algebra beweisen, da sie sich mehr oder weniger von selbst verstehen, zumindest wenn man den naiven Mengenbegriff zugrundelegt.

Als zweites wesentliches Element der statistischen Beschreibung eines Zufallsexperiments ist die Wahl der *Ereignisalgebra* \mathcal{E} . Darunter verstehen wir eine Teilmenge von $\mathcal{P}\Omega$, die unter den Operationen \cup, \cap und Komplementbildung abgeschlossen ist. Ist Ω eine Menge mit unendlich vielen Elementen, muß dies nicht bedeuten, daß die Algebra auch unter den Operationen \cup and \cap mit *unendlich vielen Mengen* abgeschlossen ist. Dies setzen wir im folgenden stets voraus. Solche Algebren heißen auch *Borel algebra*.

Jetzt wollen wir ein Maß für die *Wahrscheinlichkeit* des Eintretens eines Ereignisses angeben. Wir werden dabei einfach die *Kolmogorowschen Axiome* zugrundelegen, ohne auf die diversen philosophischen Probleme, die in der Literatur diskutiert werden, einzugehen. Auf die Interpretation für tatsächliche Experimente, deren Ausgang wir nicht vorhersehen können, kommen wir weiter unten noch zurück. Es ist übrigens eine unmittelbare Folgerung dieser Definitionen, daß jede Ereignisalgebra wenigstens Ω und die leere Menge \emptyset enthalten muß.

(P1) Das Wahrscheinlichkeitsmaß ist eine Funktion $P : \mathcal{E} \rightarrow \mathbb{R}_{\geq 0}$,

(P2) $P(\Omega) = 1$,

(P3) Für alle Ereignisse E_1 und E_2 mit $E_1 \cap E_2 = \emptyset$ soll gelten $P(E_1 \cup E_2) = P(E_1) + P(E_2)$.

Die anschauliche Idee, die hinter diesen Axiomen steckt, ist es, ein Maß zur Verfügung zu stellen, welches angibt, wie häufig ein Ereignis bei der Wiederholung eines bestimmten Zufallsexperiments eintritt.

Wie die Wahrscheinlichkeiten für ein bestimmtes Zufallsexperiment zu wählen sind, geben die Axiome allerdings nicht an, und man muß sie aus der Erfahrung gewinnen. Es wird sich zeigen, daß man zur genauen Bestimmung der Wahrscheinlichkeiten $P(E)$ Zufallsexperimente hinreichend oft ausführen muß, wobei es aufgrund von Axiom (P3) ausreicht, nur die einelementigen Ereignisse, also die Elementarereignisse selbst zu messen, d.h. man bestimmt einfach empirisch, wie häufig die Elementarereignisse eintreten. Man kann dann diese empirischen Wahrscheinlichkeiten durch fortgesetzte Anwendung von (P3) benutzen, um die gesamte Funktion P zu bestimmen.

Oft ist es nun aber unmöglich, diese Messung durchzuführen, z.B. weil die Messungen zu aufwendig sind, weil es einfach zu viele Elementarereignisse gibt, oder sie ist zu kompliziert durchzuführen. Dann kann man den Elementarereignissen einfach Schätzwahrscheinlichkeiten zuordnen. Im Falle eines Würfels kann man etwa annehmen, daß $P(\{i\}) = 1/6$ für alle $i \in \{1, \dots, 6\}$. Das bedeutet natürlich nicht, daß diese Wahrscheinlichkeitsverteilung dem tatsächlichen Würfel auch zukommt, und man kann durch reale Zufallsexperimente testen, ob die gewählte Verteilung auf den Würfel zutrifft oder nicht. Wir werden später zeigen, daß man $P(E)$ erhalten kann, wenn man N voneinander unabhängige Zufallsexperimente ausführt und die Anzahl $N(E)$ für das Auftreten der Ereignisse bestimmt. Es wird sich zeigen, daß $P(E)$ durch den Limes $N \rightarrow \infty$ gegeben ist.

Zunächst wollen wir aber einige einfache Regeln über Wahrscheinlichkeiten aus den Axiomen (P1)-(P3) ableiten:

Lemma 1. (1) Für $E \in \mathcal{P}\Omega$ gilt $P(\bar{E}) = 1 - P(E)$. Für alle $E \in \mathcal{P}\Omega$ ist $P(E) \leq 1$

(2) $P(\emptyset) = 0$.

(3) Für beliebige $A, B \in \mathcal{P}\Omega$: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

(4) Die Sylvestersche Formel: Sei $A_i \in \mathcal{P}\Omega$ mit $i = 1, \dots, n$. Dann gilt

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) - \sum_{i,j=1;i < j}^n P(A_i \cap A_j) + \dots + (-1)^{n-1} P(A_1 \cap A_2 \cap \dots \cap A_n). \quad (1.5)$$

Beweis:

(1) folgt sofort daraus, daß $E \cup \bar{E}$ und $E \cap \bar{E} = \emptyset$. Unter Verwendung von P2 und P3 folgt unmittelbar

$$1 = P(\Omega) = P(E \cup \bar{E}) = P(E) + P(\bar{E}). \quad (1.6)$$

Wegen (P1) ist $0 \leq P(\bar{E}) = 1 - P(E)$ und folglich $P(E) \leq 1$.

(2) Aus (1) ist bereits erwiesen, daß

$$P(\emptyset) = 1 - P(\bar{\emptyset}) = 1 - P(\Omega) = 0, \quad (1.7)$$

wobei im letzten Schritt P2 benutzt wurde.

(3) Es gilt $A \cup B = A \cup (B \cap \bar{A})$ und $A \cap (B \cap \bar{A}) = \emptyset$. Wegen (P3) ist also

$$P(A \cup B) = P(A) + P(B \cap \bar{A}). \quad (1.8)$$

Andererseits ist auch $B = (A \cap B) \cup (\bar{A} \cap B)$ und $(A \cap B) \cap (\bar{A} \cap B) = \emptyset$, so daß wir wieder (P3) anwenden können:

$$P(B) = P(A \cap B) + P(\bar{A} \cap B). \quad (1.9)$$

Damit können wir $P(\bar{A} \cap B)$ aus (1.8) eliminieren und erhalten sogleich die Behauptung.

(4) wird durch einen einfachen Induktionsbeweis unter Verwendung von

$$\bigcup_{i=1}^n A_i = \bigcup_{i=1}^{n-1} A_i \cup A_n \quad (1.10)$$

bewiesen. Man möge die einfache Ausführung des Beweises als Übung selbst durchführen.

Jetzt definieren wir *bedingte Wahrscheinlichkeiten*. Die Wahrscheinlichkeit für ein Ereignis $B \in \mathcal{P}\Omega$ unter der Voraussetzung, daß $A \in \mathcal{P}\Omega$ eingetreten ist, ist für $P(A) \neq 0$ definiert als

$$P_A(B) = \frac{P(A \cap B)}{P(A)}. \quad (1.11)$$

Die anschauliche Bedeutung wird klar, wenn wir uns die oben besprochene empirische Bestimmung der Wahrscheinlichkeit vor Augen führen: Wir betrachten nur noch Zufallsexperimente, bei denen das Ereignis A eingetreten ist. Ist N wieder die Gesamtzahl der durchgeführten Experimente, dann ist $P(A) = N(A)/N$, wo $N(A)$ die Zahl angibt, daß A eingetreten ist. Die bedingte Wahrscheinlichkeit, daß B eingetreten ist unter Voraussetzung, daß A eingetreten war, ist dann dadurch gegeben, daß nur noch Ereignisse, für die A eingetreten ist, berücksichtigt werden. B kann also nur eingetreten sein, wenn $B \cap A$ eingetreten ist. Es ist also $P_A(B) = N(B \cap A)/N(A) = N(B \cap A)/N \cdot N/N(A) = P(A \cap B)/P(A)$, wobei immer der Limes $N \rightarrow \infty$ impliziert gedacht wird.

Nun zeigen wir

Lemma 2. Die Funktion $P_A : \mathcal{P}\Omega \rightarrow \mathbb{R}_{\geq 0}$ erfüllt die Kolmogorowschen Axiomen (P1)-(P3).

Beweis:

(P1) Wir müssen zeigen, daß P_A für alle Ereignisse in $[0, 1]$ liegt. Es ist klar, daß $P_A(E) \geq 0$, da sowohl $P(A \cap B) \geq 0$ als auch $P(A) \geq 0$ ist. Daß es auch ≤ 1 ist, kann man aus (P2) und (P3) für die Wahrscheinlichkeitsfunktion P schließen, weil jedes Ereignis aus disjunkte Vereinigung elementarer Ereignisse geschrieben werden kann.

(P2) $P_A(\Omega) = P(A \cap \Omega)/P(A) = P(A)/P(A) = 1$.

(P3) Sei $B \cap C = \emptyset$. Dann gilt

$$P_A(B \cup C) = \frac{P[A \cap (B \cup C)]}{P(A)} = \frac{P[(A \cap B) \cup (A \cap C)]}{P(A)}. \quad (1.12)$$

Wegen $(A \cap B) \cap (A \cap C) = A \cap (B \cap C) = A \cap \emptyset = \emptyset$ folgt die Behauptung unter Verwendung von (P3) für die Wahrscheinlichkeiten P .

Definition 1. Zwei Ereignisse $A, B \in \mathcal{P}\Omega$ heißen genau dann *stochastisch unabhängig*, wenn

$$P(A \cap B) = P(A)P(B) \quad (1.13)$$

gilt.

Es ist wichtig, stets in Erinnerung zu behalten, daß stochastische Unabhängigkeit anhand der Wahrscheinlichkeitsfunktion P definiert ist und nicht eine Eigenschaft der Ereignisse als Menge ist.

1.2 Zufallsvariable

In der Physik streben wir eine *quantitative Beschreibung* von Systemeigenschaften, also Ereignissen im Sinne der obigen statistischen Theorie, an. Dazu wird das Konzept der *Zufallsvariablen* in die mathematische Statistik eingeführt. Es handelt sich dabei um eine Abbildung $\xi : X \subseteq \mathcal{P}\Omega \rightarrow \mathbb{R}$. Ein einfaches Beispiel für unser Würfelspiel ist etwa die Summe der Augenzahlen, die in einem Ereignis enthalten sind.

Ist nun ein Wahrscheinlichkeitsmaß P im Sinne der Kolmogorovschen Axiome definiert, können wir wie folgt eine *Wahrscheinlichkeitsfunktion* für die Zufallsvariablen definieren:

$$P_\xi : D_\xi \rightarrow \mathbb{R} : P_\xi(x) = P(\{r_i | \xi(r_i) \leq x\}). \quad (1.14)$$

Dabei ist D_ξ die Menge der möglichen Werte der Zufallsvariablen ξ . Weiter definieren wir die *Wahrscheinlichkeitsverteilung* als die schwache Ableitung der Wahrscheinlichkeitsfunktion im Sinne der Distributionentheorie.

Nehmen wir als Beispiel wieder das Würfelspiel und definieren eine Zufallsvariable $\xi(\{k\}) = k$. Nehmen wir wieder unsere Standardwahrscheinlichkeitsdefinition, also die Gleichverteilung $P(\{k\}) = 1/6$ für alle $k \in \Omega = \{1, \dots, 6\}$ (und die nach den Kolmogorovschen Axiomen ergänzte Definition für alle nichtelementaren Ereignisse), so erhalten wir für die Wahrscheinlichkeitsfunktion

$$P_\xi(x) = \frac{1}{6} \sum_{i=1}^6 \Theta(x - i), \quad (1.15)$$

wobei wir die *Heavisidesche Einheitstreppe* durch

$$\Theta : \mathbb{R} \rightarrow \mathbb{R} : \Theta(x) = \begin{cases} 0 & \text{für } x < 0 \\ 1 & \text{für } x \geq 0, \end{cases} \quad (1.16)$$

definiert haben. Die schwache Ableitung der Θ -Funktion ist nun die δ -Distribution, wobei wir als Testfunktionenraum den Raum der beliebig oft stetig differenzierbaren Funktionen $C_0^\infty(\mathbb{R})$ zugrundelegen wollen. Das ist der Raum der beliebig oft stetig differenzierbaren Funktionen, deren Träger $\text{supp} f := \overline{\{x | f(x) \neq 0\}}$, wo der Strich über der Menge den topologischen Abschluß bedeuten soll.

Um die schwache Ableitung von Θ zu gewinnen, erinnern wir uns, daß diese wie folgt definiert wird

$$\forall f \in C_0^\infty(\mathbb{R}) : \int dx \Theta'(x - y) f(x) = - \int dx \Theta(x - y) f'(x). \quad (1.17)$$

Daraus ergibt sich unmittelbar

$$- \int dx \Theta(x - y) f'(x) = - \int_y^\infty dx f'(x) = f(y) = \int dx \delta(x - y) f(x). \quad (1.18)$$

Für unser Beispiel (1.15) ist also die Wahrscheinlichkeitsverteilung

$$W_\xi(x) := P'_\xi(x) = \frac{1}{6} \sum_{i=1}^6 \delta(x - i). \quad (1.19)$$

Die Anwendung der Distributionentheorie hat den Vorteil, daß wir die Fälle *diskreter und kontinuierlicher Zufallsvariablen* in gleicher Weise notieren können. Für kontinuierliche Zufallsvariablen ist es oft zweckmäßiger, von den Verteilungen auszugehen und danach erst durch Integration die Wahrscheinlichkeitsfunktionen zu gewinnen.

Wir gelangen nun zum wichtigen Begriff des Erwartungswertes:

Definition 2. Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ besitzt bei gegebener Wahrscheinlichkeitsverteilung W_ξ den Erwartungswert

$$\langle f(\xi) \rangle = \int dx f(x) W_\xi(x), \quad (1.20)$$

vorausgesetzt das definierende Integral existiert.

Speziell gilt

$$P_\xi(x) = \langle \Theta(x - \xi) \rangle \quad \text{und} \quad W_\xi(x) = \langle \delta(x - \xi) \rangle. \quad (1.21)$$

Offenbar ist

$$\langle 1 \rangle = \int dx W_\xi(x) = \lim_{x \rightarrow \infty} P_\xi(x) - \lim_{x \rightarrow -\infty} P_\xi(x) = 1. \quad (1.22)$$

Sei nun $g : D_\xi \rightarrow \mathbb{R}$ und g schwach differenzierbar. Weiter besitze für $y \in \mathbb{R}$ die Abbildung $x \mapsto g(x) - y$ nur eine endliche Anzahl von einfachen Nullstellen. Dann können wir leicht die Verteilungsfunktion für die Zufallsvariable $\eta = g \circ \xi$ angeben: Gemäß (1.21) setzen wir

$$W_\eta(y) = \langle \delta[y - g(\xi)] \rangle = \int dx \delta[y - g(x)]. \quad (1.23)$$

Sei weiter $\{g_n^{-1}(y)\}_{n \in I(y)}$ die Menge aller Punkte x mit $g(x) = y$. Dann folgt daraus sofort

$$W_\eta(y) = \sum_{n \in I(y)} \frac{W_\xi[g_n^{-1}(y)]}{|g'[g_n^{-1}(y)]|}. \quad (1.24)$$

Wir geben weiter einige nützliche Definitionen, die sich später als rechentechnisch vorteilhaft erweisen werden:

Definition 3. Die charakteristische Funktion einer Wahrscheinlichkeitsverteilung W_ξ ist als deren Fouriertransformation definiert:

$$C_\xi(u) = \int dx \exp(iux) W_\xi(x) = \langle \exp(iu\xi) \rangle. \quad (1.25)$$

Dieser Definition entnehmen wir sofort die Erwartungswerte aller Potenzen von ξ :

$$M_n := \langle \xi^n \rangle = \frac{1}{i^n} \left(\frac{d}{du} \right)^n C_\xi(u) |_{u=0}, \quad (1.26)$$

M_n wird auch als das n -te Moment der Wahrscheinlichkeitsverteilung W_ξ bezeichnet. Aus dem Taylorschen Satz entnehmen wir, daß umgekehrt die Kenntnis aller Momente einer Wahrscheinlichkeitsverteilung deren vollständige Rekonstruktion gestattet:

$$C_\xi(u) = 1 + \sum_{n=1}^{\infty} (iu)^n \frac{M_n}{n!}. \quad (1.27)$$

Definition 4. Die kumulative Verteilungsfunktion ist durch den Logarithmus der charakteristischen Funktion gegeben:

$$K_\xi(u) = \ln[C_\xi(u)] = \ln \left[1 + \sum_{n=1}^{\infty} \frac{(iu)^n}{n!} M_n \right] = \sum_{n=1}^{\infty} \frac{(iu)^n}{n!} K_n. \quad (1.28)$$

Die Koeffizienten K_n werden Kumulanten genannt.

Es ist klar, daß wir auch den Logarithmus (zumindest formal) in seine Potenzreihe entwickeln können und also aus den Kumulanten auf die Momente und vice versa schließen können.

Wir geben uns für jetzt mit den ersten beiden Kumulanten zufrieden:

$$K_1 = M_1, K_2 = M_2 - M_1^2. \quad (1.29)$$

Der erste Kumulant ist also einfach das erste Moment, d.h. der Erwartungswert der Zufallsvariablen ξ . Der zweite Kumulant wird auch als *Varianz* bezeichnet und ein Maß für die Breite der Verteilung um den Mittelwert der Zufallsvariablen. Es gilt

$$K_2 = \langle (\xi - \langle \xi \rangle)^2 \rangle = \langle \xi^2 \rangle - \langle \xi \rangle^2. \quad (1.30)$$

Betrachten wir nun zwei einfache Beispiele von Verteilungen, durch ihre Kumulanten definiert werden können.

Sei zuerst nur der Mittelwert $K_1 = M_1$ der Zufallsvariablen von 0 verschieden, d.h. sei $0 = K_2 = K_3 = \dots$. Dann folgt zunächst für die charakteristische Funktion und dann durch Fourierrücktransformation

$$C_\xi(u) = \exp(iuK_1) \Rightarrow W_\xi(x) = \frac{1}{2\pi} \int du C_\xi(u) \exp(-iux) = \delta(x - K_1). \quad (1.31)$$

Das bedeutet, daß bei Verschwinden aller Kumulanten ab dem zweiten das Zufallsexperiment hinsichtlich der Variablen ξ insofern determiniert ist als bei jedem Zufallsexperiment (bis auf eine Menge vom Maß 0) stets ein Wert für ξ resultiert, der $> K_1$ ist.

Als nächstes betrachten wir den Fall, daß alle Kumulanten ab dem einschließlich dritten identisch verschwinden. Dann ist die Verteilung eine *Gaußverteilung*

$$C_\xi(u) = \exp\left(iuK_1 - K_2 \frac{u^2}{2}\right) \Rightarrow W_\xi(x) = \frac{1}{\sqrt{2\pi K_2}} \exp\left[-\frac{(x - K_1)^2}{2K_2}\right]. \quad (1.32)$$

Dies zeigt direkt, daß K_2 ein Maß für die Breite der Verteilung ist, während die höheren Kumulanten die Abweichung von einer Gaußverteilung angeben.

1.3 Vektorwertige Zufallsvariable

In der statistischen Physik kommen meist Probleme vor, die durch mehr als eine Wahrscheinlichkeitsvariable beschrieben werden. In diesem Abschnitt werden wir kurz die entsprechenden mathematischen Hilfsmittel bereitstellen, die im wesentlichen Verallgemeinerungen des Falles einer einzelnen Zufallsvariablen sind.

Wir definieren entsprechend $\xi : X \subseteq \mathcal{P}\Omega \rightarrow \mathbb{R}^r$ und eine Wahrscheinlichkeitsverteilung

$$W_\xi : \mathbb{R}^r \rightarrow \mathbb{R} : W_\xi(x) = \langle \delta^{(r)}(\xi - x) \rangle. \quad (1.33)$$

Der Erwartungswert einer Funktion f , die auf dem Wertebereich von ξ definiert (und sogleich durch 0 auf ganz \mathbb{R}^r fortgesetzt gedacht wird) ist wieder durch

$$\langle f(\xi) \rangle = \int_{\mathbb{R}^r} d^r x f(x) W_\xi(x) \quad (1.34)$$

festgelegt.

Die Verteilung für eine niederdimensionale Teilmenge von D_ξ kann dann wie folgt bestimmt werden. Wir nehmen also an, die Teilmenge liege ganz in einem Unterraum, die von den ersten $i < r$ Basisvektoren aufgespannt werden. Dann ist die Verteilung für die Zufallsvariablen ξ_1, \dots, ξ_i durch

$$W_i(x_1, \dots, x_i) = \int_{\mathbb{R}^{r-i}} dx_{i+1} \cdots dx_r W_r(x_1, \dots, x_r) \quad (1.35)$$

definiert. Dies ist konsistent mit der allgemeinen Definition (1.33), denn es ist ja

$$\delta^{(r)}(x - \xi) = \prod_{k=1}^r \delta(x_k - \xi_k) \quad (1.36)$$

folglich

$$\begin{aligned} W_i(x_1, \dots, x_i) &= \left\langle \delta^{(i)}(x_1 - \xi_1, \dots, x_i - \xi_i) \right\rangle = \\ &= \int_{\mathbb{R}^{r-i}} dx_{i+1} \cdots dx_r \left\langle \delta^{(r)}(x_1 - \xi_1, \dots, x_r - \xi_r) \right\rangle \end{aligned} \quad (1.37)$$

Das heißt mit (1.33) ergibt dies (1.35).

Wie im Falle einer einzelnen Zufallsvariablen definiert die Fouriertransformation der Verteilung die *charakteristische Funktion*

$$C_r(u) = \langle \exp(iu\xi) \rangle = \int_{\mathbb{R}^r} \exp(iux) W_r(x). \quad (1.38)$$

Die Momente sind ebenfalls analog wie bei einer einzelnen Zufallsvariablen gegeben:

$$M_{n_1 \dots n_r} = \left\langle \prod_{j=1}^r \xi_j^{n_j} \right\rangle = \left(\frac{1}{i} \right)^{\sum_{j=1}^r n_j} \prod_{j=1}^r \left(\frac{\partial}{\partial u_j} \right)^{n_j} C_r(u) \Big|_{u=0}. \quad (1.39)$$

Wieder läßt sich die charakteristische Funktion aus den Momenten über die Taylorentwicklung zurückgewinnen:

$$C_r(u) = \sum_{n_1 \dots n_r=1}^{\infty} M_{n_1 \dots n_r} \prod_{j=1}^r \frac{(iu_j)^{n_j}}{n_j!}. \quad (1.40)$$

Kehren wir (1.38) um, erhalten wir schließlich aus der charakteristischen Funktion die Verteilung zurück:

$$W_r(x) = \left(\frac{1}{2\pi} \right)^r \int_{\mathbb{R}^r} d^r u \exp(-iux) C_r(u). \quad (1.41)$$

Die Kumulanten sind

$$K_{n_1 \dots n_r} = \left(\frac{1}{i} \right)^{\sum_{j=1}^r n_j} \sum_{j=1}^r \prod_{j=1}^r \left(\frac{\partial}{\partial u_j} \right)^{n_j} \ln[C_r(u)] \Big|_{u=0}. \quad (1.42)$$

Die charakteristische funktion läßt wiederum aus den Kumulanten rekonstruieren

$$C_r(u) = \exp \left[\sum_{n_1 \dots n_r \in \mathbb{N}} K_{n_1 \dots n_r} \prod_{j=1}^r \frac{(iu_j)^{n_j}}{n_j!} \right]. \quad (1.43)$$

Wir werden in der statistischen Physik übrigens eine enge Verwandtschaft der Momenten mit den Greenschen Funktionen der Quantenfeldtheorie und eine ebensolche der Kumulanten mit den eigentlichen Vertexfunktionen der Quantenfeldtheorie erkennen, so daß die Rekonstruktion der Verteilung aus den Kumulanten eine bedeutende Rolle spielt.

Jetzt können wir unter Verwendung von (1.35) die bedingte Wahrscheinlichkeitsverteilung unter der Voraussetzung, daß die erste Komponente $\xi_1 = x_1$ ist angeben:

$$P_{x_1|x_2, \dots, x_r} = \frac{W_r(x_1, \dots, x_r)}{W_{r-1}(x_2, \dots, x_r)} = \frac{W_r(x_1, \dots, x_r)}{\int_{\mathbb{R}} dx_1 W_r(x_1, \dots, x_r)}. \quad (1.44)$$

Per Definition sind die Wahrscheinlichkeitsvariablen ξ_1 und ξ_2 stochastisch unabhängig, wenn die Wahrscheinlichkeitsverteilung wie folgt faktorisiert:

$$W_2(x_1, x_2) = P(x_1|x_2) \int_{\mathbb{R}} dx_1 W_2(x_1, x_2) =: W_1^{(1)}(x_1) W_2^{(2)}(x_2). \quad (1.45)$$

Umgekehrt definieren wir zwei Zufallsvariable ξ_1 und ξ_2 als stochastisch unabhängig, wenn die Verteilungsfunktion in der oben angegebenen Weise in ein Produkt aus zwei Verteilungen jeweils einer Zufallsvariablen faktorisiert.

Der andere Extremfall ist der eines vollständigen Determinismus von ξ_1 , wenn ξ_2 bekannt ist, d.h. es existiert dann ein funktionaler Zusammenhang $\xi_1 = f(\xi_2)$ der Zufallsvariablen, und die bedingte Wahrscheinlichkeit ist dann notwendig durch

$$P(x_1|x_2) = \delta[x_1 - f(\xi_2)] \quad (1.46)$$

beschreiben. Aus (1.44) ergibt sich dann

$$W_2(x_1, x_2) = \delta[x_1 - f(x_2)] W_1(x_2). \quad (1.47)$$

Es gibt verschiedene Maße für die stochastische Unabhängigkeit oder für die *Korrelation* zweier Zufallsvariablen. Die bekannteste die die sog. *Kreuzkorrelationsfunktion*

$$\kappa(\xi_1, \xi_2) = \langle \xi_1 \xi_2 \rangle - \langle \xi_1 \rangle \langle \xi_2 \rangle. \quad (1.48)$$

Mit (1.45) ist sofort klar, daß für stochastisch unabhängige Zufallsvariable, die Kreuzkorrelation verschwindet. Es muß ausdrücklich betont werden, daß die Umkehrung nicht gilt: Aus dem Verschwinden der Kreuzkorrelation kann *nicht* auf die stochastische Unabhängigkeit der Zufallsvariablen geschlossen .

Ein anderes Korrelationsmaß ist die *relative Kreuzkorrelation*:

$$R(\xi_1, \xi_2) = \frac{\kappa(\xi_1, \xi_2)}{\sigma(\xi_1)\sigma(\xi_2)} \text{ mit } \sigma(\xi_j) = \sqrt{\langle \xi_j^2 \rangle - \langle \xi_j \rangle^2}, \quad j = 1, 2. \quad (1.49)$$

Um diese Betrachtungen auf mehr als zwei Zufallsvariable zu verallgemeinern, sei bemerkt, daß (1.48) der K_{11} -Kumulant gemäß (1.43) ist. Die Kreuzkorrelation für r Zufallsvariable wird daher durch

$$\kappa(\xi_1, \dots, \xi_r) = K_{1\dots r} = \prod_{j=1}^r \left(\frac{1}{i} \frac{\partial}{\partial u_j} \right) \ln C_r(u_1, \dots, u_r) \Big|_{u=0} \quad (1.50)$$

definiert. Diese Funktion verschwindet übrigens bereits, wenn nur eine einzige Zufallsvariable, z.B. ξ_1 , stochastisch unabhängig von den übrigen ist, d.h. wenn

$$W_r(x) = W_1^{(1)}(x_1) W_{r-1}^{(2)}(x_2, \dots, x_r), \quad (1.51)$$

wobei $W_{r-1}^{(2)}$ nicht faktorisieren muß, gilt.

1.4 Stochastische Prozesse

Wir nähern uns nun schon eher physikalischen Fragestellungen, nämlich wie man die zeitliche Entwicklung von Zufallsexperimenten beschreiben kann. Dabei setzen wir voraus, daß die Zufallsexperimente an Systemen ausgeführt werden, die sich zeitlich verändern können. Dies nennt man einen *stochastischen Prozeß*. Wir betrachten zunächst ein Zufallsexperiment, bei dem eine einzelne Zufallsvariable ξ gemessen wird. Die Verallgemeinerung auf mehr als eine Zufallsvariable ist ohne weitere Komplikationen möglich.

Die Zeitabhängigkeit des Systems ist durch die Zeitabhängigkeit der Zufallsvariablen ξ gegeben und unter Verwendung von (1.33) erhalten wir für die Wahrscheinlichkeitsverteilung

$$W_1(x_1; t) = \langle \delta[x_1 - \xi(t)] \rangle, \quad (1.52)$$

die ihrerseits nun ebenfalls von der Zeit abhängt. Die *empirische* Erwartungswertbildung ist hier so zu verstehen, daß ein hinreichend großes Ensemble von voneinander unabhängigen gleichartig präparierten Systemen vorliegt, an denen die Zufallsvariable ξ zur Zeit t gemessen wird.

Jetzt können wir aber auch danach fragen, wie die Wahrscheinlichkeitsverteilung dafür aussieht, daß ξ zu den Zeiten t_1, \dots, t_n beziehentlich die Werte x_1, \dots, x_n annimmt. Wir wenden wieder (1.33) an, um diese Verteilung zu bestimmen:

$$W_n(x_n, t_n; \dots, x_1, t_1) = \left\langle \prod_{j=1}^n \delta[x_j - \xi(t_j)] \right\rangle. \quad (1.53)$$

Ist W_n von allen Zeiten t_j unabhängig, dann heißt der betreffende Prozeß *stationär*.

Die bedingte Wahrscheinlichkeit, daß $\xi(t_n)$ den Wert x_n unter der Voraussetzung, daß es die Werte x_1, \dots, x_{n-1} zu den früheren Zeiten $t_1 < t_2 < \dots < t_n$ angenommen hat, ist gemäß (1.44) durch

$$P(x_n, t_n | x_{n-1}, t_{n-1}; \dots; x_1, t_1) = \frac{W_n(x_n, t_n; \dots; x_1, t_1)}{\int_{\mathbb{R}} dx_n W_n(x_n, t_n; \dots; x_1, t_1)} \quad (1.54)$$

gegeben.

Üblicherweise werden die stochastischen Prozesse grob in folgende Typen unterteilt.

- Vollständig stochastische Prozesse

Ein vollständig stochastischer Prozeß ist dadurch charakterisiert, daß die Wahrscheinlichkeitsverteilung (1.53) wie folgt faktorisiert

$$W_n(x_n, t_n; \dots; x_1, t_1) = \prod_{j=1}^n W_1(x_j, t_j). \quad (1.55)$$

Das bedeutet, daß die statistischen Eigenschaften der Zufallsvariable ξ zur Zeit t_n unkorreliert mit den Eigenschaften der Zufallsvariablen zu jeder früheren Zeit sind. Dies erwarten wir von physikalischen Vorgängen nicht, denn aufgrund der unterliegenden *kausalen* Naturgesetze erwarten wir, daß die statistischen Eigenschaften von Mittelwerten der Größen von Systemen, deren Zustand nur unvollständig bekannt ist, ebenfalls eine bestimmte Kausalstruktur aufweisen, d.h. daß zumindest bei hinreichender statistischer Kenntnis über die Vorgeschichte des Systems die statistischen Eigenschaften des Systems zu einem späteren Zeitpunkt berechenbar sind.

- Markovprozesse

Bei einem Markovprozeß sind die bedingten Wahrscheinlichkeiten (1.54) unabhängig von den Bedingungen $\xi(t_1) = x_1, \dots, \xi(t_{n-2}) = x_{n-2}$. Entscheidend ist nur die Bedingung zum letzten Zeitpunkt vor der Zeit t_n , d.h.

$$P(x_n, t_n | x_{n-1}, t_{n-1}; \dots; x_1, t_1) = P(x_n, t_n | x_{n-1}, t_{n-1}). \quad (1.56)$$

Mit (1.54) und (1.56) kann man die Wahrscheinlichkeiten durch die eine bedingte Wahrscheinlichkeit auf der rechten Seite von (1.56) und die Wahrscheinlichkeitsverteilung W_1 ausdrücken: ausdrücken

$$W_n(x_n, t_n; \dots; x_1, t_1) = \left[\prod_{k=1}^{n-1} P(x_{k+1}, t_{k+1} | x_k, t_k) \right] W_1(x_1, t_1), \quad (1.57)$$

wodurch deutlich wird, daß $P(x_{k+1}, t_{k+1} | x_k, t_k)$ die *Übergangswahrscheinlichkeit* dafür ist, daß die Zufallsvariable ξ zur Zeit t_{k+1} den Wert x_{k+1} annimmt, vorausgesetzt, sie hatte den Wert x_k zur früheren Zeit t_k .

- Allgemeine Prozesse

Im allgemeinsten Falle hängt (1.54) von allen Zeiten t_k ab, d.h. das System „erinnert“ sich an alle Messungen der Zufallsvariablen die jemals früher an ihm stattgefunden haben.

In der Physik treten solche Prozesse tatsächlich auf. Bestimmte Makromoleküle „erinnern“ sich an ihre gesamte Vorgeschichte. Für solche Prozesse ist es unmöglich, deren statistische Verteilungen durch eine einzelne Anfangsverteilung festzulegen. Vielmehr benötigt man ein „Gedächtnisfunktional“, das die gesamte Vorgeschichte des Systems berücksichtigt.

1.5 Information

Jetzt kommen wir auf die Frage zurück, wie wir Wahrscheinlichkeitsverteilungen für ein bestimmtes Zufallsexperiment aufgrund irgendwelcher, i.a. unvollständiger, Kenntnisse über das System zuordnen können. Dabei ist es das Problem, eine solche Verteilung zu finden, die nicht mehr Kenntnisse vortäuschen als tatsächlich vorliegen. Man sollte also bestrebt sein, Wahrscheinlichkeitsverteilungen mit *geringstem Vorurteil* zu finden, die mit der gegebenen Information über die Zufallsgröße verträglich ist.

Dazu müssen wir ein *Maß für die fehlende Information* definieren, welches einer beliebigen Wahrscheinlichkeitsverteilung zuzuordnen ist. Dabei ist klar, daß dieses Maß in einem bestimmten Sinne den Vergleich des unvollständigen Informationsgehalts gegenüber der vollständigen Kenntnis des Systemzustandes wiedergeben soll.

Wir beginnen mit dem einfachsten Fall eines Zufallsexperiments mit einer endlichen Zahl möglicher Ausgänge. Die Ereignismenge sei $\Omega = \{1, \dots, n\}$ und die Wahrscheinlichkeitsverteilung sei durch die Laplacesche Gleichverteilung gegeben: $P(\{i\}) = 1/n$ für $i \in \Omega$.

Jetzt vergleichen wir zwei verschiedene Zufallsexperimente, wo gilt $|\Omega_\mu| = N_\mu$ mit $\mu \in \{1, 2\}$. Die Wahrscheinlichkeitsverteilungen seien beidemal die Gleichverteilung. Dann verlangen wir vom Maß I für die fehlende Information unter diesen Voraussetzungen, daß

$$I(N_1) > I(N_2), \text{ falls } N_1 > N_2. \quad (1.58)$$

Ist $N = 1$, also kann überhaupt bei einem Experiment nur ein Ausgang möglich sein, fehlt überhaupt keine Information, wir wissen ja, was bei jedem Experiment herauskommt. Unser Informationsmaß I soll dann also verschwinden:

$$I(1) = 0. \quad (1.59)$$

Jetzt formulieren wir diesen Spezialfall anders, indem wir uns vorstellen, das Experiment bestehe darin, einen Gegenstand zufällig in einen Kasten mit N Fächern zu werfen. Der Ausgang des Experiments ist die Nummer der Zelle, in den der Gegenstand gefallen ist.

Jetzt stellen wir uns vor, jedes Fach sei noch in M Zellen unterteilt. Wir haben also insgesamt NM Zellen. Wieder Gleichverteilung für alle Zellen vorausgesetzt, ist also die fehlende Information $I(NM)$. Andererseits können wir aber das Objekt in seiner Zelle dadurch lokalisieren, daß wir zunächst fragen, in welches Fach es wohl gefallen sein mag. Die fehlende Information dafür ist $I(N)$. Weiter können wir fragen, in welche der M Zellen des entsprechenden Faches der Gegenstand gefallen ist. Die fehlende Information ist $I(M)$. Nehmen wir nun an, es gäbe keinerlei Bevorzugung einer Zelle in Abhängigkeit vom gerade getroffenen Fach, ist es sinnvoll anzunehmen, daß die gesamte fehlende Information

$$I(NM) = I(N) + I(M) \quad (1.60)$$

ist.

Ist nun $N/L \in \mathbb{N}$ kann man diese Gleichung auch in der Gestalt

$$I(N/L) = I(N) - I(L) \quad (1.61)$$

schreiben.

Um nun ein Maß für alle positive rationale Zahlen zu definieren, nehmen wir an, diese Gleichung gelte überhaupt für alle Paare $N, L \in \mathbb{N}_{>0}$. Wir nehmen ferner an, man könne diese Funktion stetig nach allen positiven reellen Zahlen fortsetzen.

Unter diesen Annahmen liegt das Informationsmaß für Gleichverteilungen bis auf einen positiven Faktor eindeutig fest. Sei dazu $\ln x = M/N$ mit $M, N \in \mathbb{N}_{>0}$. Dann gilt wegen (1.60) und (1.61) und der Stetigkeitsannahme

$$I(x^N) = I(\exp M) \Rightarrow NI(x) = mI(e). \quad (1.62)$$

Damit folgt aber

$$I(x) = \frac{M}{N}I(e) = I(e) \ln x. \quad (1.63)$$

Definieren wir schließlich der Bequemlichkeit noch $I(e) = 1$ liegt das Maß für die fehlende Information für N gleichverteilte Ereignisse fest:

$$I(N) = \ln N. \quad (1.64)$$

Jetzt wollen wir diese Definition für das Maß der fehlenden Information für den Fall verallgemeinern, daß wir ein Zufallsexperiment mit N möglichen Ausgängen haben und die Wahrscheinlichkeiten für jedes Elementarereignis P_i willkürlich vorgegeben sind. Es muß natürlich nach den Kolmogorowschen Axiomen $P_i \geq 0$ und $\sum_{i=1}^n P_i = 1$ sein.

Wir zeigen nun, daß wir das Informationsmaß für diesen Fall dadurch auf den der gleichverteilten Elementarereignisse zurückführen können, daß wir uns das Experiment N -mal wiederholt

denken, wobei die einzelnen Experimente voneinander unabhängig sind. Die Wahrscheinlichkeit, daß $\{k\}$ eintritt ist per definitionem stets gleich wahrscheinlich $p = P_k$. Die Wahrscheinlichkeit, daß $\{k\}$ eintritt ist natürlich $q = 1 - p$. Wir zeigen jetzt, daß wir die Wahrscheinlichkeit p beliebig genau bestimmen können, wenn wir $N \rightarrow \infty$ machen, d.h. das Zufallsexperiment beliebig oft wiederholen. Sei dafür $N(k, N)$ die Zahl der Fälle, daß k eingetreten ist, wenn N voneinander unabhängige Zufallsexperimente vorgenommen wurden. Wir können den Ausgang der N Experimente als N -Tupel hinschreiben. Die Wahrscheinlichkeit, daß ein bestimmtes N -Tupel mit $N(k, N)$ Einträgen k erscheint ist wegen der Unabhängigkeit der Experimente offenbar durch $p^{N(k, N)} q^{N - N(k, N)}$ gegeben.

Allerdings ist es für die Wahrscheinlichkeit, daß bei N unabhängigen Zufallsexperimenten genau $N(k, N)$ -mal das Ereignis $\{k\}$ eingetroffen wird, die Reihenfolge der Experimentausgänge belanglos, so daß wir schließlich für die Wahrscheinlichkeit, daß bei N unabhängigen Zufallsexperimenten $N(k, N)$ -mal das Ereignis $\{k\}$ eingetreten ist, die Wahrscheinlichkeit zu

$$P[N(k, N)] = \binom{N}{N(k, N)} p^{N(k, N)} q^{N - N(k, N)} \quad (1.65)$$

ermitteln. Jetzt wollen wir die Momente dieser Verteilung berechnen. Dazu bedienen wir uns eines Tricks, der in der Wahrscheinlichkeitstheorie oft zum Erfolg führt. Wir definieren die Erzeugende Funktion

$$f(x) = \sum_{j=0}^N \binom{N}{j} p^j q^{N-j} x^j = (px + q)^N, \quad (1.66)$$

wobei wir das Binomialtheorem benutzt haben.

Für den Erwartungswert finden wir z.B.

$$\langle N(k, N) \rangle = \sum_{j=0}^N j \binom{N}{j} p^j q^{N-j} = f'(1) = Np \quad (1.67)$$

und für das zweite Moment

$$\langle N(k, N)^2 \rangle = [x f'(x)]' |_{x=1} = Np + N(N-1)p^2. \quad (1.68)$$

Der Erwartungswert für $N(k, N)/N$ und seine Varianz ist demnach

$$\left\langle \frac{N(k, N)}{N} \right\rangle = p, \quad \sigma^2_{\left(\frac{N(k, N)}{N}\right)} = \frac{pq}{N}. \quad (1.69)$$

Es ist leicht zu zeigen, daß auch alle höheren Kumulanten für $N \rightarrow \infty$ verschwinden, also zur determinierten Verteilung für $N(k, N)$ führen.

Die Anzahl überhaupt möglicher Resultate ist die Anzahl verschiedener N -Tupel, die im Limes großer Versuchsanzahl $N \rightarrow \infty$ überhaupt möglich sind, wobei das Vertauschen gleicher Ergebnisse jedoch gleichgültig ist. Für $N \rightarrow \infty$ ist aber nach dem obigen Resultat $N(j, N) = P_j N$ und damit die totale Zahl von Ereignissen für $N \rightarrow \infty$ durch

$$\frac{N!}{\prod_{j=1}^n (NP_j)!} \quad (1.70)$$

gegeben. Unter Verwendung von (1.64) für gleichverteilte Zufallsexperimente ergibt sich also im Limes großer N

$$I(N) = \ln \frac{N!}{\prod_{j=1}^n (NP_j)!} = \ln N! - \sum_{j=1}^n \ln[(NP_j)!]. \quad (1.71)$$

Für $N \rightarrow \infty$ können wir aber die *Stirlingsche* Formel benutzen:

$$\ln N! \underset{N \rightarrow \infty}{\cong} N \ln N - N + O(\ln N). \quad (1.72)$$

Es folgt also

$$I(N) \underset{N \rightarrow \infty}{\cong} -N \sum_{j=1}^n P_j \ln P_j + O(\ln N). \quad (1.73)$$

Die fehlende Information pro Experiment ist also

$$I[P] = \lim_{N \rightarrow \infty} \frac{I(N)}{N} = - \sum_{j=1}^n P_j \ln P_j = - \langle \ln P \rangle. \quad (1.74)$$

Man kann dieselbe Formel auch für ein Zufallsexperiment mit abzählbar unendlich vielen Ausgängen benutzen:

$$I[P] = - \sum_{j=1}^{\infty} P_j \ln P_j. \quad (1.75)$$

Im Falle kontinuierlicher Verteilungen, müssen wir noch einen Limes aus diskreten Verteilungen bilden. Sei ξ ein d -dimensionaler Zufallsvektor. Dann kann man den Definitionsbereich in Würfel vom Volumen $\Delta^d x$ einteilen, und man kann eine diskrete Verteilung

$$P_j = W_d(x_j) \Delta^d x \quad (1.76)$$

definieren, wo x_j ein beliebiger Wert aus dem j Würfel ist. Gemäß (1.74) bzw. (1.75) ist die fehlende Information

$$I(W, \Delta^d x) = - \sum_j P_j \ln P_j = - \sum_j W_d(x_j) \Delta^d x \ln[W_d(x_j) \Delta^d x]. \quad (1.77)$$

Wir werden kontinuierliche Verteilungen stets als Limes entsprechender diskreter Verteilungen behandeln, weshalb sich hier eine allgemeine Definition der fehlenden Information für kontinuierliche Verteilungen erübrigt.

Für eine rigorosere Behandlung des Informationsbegriffes sei auf [Hob87] verwiesen, wo der Begriff der relativen Information eingeführt wird, die den Informationsgehalt einer Verteilung p_i relativ zu einer Referenzverteilung $p_i^{(0)}$ definiert. Dann können kontinuierliche Verteilungen problemlos durch

$$I(P; P^{(0)}) = \int dx W_\xi(x) \ln \left[\frac{W_\xi(x)}{W_\xi^{(0)}(x)} \right] \quad (1.78)$$

definiert werden. Diese Definition hat den Vorteil, daß im Logarithmus eine dimensionslose Größe auftritt, wie es sein muß, und (1.78) invariant unter beliebigen diffeomorphen Transformationen von einer Zufallsvariablen ξ zu einer anderen Zufallsvariable η ist, da W_ξ per constructionem eine Dichte im Sinne der Differentialgeometrie ist.

Der obige Informationsbegriff fällt mit dem jetzigen zusammen, wenn man als Referenzverteilung die Gleichverteilung verwendet.

1.6 Erinnerung an die Quantentheorie

Wir setzen voraus, daß der Leser mit der Quantentheorie, formuliert im Diracschen „Bra-Ket-Formalismus“ vertraut ist. Meine ausführliche Zusammenfassung dieser Grundlagen findet sich auf meiner Webpage [Hee98]. Von der zahlreichen Lehrbuchliteratur seien nur die folgenden beiden zitiert: [Fic79, Bal98]. Hier wollen wir lediglich die Notation etablieren und den Rahmen zur Behandlung der Quantenstatistik bereitstellen.

Wir stellen in diesem Abschnitt die Grundlagen der Quantentheorie zusammen, und zwar zunächst ohne auf die *Quantenstatistik* einzugehen. Die Quantentheorie in diesem Sinne behandelt nur die Frage, wie ein System überhaupt quantentheoretisch beschrieben wird (analog zur *Kinematik* für ein klassisches mechanisches System) und wie bei vollständiger Kenntnis des Anfangszustandes die zeitliche Entwicklung der Größen berechnet werden kann (analog zur Dynamik in der klassischen Mechanik, die z.B. durch die Newtonschen Bewegungsgleichungen für Punktteilchen bestimmt ist).

- (Q1) Der Zustand eines Quantensystems wird durch den Strahl in einem Hilbertraum \mathcal{H} vollständig beschrieben. Ein Strahl ist dabei für $0 \neq |\psi\rangle \in \mathcal{H}$ durch die folgende Äquivalenzklasse von Vektoren definiert:

$$[|\psi\rangle] = \{c|\psi\rangle \mid |\psi\rangle \in \mathcal{H}, c \in \mathbb{C} \setminus \{0\}\}. \quad (1.79)$$

- (Q2) Jeder Observablen O des Systems wird ein *selbstadjungierter Operator* $\mathbf{O} : \mathcal{D} \rightarrow \mathcal{D}$, welcher zumindest auf einem dichten Teilraume $\mathcal{D} \subseteq \mathcal{H}$ definiert ist, zugeordnet.

Das Resultat einer exakten Messung der Observablen O ist notwendig ein Eigenwert des entsprechenden Operators \mathbf{O} .

Es sei weiter $\text{Eig}(\mathbf{O}, o)$ der Eigenraum zum Eigenwert o und

$$\mathbf{P}(o) = \sum_k |\mathbf{O}, o, k\rangle \langle \mathbf{O}, o, k| \quad (1.80)$$

der *Projektionsoperator* auf diesen Eigenraum. Dabei sind die $|\mathbf{O}, o, k\rangle$ eine beliebige orthonormierte Basis des Eigenraums. Dann ist

$$p(o) = \langle \psi | \mathbf{P}(o) | \psi \rangle = \sum_k |\langle \psi | \mathbf{O}, o, k \rangle|^2 \text{ mit } \langle \psi | \psi \rangle = 1 \quad (1.81)$$

die Wahrscheinlichkeit dafür, bei einer Messung der Observablen O den Meßwert o zu erhalten.

Zwei Observablen O_1 und O_2 heißen kompatibel, wenn die sie repräsentierenden selbstadjungierten Operatoren kommutieren.

Ein Menge $\{\mathbf{O}_j\}_{j \in \{1, \dots, n\}}$ von kompatiblen Observablen heißt *vollständiger Satz kompatibler Observabler*, wenn deren simultaner Eigenraum $\text{Eig}[(\mathbf{O}_j), (o_j)]$ eindimensional ist.

Bemerkungen:

Anstatt mit den praktisch etwas unbequemen Strahlen zu operieren, können wir auch den Projektionsoperator

$$\mathbf{P}_{|\psi\rangle} = \frac{|\psi\rangle \langle \psi|}{\|\psi\|^2} \quad (1.82)$$

zur Zustandsfestlegung verwenden. Er ist unabhängig von der Wahl des Repräsentanten im Strahl $[|\psi\rangle]$ und definiert umgekehrt eindeutig den Strahl, denn $|\psi\rangle$ ist offensichtlich sein einziger Eigenvektor zum Eigenwert 1. Solche Zustände eines Systems bezeichnen wir genauer auch als *reine Zustände*.

Hier muß betont werden, daß die statistischen Aussagen, die durch dieses Postulat getroffen werden, *nicht* notwendig sind, weil wir den genauen Zustand des Systems nicht kennen. Wir setzen ja gerade voraus, daß sich das System in dem Zustand $[|\psi_1\rangle]$ befindet und daß die Zuordnung dieses Strahls im Hilbertraum die vollständige Festlegung des Systemzustandes bedeutet. Vielmehr spiegeln diese Wahrscheinlichkeitsaussagen den prinzipiellen *Indeterminismus*, der durch die Quantentheorie gerade beschrieben werden soll, wieder. Es kann zwar sein, daß, obwohl man den Zustand $[|\psi\rangle]$ des Systems mit Sicherheit kennt, die Frage, welchen Wert die Messung einer Observablen O liefert, nicht mit Sicherheit beantwortet werden kann. Der Meßwert ist gemäß Postulat (Q2) nur dann determiniert, wenn $|\psi\rangle$ ein Eigenvektor des dieser Observablen zugeordneten selbstadjungierten Operators \mathbf{O} ist, und der Meßwert ist dann mit Sicherheit der dazugehörige Eigenwert.

Die Quantentheorie macht dann jedoch immerhin die statistische Aussage, daß (1.81) die Wahrscheinlichkeit dafür angibt, bei der Messung der Observablen O gerade den Eigenwert o des dazugehörigen Operators \mathbf{O} zu finden. (Q2) sagt darüber hinaus aus, daß eine Observable gar keine anderen Werte als diese Eigenwerte annehmen kann.

Wir müssen nun noch kurz rekapitulieren, daß (Q2) eine mit den Kolmogorovschen Axiomen verträgliche Definition von Wahrscheinlichkeiten für das „Zufallsexperiment“ der Messung der Observablen O liefert. Zunächst ist klar, daß in dem gegebenen Falle die Ergebnismenge durch $\Omega = \{o | o \text{ ist Eigenwert von } \mathbf{O}\}$ gegeben ist. Die Ereignisalgebra ist die Potenzmenge von Ω , d.h. $E = \mathcal{P}\Omega$. Weiter müssen sich die Ergebnisse gegenseitig ausschließen. Dies ist jedoch der Fall, denn vorausgesetzt, das System sei in einem Zustand präpariert, in dem mit Sicherheit der bei der Messung von O der Meßwert o ist, befindet sich wegen (Q2) das System in einem Eigenzustand $|o\rangle$ des Operators \mathbf{O} zum Eigenwert o . Nun stehen alle Eigenvektoren zu einem von o verschiedenen Eigenwert o' aufeinander senkrecht, denn es gilt:

$$\langle o' | \mathbf{O} | o \rangle = o \langle o' | o \rangle = \langle \mathbf{O} o' | o \rangle = o' \langle o' | o \rangle \Rightarrow (o - o') \langle o' | o \rangle = 0. \quad (1.83)$$

Daraus folgt, daß für $o \neq o'$ die Eigenvektoren wie behauptet aufeinander senkrecht stehen. Daraus folgt unmittelbar, daß $\mathbf{P}[\text{Eig}(\mathbf{O}, o) | o'] = 0$ und folglich gemäß (1.81) die Wahrscheinlichkeit, daß o' gemessen wird, verschwindet.

Weiter bilden die Eigenvektoren eines selbstadjungierten Operators ein vollständiges Orthonormalsystem im Hilbertraum, so daß gemäß (Q2) auch

$$p(\Omega) = \sum_o p(o) = \langle \psi | \psi \rangle = 1 \quad (1.84)$$

gilt. Somit sind die Kolmogorovschen Axiome erfüllt.

Es bleibt zu bemerken, daß der Hilbertraum für reale Systeme (z.B. Teilchen) i.a. unendlichdimensional ist und selbstadjungierte Operatoren auch ein kontinuierliches Spektrum oder kontinuierliche Anteile im Spektrum haben können. Dann sind die obigen Axiome im Sinne der Spektraltheorie um *verallgemeinerte Eigenvektoren*, welche im Rahmen der Distributionentheorie strikt begründet werden können, zu verallgemeinern. Wir setzen voraus, daß für Werte im kontinuierlichen Spektrum verallgemeinerte Eigenvektoren existieren, die „auf die δ -distribution“ normiert werden können:

$$\langle o | o' \rangle = \delta(o - o'). \quad (1.85)$$

Die *Vollständigkeitsrelation* kann dann in der Gestalt

$$\sum_{o,k} |\mathbf{O}, o, k\rangle \langle \mathbf{O}, o, k| + \int_E do |o\rangle \langle o| = 1 \quad (1.86)$$

wobei die o in der Summe die Werte im diskreten Spektrum von \mathbf{O} und die k die zueinander orthogonal gewählten normierten Basisvektoren des zugehörigen Eigenraums durchlaufen. E ist die Teilmenge der reellen Zahlen, die das kontinuierliche Spektrum durchlaufen.

Ist nun o im kontinuierlichen Spektrum, so ist

$$p(o) = |\langle o | \psi \rangle|^2 \quad (1.87)$$

für einen auf 1 normierten Vektor $|\psi\rangle$ die *Wahrscheinlichkeitsverteilung* (im Sinne der in Abschnitt 1.2 erklärten Theorie der Zufallsgrößen) für das Meßresultat o bei der Messung der Observablen O . Man macht sich schnell klar, daß zusammen mit (1.85) der in Abschnitt 1.2 entwickelte Formalismus der Wahrscheinlichkeitsverteilungen und Wahrscheinlichkeitsfunktionen allgemein anwendbar gültig ist.

Eine der wohl berühmtesten Folgerungen aus dem Formalismus der Quantentheorie sind wohl Heisenbergs *Unschärferelationen*:

$$\Delta A \Delta B \geq \frac{1}{2} |\langle [\mathbf{A}, \mathbf{B}] \rangle|. \quad (1.88)$$

ΔO bezeichnet dabei die Varianz einer Observablen O . Die Unschärferelation lehrt uns, daß zwei Observablen unabhängig vom Systemzustand nur dann simultan scharfe Werte annehmen können, wenn die entsprechenden Operatoren vertauschen.

Der Erwartungswert einer Observablen für ein System, welches im Zustand $|\psi\rangle$ präpariert ist, ist $\langle \psi | \mathbf{O} | \psi \rangle$. Daß diese Vorschrift verträglich mit dem Postulat (Q2) ist, resultiert aus der Anwendung der Vollständigkeitsrelation (1.86):

$$\langle \psi | \mathbf{O} | \psi \rangle = \int do \langle \psi | \mathbf{O} | o \rangle \langle o | \psi \rangle = \int do o |\langle \psi | o \rangle|^2. \quad (1.89)$$

Dabei sind wir der Kürze der Schreibweise wegen von einem rein kontinuierlichen Spektrum ausgegangen. Dies zeigt, daß in der Tat der Erwartungswert wie in (1.20) definiert ist, denn $|\langle \psi | o \rangle|^2$ ist ja gemäß der obigen Erörterungen die Wahrscheinlichkeitsdichte dafür, daß eine Messung der Observablen O an dem System, welches im Zustand $|\psi\rangle$ präpariert ist, das Meßresultat o ergibt.

Schließlich kommen wir zur *Dynamik* quantenmechanischer Zustände und Operatoren.

(Q3) Die *Zeit* ist im Rahmen der Quantentheorie ein reeller Parameter. Weiter nehmen wir an, daß es einen selbstadjungierten Operator \mathbf{H} gibt, so daß der Zeitableitung einer Observablen O der Operator

$$\dot{\mathbf{O}} = \frac{1}{i} [\mathbf{O}, \mathbf{H}] + \partial_t \mathbf{O} \quad (1.90)$$

zugeordnet wird. Die partielle Zeitableitung ∂_t trägt einer eventuellen expliziten Zeitabhängigkeit des Operators \mathbf{O} Rechnung.

(Q4) Ist $|\psi, t\rangle$ der Zustand des Systems zur Zeit t , so gilt

$$\dot{\mathbf{P}}_{|\psi, t\rangle} = 0, \quad \text{mit } \mathbf{P}_{|\psi, t\rangle} = |\psi\rangle \langle \psi|, \quad \langle \psi | \psi \rangle = 1. \quad (1.91)$$

Bemerkungen: Die Operatoren, die den grundlegenden dynamischen Observablen des Systems, wie Orts- und Impulsoperatoren, besitzen keine explizite Zeitabhängigkeit. Weiter ist es wichtig zu bemerken, daß $\dot{\mathbf{O}}$ nicht mit der totalen (mathematischen) Zeitableitung $d\mathbf{O}/dt$ des Operators verwechselt werden darf. Beide Operationen sind nur im *Heisenbergbild* der Zeitentwicklung identisch. Wir kommen auf verschiedene Bilder der Zeitentwicklung gleich noch zurück, denn wie man die Zeitabhängigkeiten auf Operatoren und Zustände verteilt, ist weitgehend willkürlich und wird je nach Anwendungsfall geschickt festgelegt. Die physikalischen Aussagen der Theorie sind von dieser Wahl des Bildes natürlich unabhängig. Die Freiheit der Bildwahl ist dadurch begründet, daß alle physikalischen Vorhersagen invariant unter zeitabhängigen unitären Transformationen, sog. *Bildtransformationen* von Operatoren und Zuständen sind. Man kann daher $\dot{\mathbf{O}}$ als *bzgl. Bildtransformationen kovariante Zeitableitung* auffassen.

Wie wir unten noch sehen werden, bedeutet (Q4), daß die Quantentheorie eine kausale Theorie ist, d.h. ist der Zustand zur Zeit t_0 und der Hamiltonoperator bekannt, ist der Zustand zu jeder späteren Zeit determiniert. Das bedeutet, daß wenn $|\psi, t\rangle$ und $|\phi, t\rangle$ zwei Zustände sind, die sich zeitlich gemäß der Dynamik des Systems bewegen, die Wahrscheinlichkeitsamplituden $\langle\psi, t|\phi, t\rangle = \text{const}$ sind.

1.7 Wahl des Bildes der Zeitentwicklung

Jetzt vergegenwärtigen wir uns, wie aus den obigen Postulaten die mathematische Zeitabhängigkeit von Operatoren und Zuständen festzulegen ist. Wie oben schon erwähnt kann diese nur bis auf zeitabhängige unitäre Ähnlichkeitstransformationen von Zuständen und Operatoren festliegen. Eine bestimmte Wahl der Zeitentwicklung definiert das *Bild* der Zeitentwicklung.

Theorem 1. *Das Bild der Zeitentwicklung ist eindeutig durch die Wahl eines hermiteschen Operators \mathbf{X} , der eine Funktion der Zeit (kein Funktional der Vorgeschichte des Systems) sein darf, festgelegt.*

Dieser Operator ist Erzeuger der Zeitentwicklung für nicht explizit zeitabhängige (fundamentale) Operatoren, die Observablen beschreiben. Er definiert den unitären Zeitentwicklungsoperator für Operatoren durch das Anfangswertproblem

$$i\partial_t \mathbf{A}(t, t_0) = -\mathbf{X}(t)\mathbf{A}(t, t_0), \quad \mathbf{A}(t_0, t_0) = \mathbf{1}. \quad (1.92)$$

Dann gilt für alle Observablen, die nicht explizit zeitabhängig sind.

$$\mathbf{O}(t) = \mathbf{A}(t, t_0)\mathbf{O}(t_0)\mathbf{A}^\dagger(t, t_0). \quad (1.93)$$

Der Erzeugungsoperator für die Zeitentwicklung von Zuständen ist dann notwendig durch den hermiteschen Operator $\mathbf{Y} = \mathbf{H} - \mathbf{X}$ gegeben, wobei \mathbf{H} der in (1.90) definierte Hamiltonoperator des Systems ist. Der unitäre Zeitentwicklungsoperator für Zustände ist dann durch die Differentialgleichung

$$i\partial_t \mathbf{C}(t, t_0) = +\mathbf{Y}(t)\mathbf{C}(t, t_0), \quad \mathbf{C}(t_0, t_0) = \mathbf{1} \quad (1.94)$$

gegeben. Dieser unitäre Operator definiert die Zeitentwicklung der Zustände wie folgt:

$$|\psi, t\rangle \langle\psi, t| = \mathbf{C}(t, t_0) |\psi, t_0\rangle \langle\psi, t_0| \mathbf{C}^\dagger(t, t_0). \quad (1.95)$$

Beweis. Es sind zwei Behauptungen zu zeigen: Erstens müssen wir uns versichern, daß die Bewegungsgleichungen für die Zeitentwicklungsoperatoren in sich konsistent, d.h. mit den Anforderungen an einen Zeitentwicklungsoperator für die Observablen-Operatoren verträglich, sind.

Zum zweiten sollte die Zeitableitung von Erwartungswerten von Observablen bildunabhängig sein und mit dem Erwartungswert der „kovarianten Zeitableitung“ (1.90) übereinstimmen, wobei sich im gegebenen Falle die Kovarianz auf allgemeine Bildtransformationen bezieht, d.h. auf die Änderung der Wahl von \mathbf{X} mit der im Satz oben angegebenen Bedeutung.

Zum ersten Punkt: Da die Kommutatorrelationen von Operatoren zu gleichen Zeiten durch physikalische Symmetrieeigenschaften (z.B. Raumzeit-Symmetrien) definiert sind, müssen diese unter der Zeitentwicklung erhalten bleiben, d.h. es muß eine Ähnlichkeitstransformation der Art $\mathbf{O}(t) = \mathbf{A}(t, t_0)\mathbf{O}(t_0)\mathbf{A}^{-1}(t, t_0)$ vorliegen, wobei der Zeitentwicklungsoperator \mathbf{A} insbesondere invertierbar sein muß.

Da weiter die Operatoren, die die Observablenalgebra konstituieren, allesamt zu jeder Zeit t selbstadjungiert sein müssen, sollte $\mathbf{A}^{-1} = \mathbf{A}^\dagger$, d.h. \mathbf{A} unitär, sein.

Für $t > t_0$ sollte der Zeitentwicklungsoperator $\mathbf{A}(t, t_0)$ eine Funktion von t und t_0 allein sein. Betrachten wir jetzt zunächst die Zeitentwicklung der Operatoren vom Zeitpunkt t_0 zu einem Zeitpunkt t_1 mit $t_0 < t_1$. Wir können die so gewonnenen Operatoren $\mathbf{O}(t_1)$ als neue Anfangsbedingungen für die Zeitentwicklung ansehen und diese wiederum mit dem zugehörigen Zeitentwicklungsoperator $\mathbf{A}(t_2, t_1)$ zum Operator zur Zeit $t_2 > t_1$ transformieren. Diese Hintereinanderausführung von zwei Zeitentwicklungen muß aber andererseits auch durch den einzelnen Zeitentwicklungsoperator $\mathbf{A}(t_2, t_0)$ gegeben sein. Es folgt also als Konsistenzbedingung:

$$\forall t_0 < t_1 < t_2 \in \mathbb{R} : \mathbf{A}(t_2, t_1)\mathbf{A}(t_1, t_0) = \mathbf{A}(t_2, t_0). \quad (1.96)$$

Aus der Unitarität von $\mathbf{A}(t, t_0)$ schließen wir

$$\mathbf{A}\mathbf{A}^\dagger = \mathbf{1} = \text{const.} \Rightarrow (i\partial_t\mathbf{A})\mathbf{A}^\dagger = -\mathbf{A}(i\partial_t\mathbf{A})^\dagger, \quad (1.97)$$

so daß der Operator $\mathbf{X} = -i(\partial_t\mathbf{A})\mathbf{A}^\dagger$ selbstadjungiert ist: $\mathbf{X}^\dagger = \mathbf{X}$. Aus (1.96) folgern wir

$$[i\partial_t\mathbf{A}(t, t_0)]\mathbf{A}^\dagger(t, t_0) = [i\partial_t\mathbf{A}(t, t_1)]\mathbf{A}^\dagger(t, t_1) := -\mathbf{X}(t), \quad (1.98)$$

woraus klar wird, daß \mathbf{X} nicht von der Anfangszeit t_0 der Zeitentwicklung abhängt, sondern tatsächlich eine Funktion der Zeit t allein sein muß, d.h. die Zeitabhängigkeit von \mathbf{X} ist *lokal*.

Der Beweis, daß auch \mathbf{Y} , der Generator der Zeitentwicklung für die Zustände also, die gleichen Eigenschaften besitzt, ist vollkommen analog zu führen und sei dem Leser zur Übung überlassen.

Zum zweiten Problem: Wir müssen zeigen, daß die Zeitentwicklung mit den im vorigen Abschnitt angegebenen Postulaten der Quantentheorie konsistent ist: Aus (1.93) erhalten wir unter Verwendung der Definition (1.92) für einen Operator, der auch explizit zeitabhängig sein kann:

$$\frac{d\mathbf{O}(t)}{dt} = \frac{1}{i} [\mathbf{O}(t), \mathbf{X}(t)] + \partial_t\mathbf{O}(t). \quad (1.99)$$

Diese Gleichung kann aber ihrerseits in Form der bildinvarianten Zeitableitung (1.90) geschrieben werden:

$$\frac{d\mathbf{O}(t)}{dt} = \dot{\mathbf{O}} - \frac{1}{i} [\mathbf{O}, \mathbf{H} - \mathbf{X}]. \quad (1.100)$$

Wir sehen, daß die Gleichungen (1.99) und (1.100) zusammen mit vorgegebenen Anfangswerten für den Operator zur Zeit t_0 eindeutig durch die Anwendung des unitären Zeitevolutionsoperator gelöst werden, welcher die Gleichung (1.92) erfüllt.

Wie jeder Operator erfüllt auch der statistische Operator diese Gleichung. Wegen (1.91), also der Annahme, daß der statistische Operator kovariant konstant ist, weil er beständig dem System zugeordnet sein soll, ergibt sich daraus

$$\frac{d\mathbf{P}_{|\psi,t\rangle}}{dt} = -\frac{1}{i} [\mathbf{P}_{|\psi,t\rangle}, \mathbf{Y}] \quad (1.101)$$

Diese Gleichung wird eindeutig durch einen unitären Zeitentwicklungsoperator \mathbf{C} , welcher (1.94) erfüllt, gelöst.

Es folgt daraus, daß sich (bis auf einen unerheblichen Phasenfaktor) der Zustandsket gemäß der unitären Zeitentwicklung

$$|\psi, t\rangle = \mathbf{C}(t, t_0) |\psi, t_0\rangle \quad (1.102)$$

bewegt. Insbesondere bleibt der Zustand vermöge der Zeitentwicklung ein reiner Zustand, und die Normierung bleibt erhalten. □

1.8 Formale Lösung der Bewegungsgleichungen

Zum späteren Gebrauch ist es nützlich zu wissen, wie man die Zeitentwicklungsoperatoren formal aus ihren Generatoren bestimmen kann. Betrachten wir als Beispiel die Bestimmung von \mathbf{A} durch Lösung der Bewegungsgleichung (1.92).

Das Hauptproblem ergibt sich daraus, daß der selbstadjungierte Operator $\mathbf{X}(t)$ von der Zeit abhängt, und diese Operatoren zu verschiedenen Zeiten nicht notwendig vertauschen. Deshalb können wir nicht einfach die Gleichung so lösen, wie wir es von Differentialgleichungen gewöhnlicher Funktionen in \mathbb{C} oder \mathbb{R} her gewohnt sind.

Als erstes können wir unter Verwendung der Anfangsbedingung (1.92) integrieren und erhalten so die lineare Integralgleichung

$$\mathbf{A}(t, t_0) = \mathbf{1} + i \int_{t_0}^t d\tau \mathbf{X}(\tau) \mathbf{A}(\tau, t_0). \quad (1.103)$$

Diese Integralgleichung läßt sich formal durch die folgende Iteration lösen:

$$\mathbf{A}_n(t, t_0) = \mathbf{1} + i \int_{t_0}^t \mathbf{X}(\tau) \mathbf{A}_{n-1}(\tau, t_0) d\tau, \quad \mathbf{A}_0(t, t_0) = \mathbf{1}. \quad (1.104)$$

Die Lösung der Gleichung sollte durch den formalen Limes $n \rightarrow \infty$ bestimmt sein. Wir wollen hier nicht auf die komplizierten Konvergenzfragen einer solchen Iteration eingehen.

Durch vollständige Induktion kann man leicht zeigen, daß die explizite Lösung die Form

$$\begin{aligned} \mathbf{A}(t, t_0) &= \sum_{k=0}^{\infty} \mathbf{A}^{(k)}(t, t_0) \text{ mit} \\ \mathbf{A}^{(k)}(t, t_0) &= \int_{t_0}^t d\tau_1 \int_{t_0}^{\tau_1} d\tau_2 \dots \int_{t_0}^{\tau_{k-1}} d\tau_k \mathbf{X}(\tau_1) \mathbf{X}(\tau_2) \dots \mathbf{X}(\tau_k) \end{aligned} \quad (1.105)$$

besitzt.

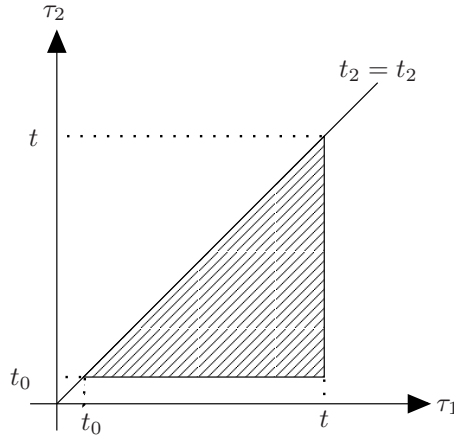


Abbildung 1.1: Integrationsbereich in (1.106)

Diese Reihe läßt sich noch vereinfachen. Betrachten wir zunächst $\mathbf{A}^{(2)}$:

$$\int_{t_0}^t d\tau_1 \int_{t_0}^{\tau_1} d\tau_2 \mathbf{X}(\tau_1) \mathbf{X}(\tau_2). \quad (1.106)$$

Der Integrationsbereich dieses Doppelintegrals ist das in Abb. 1.1 gezeigte Dreieck in der $\tau_1\tau_2$ -Ebene. Unter Anwendung des Theorems von Fubini können wir die Integrationsreihenfolge vertauschen

$$\mathbf{A}^{(2)} = \int_{t_0}^t d\tau_1 \int_{\tau_1}^t d\tau_2 \mathbf{X}(\tau_2) \mathbf{X}(\tau_1). \quad (1.107)$$

Es zeigt sich, daß in (1.106) und (1.107) die Operatoren stets so angeordnet sind, daß die Operatoren zu späteren Zeitpunkten weiter links zu stehen kommen. Für diese Anordnungsvorschrift definieren wir den *Zeitordnungsoperator* \mathcal{T}_c . Damit schreibt sich dann die Summe aus (1.106) und (1.107)

$$2\mathbf{A}^{(2)}(t, t_0) = \mathcal{T}_c \int_{t_0}^t d\tau_1 \int_{t_0}^t d\tau_2 \mathbf{X}(\tau_1) \mathbf{X}(\tau_2). \quad (1.108)$$

Diese Betrachtung läßt sich für alle Glieder der Reihe (1.105) verallgemeinern:

$$\mathbf{A}^{(k)}(t, t_0) = \frac{1}{k!} \mathcal{T}_c \int_{t_0}^t d\tau_1 \cdots \int_{t_0}^t d\tau_n \mathbf{X}(\tau_1) \cdots \mathbf{X}(\tau_n). \quad (1.109)$$

Dies läßt sich durch vollständige Induktion wie folgt nachweisen: Angenommen die Behauptung sei wahr für $k = n - 1$, d.h. wir können (1.109) auf die ersten $n - 1$ Integrationen, die $\mathbf{A}^{(n)}$ definieren, anwenden:

$$\mathbf{A}^{(n)}(t, t_0) = \frac{1}{(n-1)!} \int_{t_0}^t d\tau_1 \mathbf{X}(\tau_1) \mathcal{T}_c \int_{t_0}^{\tau_1} d\tau_2 \cdots \int_{t_0}^{\tau_1} d\tau_n \mathbf{X}(\tau_2) \cdots \mathbf{X}(\tau_n). \quad (1.110)$$

Jetzt können wir die gleiche Rechnung, die wir für $\mathbf{A}^{(2)}$ ausgeführt haben, auf das äußerste Integral anwenden, und zwar jeweils mit jedem der inneren Integrale. Addieren wir alle möglichen Paarungen einschließlich zum ursprünglichen Integral (1.110) und dividieren durch n ergibt sich die Behauptung für $k = n$:

$$\mathbf{A}^{(n)}(t, t_0) = \frac{1}{n!} \mathcal{T}_c \int_{t_0}^t d\tau_1 \cdots \int_{t_0}^t d\tau_n \mathbf{X}(\tau_1) \cdots \mathbf{X}(\tau_n). \quad (1.111)$$

Die Reihe (1.105) kann also formal durch

$$\mathbf{A}(t, t_0) = \mathcal{T}_c \exp \left[i \int_{t_0}^t d\tau \mathbf{X}(\tau) \right] \quad (1.112)$$

summiert werden. Genauso finden wir für den Operator $\mathbf{C}(t, t_0)$

$$\mathbf{C}(t, t_0) = \mathcal{T}_c \exp \left[-i \int_{t_0}^t d\tau \mathbf{Y}(\tau) \right]. \quad (1.113)$$

1.9 Ein Beispiel: Das freie Teilchen

Nur in den einfachsten Fällen lassen sich die Zeitentwicklungsoperatoren explizit angeben. Ein Beispiel, für das dies gelingt, ist ein freies Teilchen. Der Einfachheit halber betrachten wir nur die Bewegung in einer Raumrichtung. Die Observablenalgebra ist dann durch Orts- und Impulsoperator definiert, die die Heisenbergkommutatorrelation

$$\frac{1}{i} [\mathbf{x}, \mathbf{p}] = \mathbf{1} \quad (1.114)$$

erfüllen. Man kann sie am sichersten aus der Raumzeitsymmetrie begründen, wonach \mathbf{p} der Erzeuger von Raumzeittranslationen in Analogie zur klassischen Mechanik unter Anwendung des Noethertheorems ist. Für eine genauere Herleitung der Vertauschungsrelation aus den Symmetrieprinzipien sei wieder auf [Hee98] verwiesen.

Wir legen nun das *Heisenbergbild* der Zeitentwicklung zugrunde, welches im gegebenen Falle besonders leicht zum Ziele führt. In diesem Bild werden die Generatoren der Zeitentwicklung so gewählt, daß $\mathbf{C} = 1$, also $\mathbf{Y} = 0$ ist. Damit gilt $\mathbf{X} = \mathbf{H}$, wobei der Hamiltonoperator durch

$$\mathbf{H} = \frac{\mathbf{p}^2}{2m}, \quad (1.115)$$

definiert ist. Dieses Bild zeichnet sich dadurch aus, daß die kovariante (physikalische) Zeitableitung mit der totalen mathematischen Zeitableitung der Operatoren übereinstimmt.

Aus (1.99) entnehmen wir die Zeitentwicklung der Observablen \mathbf{x} und \mathbf{p} :

$$\frac{d\mathbf{p}}{dt} = \frac{1}{i} [\mathbf{p}, \mathbf{H}] = 0, \quad \frac{d\mathbf{x}}{dt} = \frac{1}{i} [\mathbf{x}, \mathbf{H}] = \frac{\mathbf{p}}{m}. \quad (1.116)$$

Dies ist den klassischen Bewegungsgleichungen weitestgehend analog, nur daß wir hier Operatoren und nicht reelle Zahlen vorliegen haben. Da allerdings \mathbf{p} nach der ersten Bewegungsgleichung zeitlich konstant ist, ist die Lösung der Operatorgleichung vollkommen analog zur Lösung im klassischen Falle mit reellen Zahlen. Es treten nämlich keine Operatorordnungsprobleme für Operatoren zu verschiedenen Zeiten auf. Es ist also

$$\mathbf{p}(t) = \mathbf{p}(0) = \text{const}, \quad \mathbf{x}(t) = \mathbf{x}(0) + \frac{\mathbf{p}(0)}{m}t. \quad (1.117)$$

Dabei haben wir ohne Beschränkung der Allgemeinheit den Anfangszeitpunkt $t_0 = 0$ gewählt.

Betrachten wir nun die Zeitentwicklung der Wellenfunktion, welche durch die Matrixelemente des Zustandskets und eines vollständigen Satzes (verallgemeinerter) Eigenkets gegeben wird. Es ist klar, daß die Zeitabhängigkeit von Matrixelementen *bidunabhängig* ist. Hier benutzen wir

zur Berechnung des Heisenbergbild. Die gesamte Zeitabhängigkeit kommt dann vom Eigenket der gewählten Basis, weil ja die Observablenoperatoren die volle Zeitabhängigkeit tragen, während die Zustandskets zeitlich konstant sind. Aus (1.116) erhalten wir, abgesehen von einem irrelevanten Phasenfaktor,

$$|p, t\rangle = \exp(i\mathbf{H}t) |p, 0\rangle = \exp\left(i\frac{p^2}{2m}t\right) |p, 0\rangle, \quad (1.118)$$

und die Zeitentwicklung der *Wellenfunktion* ist einfach

$$\psi(p, t) = \langle p, t | \psi \rangle = \exp\left(-i\frac{p^2}{2m}t\right) \psi(p, 0). \quad (1.119)$$

Dies kann durch einen linearen Integraloperator beschrieben werden

$$\psi(p, t) = \int dp' \underbrace{\langle p, t | p', 0 \rangle}_{U(t, p; 0, p')} \langle p', 0 | \psi \rangle = \int dp' U(t, p; 0, p') \psi(p', 0). \quad (1.120)$$

Aus (1.117) erhalten wir

$$U(t, p, 0, p') = \exp\left(-i\frac{p^2}{2m}t\right) \delta(p - p'). \quad (1.121)$$

Es ist wichtig zu bemerken, daß wie in diesem Beispiel offensichtlich, diese *Propagatoren* i.a. Distributionen, keine Funktionen sind.

Als nächstes wollen wir den Propagator in der Ortsdarstellung berechnen. Dazu gibt es im Prinzip zwei Ansätze: Zum einen kann man (1.121) mittels einer Fouriertransformation in die Ortsdarstellung umrechnen, zum anderen aber auch wieder direkt die Heisenbergdarstellung zur Bestimmung des zeitabhängigen Eigenvektors verwenden:

$$\mathbf{x}(t) |x, t\rangle = \left(\mathbf{x}(0) + \frac{\mathbf{p}(0)}{m}t\right) |x, t\rangle = x |x, t\rangle. \quad (1.122)$$

Multiplizieren dieser Gleichung mit $\langle x', 0 |$ ergibt unter Verwendung der Ortsdarstellung des Impulsoperators $\langle x', 0 | \mathbf{p}(0) | \psi \rangle = -i\partial_{x'}\psi(x')$:

$$(x' - x) \langle x', 0 | x, t \rangle = \frac{it}{m} \partial_{x'} \langle x', 0 | x, t \rangle. \quad (1.123)$$

Diese Gleichung wird offensichtlich formal durch

$$U(t, x; 0, x')^* = \langle x', 0 | x, t \rangle = N \exp\left[-i\frac{m}{2t}(x' - x)^2\right] \quad (1.124)$$

gelöst. Zur Bestimmung des Normierungsfaktors ist die Anfangsbedingung

$$U(0, x; 0, x') = \delta(x - x') \quad (1.125)$$

zu berücksichtigen. Da die Zeitentwicklung unitär ist, ändert sich das Normierungsintegral nicht:

$$\int dx' U(0, x; t, x') = 1. \quad (1.126)$$

Um dieses Integral unter Verwendung von (1.124) zu berechnen, müssen wir es zunächst regularisieren, um die Distribution als schwachen Limes einer integrierbaren Funktion zu erhalten. Dies kann einfach durch die Analytische Fortsetzung der Funktion zu komplexen Zeiten $t \rightarrow t - i\epsilon$ mit $\epsilon > 0$ erreicht werden. Nachdem das Integral berechnet und N bestimmt wurde, ist dann der Limes $t \rightarrow 0^+$ zu nehmen. Die einfache Rechnung sei dem Leser überlassen. Das endgültige Resultat lautet:

$$U(t, x; 0, x') = \sqrt{\frac{mi}{2\pi t}} \exp \left[-i \frac{m}{2t} (x' - x)^2 \right]. \quad (1.127)$$

Auch bei der alternativen Rechnung als Fouriertransformierte der Impulsdarstellung (1.121) bedient man sich selbstverständlich mit Gewinn derselben Regularisierung.

Eine dritte Möglichkeit ist, daß wir von einer Darstellung, in der auch die Zeit fouriertransformiert (die konjugierte Variable ist hier die Frequenz ω) wird, ausgehen und den Propagator direkt als Greensche Funktion der zeitabhängigen freien Schrödingergleichung bestimmen. Hier kommt die Problematik mit den Distributionen durch die Frage, wie das Fourierintegral zur „Zurückrechnung“ in den Zeitbereich zu definieren ist, ins Spiel, weil charakteristischerweise Pole auf der reellen ω -Achse auftauchen. Hierbei wird diese Frage eindeutig dadurch gelöst, daß U kausal sein muß, also nur Zeiten $< t$ in die Berechnung der Wellenfunktion zur Zeit t eingehen dürfen. In dem jetzigen Falle der nichtrelativistischen Quantentheorie ergibt dies Randbedingungen (bzw. „ $i\eta$ -Vorschriften“), die zur *retardierten Greenschen* Funktion führen.

Wir werden auf diese Rechnungen im Rahmen der feldtheoretischen Beschreibung, zu der wir im übernächsten Kapitel übergehen werden, weil sie die natürliche Grundlage zur Beschreibung von Vielteilchensystemen ist und zudem auch auf den relativistischen Fall verallgemeinerbar ist, noch ausführlich zu sprechen kommen.

Bemerkung: Die Schrödingergleichung ist eine partielle Differentialgleichung, die die Zeitentwicklung der *Wellenfunktion* für ein (nichtrelativistisches) Teilchen beschreibt. Diese Gleichung ist *bildunabhängig*, d.h. sie läßt sich in jedem Bild der Zeitentwicklung herleiten.

Betrachten wir als Beispiel die Schrödingergleichung der Wellenfunktion in der Ortsdarstellung in einem beliebigen Bild der Zeitentwicklung. Die Ortseigenvektoren sind zeitabhängig, da sich gemäß (1.93) der Ortsoperator gemäß

$$\mathbf{x}(t) = \mathbf{A}(t, t_0) \mathbf{x}(t_0) \mathbf{A}^\dagger(t, t_0) \quad (1.128)$$

bewegt (wobei wir benutzt haben, daß \mathbf{x} nicht explizit zeitabhängig ist, sondern nur durch diese unitäre Ähnlichkeitstransformation zeitabhängig ist). Damit folgt

$$\mathbf{x}(t) |x, t\rangle = x |x, t\rangle \Rightarrow |x, t\rangle = \mathbf{A}(t, t_0) |x, t_0\rangle. \quad (1.129)$$

Verwenden wir nun (1.102) zur Zeitentwicklung der Zustandskets, folgt

$$\psi(t, x) = \langle x, t | \psi, t \rangle = \langle x, t_0 | \mathbf{A}^\dagger(t, t_0) \mathbf{C}(t, t_0) | \psi, t_0 \rangle. \quad (1.130)$$

Bilden wir nun die Zeitableitung, folgt wegen (1.92) und (1.94)

$$i\partial_t \psi(t, x) = \langle x, t_0 | \mathbf{A}^\dagger(t, t_0) [\mathbf{X}(t) + \mathbf{Y}(t)] \mathbf{C}(t, t_0) | \psi, t_0 \rangle \quad (1.131)$$

Verwenden wir nun $\mathbf{X} + \mathbf{Y} = \mathbf{H}$, (1.129) und (1.102), folgt sogleich

$$i\partial_t \psi(t, x) = \langle x, t | \mathbf{H}(t) | \psi, t \rangle. \quad (1.132)$$

Für das freie Teilchen ergibt sich wegen $\mathbf{p}(t) |x, t\rangle = i\partial_x |x, t\rangle$ (übrigens ist dies eine ebenfalls bildunabhängige Gleichung) damit die Schrödingergleichung des freien Teilchens:

$$i\partial_t\psi(t, x) = -\frac{1}{2m}\partial_x^2\psi(t, x). \quad (1.133)$$

1.10 Quantenstatistik

Jetzt gehen wir zur Quantenstatistik, dem eigentlichen Gegenstand dieses Artikels, über. Wir werden hier genau darstellen, wie wir die in den vorigen Abschnitten gegebenen mathematischen Grundlagen der Statistik und der Informationstheorie auf die Quantentheorie eines Systems, dessen exakter quantenmechanischer Zustand $[[\psi]]$ zur Anfangszeit nicht bekannt ist, aufgrund eventuell vorhandener Information mit Hilfe des Jaynesschen Prinzips des geringsten Vorurteils statistisch beschreiben können.

Wir müssen dazu natürlich zunächst erst einmal definieren, wie die „Zufallsexperimente“ in der Quantentheorie überhaupt beschrieben werden können. In der Quantenwelt sind allerdings verschiedene Arten statistischer Beschreibung zu unterscheiden. Die eine beruht auf dem grundlegend *indeterministischen* Charakter der Quantentheorie. Sie wurde oben bereits beschrieben: Selbst wenn wir den Zustand $[[\psi]]$ des Systems exakt kennen, können wir i.a. nicht vorhersagen, welchen Meßwert o einer Observablen O wir finden werden. Vielmehr sind die Wahrscheinlichkeiten für das Vorfinden eines bestimmten Wertes für die Observable O durch (1.81) bestimmt.

1.11 Gemische und statistische Operatoren

Die Quantenstatistik beschäftigt sich jedoch mit der Frage, wie wir Systeme beschreiben können, von denen der Zustand nicht exakt bekannt ist. Wir können uns etwa vorstellen, daß ein Ensemble des Systems vorliegt, das aus verschiedenen *voneinander unabhängigen Quellen* Q_j ($j \in \{1, \dots, n\}$) stammt, von denen jede das System in einem exakt bekannten Zustand $[[\psi_j]]$ liefert. Die relativen Intensitäten der jeweiligen Quellen seien mit λ_j bezeichnet. Man spricht in einem solchen Falle von Präparation eines Ensembles oft von einem *Gemisch* und bezeichnet den dabei präparierten Zustand entsprechend als *gemischten Zustand*¹.

Es ist zwar $\sum_{j=1}^n \lambda_j = 1$, aber wichtig zu bemerken, daß die λ_j i.a. *nicht die Wahrscheinlichkeiten sind, ein System aus dem Ensemble im Zustand $[[\psi_j]]$ zu finden*, denn die $|\psi_j\rangle$ müssen nicht orthogonal sein, und $\mathcal{P}\{[[\psi_j]]\}_{j \in \{1, \dots, n\}}$ kann dann *keine Ereignisalgebra* mit $\Omega = \{[[\psi_j]]\}_{j \in \{1, \dots, n\}}$ als Ergebnismenge bilden, weil sich die Zustände eben nicht gegenseitig ausschließen, wenn die $|\psi_j\rangle$ nicht orthogonal zueinander sind.

Jetzt betrachten wir einen beliebigen Zustand $[[\phi]]$ und fragen nach der Wahrscheinlichkeit, einen aus dem Gemisch entnommenen Vertreter des Systems in diesem Zustand anzutreffen. Die dazugehörige Observable ist $\mathbf{O} = |\phi\rangle\langle\phi|$. Wegen $\mathbf{O} = \mathbf{O}^2$ sind die möglichen Eigenwerte 0 und 1. Wird der Eigenwert 1 gemessen, ist das System im Zustand $[[\phi]]$ gefunden worden, andernfalls nicht.

Wir wissen, daß sich der Vertreter des Gemischs in einem der Zustände $[[\psi_j]]$ befindet. Ist das der Fall, ist gemäß (1.81) die Wahrscheinlichkeit, daß das System im Zustand $[[\phi]]$ gefunden

¹Dieser Abschnitt verdankt einer Diskussion in der Newsgroup `de.sci.physik` viele Anregungen. Besonders danken möchte ich in diesem Zusammenhang Norbert Dragon, Roland Franzius und Ilja Schmelzer

wird, $p(1) = |\langle \phi | \psi_j \rangle|^2$, wobei wir annehmen, daß die beteiligten Zustandskets normiert sind. Entsprechend der Anteile der Systeme in dem Gemisch liegt es nahe, die Wahrscheinlichkeit, daß $[[\phi]]$ festgestellt wird, mit

$$p([[\phi]]) = \sum_{j=1}^n \lambda_j \langle \phi | \psi_j \rangle \langle \psi_j | \phi \rangle = \langle \phi | \left(\sum_{j=1}^n \lambda_j |\psi_j\rangle \langle \psi_j| \right) | \phi \rangle \quad (1.134)$$

anzunehmen. Man bezeichnet den in Klammern stehenden Operator als

$$\mathbf{R} = \sum_{j=1}^n \lambda_j |\psi_j\rangle \langle \psi_j| \quad (1.135)$$

als den zu dem präparierten Gemisch gehörigen *statistischen Operator* des Systems. Damit läßt sich (1.134) durch

$$p([[\phi]]) = \langle \phi | \mathbf{R} | \phi \rangle \quad (1.136)$$

ausdrücken.

Die Eigenschaften des Dichteoperators sind schnell geklärt: Es ist offenbar ein *hermitescher Operator*. Wegen $\lambda_j > 0$ ist

$$p([[\phi]]) \geq 0, \quad (1.137)$$

der statistische Operator also *positiv semidefinit*.

Jetzt können wir auch den Erwartungswert einer an den Konstituenten des Gemisches gemessenen Observablen \mathbf{O} angeben:

$$\langle O \rangle = \sum_{n=1}^{\infty} o_k p([[\mathbf{O}, k]]) = \sum_{n=1}^{\infty} o_k \langle \mathbf{O}, k | \mathbf{R} | \mathbf{O}, k \rangle = \sum_{n=1}^{\infty} \langle \mathbf{O}, k | \mathbf{R} \mathbf{O} | \mathbf{O}, k \rangle := \text{Tr}(\mathbf{R} \mathbf{O}). \quad (1.138)$$

Dabei ist $\{|\mathbf{O}, k\rangle\}_{k \in \mathbb{N}_{>0}}$ das vollständige Orthonormalsystem aus Eigenvektoren von \mathbf{O} . Wir beschränken uns dabei auf den Fall eines diskreten Spektrums. Für kontinuierliche Spektren sind die verallgemeinerten Eigenvektoren und die Summen durch die entsprechenden Integrale zu ersetzen.

Die Spur eines Operators ist nun aber unabhängig vom gewählten vollständigen Orthogonalsystem immer durch die Summe der Diagonalelemente des entsprechenden Operators gegeben. Mögen dazu $|\phi_k\rangle$ und $|\psi_k\rangle$ zwei beliebige vollständige Orthonormalsysteme bilden. Dann ist für einen Operator \mathbf{A}

$$\begin{aligned} \sum_{k=1}^{\infty} \langle \psi_k | \mathbf{A} | \psi_k \rangle &= \sum_{k,l,m=1}^{\infty} \langle \psi_k | \phi_l \rangle \langle \phi_l | \mathbf{A} | \phi_m \rangle \langle \phi_m | \psi_k \rangle \\ &= \sum_{l,m=1}^{\infty} \delta_{lm} \langle \phi_l | \mathbf{A} | \phi_m \rangle = \sum_{l=1}^{\infty} \langle \phi_l | \mathbf{A} | \phi_l \rangle. \end{aligned} \quad (1.139)$$

Wir kommen nun zu der Frage, wie im Rahmen der Quantentheorie Zufallsexperimente im Sinne der im Abschnitt 1 besprochenen mathematischen Statistik zu formalisieren sind.

Es ist klar, daß der Fall, daß das System in einem reinen Zustand $[[\psi]]$ präpariert wurde, wie er in (Q1) postuliert wurde, unter dieses allgemeinere Schema fällt. Es ist der statistische Operator dann durch

$$\mathbf{R} = |\psi\rangle \langle \psi| \quad (1.140)$$

gegeben.

Ein Zufallsexperiment an einem Quantensystem ist stets die Messung einer oder die simultane Messung mehrerer Observabler. Das ergibt nur dann einen Sinn, wenn im letzteren Falle die Observablen miteinander verträglich sind. Im maximalen Falle, kann ein vollständiger Satz miteinander kompatibler Observabler gemessen werden.

Jetzt sind wir in der Lage, Zufallsexperimente an Quantensystemen eindeutig zu formalisieren: Die *Ergebnismenge* Ω für dieses Zufallsexperiment sind die möglichen Meßwerte der Observablen O , also die Werte im Spektrum des dazugehörigen selbstadjungierten Operators \mathbf{O} : $\Omega = \text{Spec}(\mathbf{O}) \subseteq \mathbb{R}$. Im folgenden gehen wir der Einfachheit von einem diskreten Spektrum aus. Die Verallgemeinerung auf kontinuierliche Spektren ist wieder in der üblichen Weise zu vollziehen.

Die Ereignisalgebra ist dann die Potenzmenge $E = \mathcal{P}\Omega$.

Jetzt müssen wir nur die Wahrscheinlichkeiten für die Elementarereignisse bestimmen. Ist das Ensemble wie oben beschrieben in dem durch \mathbf{R} beschriebenen Gemisch präpariert, dann ist nach (1.136) die Wahrscheinlichkeit, ein aus ihm herausgegriffenes System im Zustand $|(o_j), k\rangle$ anzutreffen $\langle(o_j), k | \mathbf{R} | (o_j), k\rangle$. Dabei werden die den Eigenraum $\text{Eig}[(\mathbf{O}_j), (o_j)]$ aufspannenden simultanen Eigenvektoren $|(o_j), k\rangle$ der Operatoren (\mathbf{O}_j) orthonormiert gewählt. Dann schließen sich die Ereignisse $\{(o_j)\}$ nämlich gegenseitig aus, und es gilt für die Wahrscheinlichkeit, die simultanen Eigenwerte (o_j) zu messen:

$$p[(o_j)] = \sum_k \langle(o_j), k | \mathbf{R} | (o_j), k\rangle. \quad (1.141)$$

Das läßt sich mit Hilfe des Projektors

$$\mathbf{P}[(o_j)] = \sum_k |(o_j), k\rangle \langle(o_j), k| \quad (1.142)$$

auf den Eigenraum $\text{Eig}[(\mathbf{O}_j), (o_j)]$ in der Form

$$p[(o_j)] = \text{Tr}(\mathbf{R}\mathbf{P}[(o_j)]) \quad (1.143)$$

schreiben.

Der Projektor $\mathbf{P}(o_j)$ ist unabhängig von der Wahl der Basis des Eigenraums. Sei nämlich $|(o_j)', k'\rangle$ eine andere Basis, dann gilt

$$\begin{aligned} \sum_{k'} |(o_j)', k'\rangle \langle(o_j)', k'| &= \sum_{k_1, k_2, k'} |(o_j), k_1\rangle \langle(o_j), k_1 | (o_j)', k'\rangle \langle(o_j)', k' | (o_j), k_2\rangle \langle(o_j), k_2| \\ &= \sum_{k_1, k_2} |(o_j), k_1\rangle \underbrace{\langle(o_j), k_1 | (o_j), k_2\rangle}_{\delta(k_1, k_2)} \langle(o_j), k_2| \\ &= \sum_{k_1} |(o_j), k_1\rangle \langle(o_j), k_1| := \mathbf{P}[(o_j)]. \end{aligned} \quad (1.144)$$

Dabei haben wir in der ersten Zeile beim Einschleiben der Zerlegung der 1 durch das VONS $\{|(o_j), k\rangle\}_{(o_j), k}$ ausgenutzt, daß die Eigenräume zu verschiedenen simultanen Eigenwerten (o_j) orthogonal aufeinander stehen. Das gleiche wurde beim Übergang zur zweiten Zeile beachtet, wo für den einen Eigenraum die neue Basis statt der alten benutzt wurde.

Schließlich müssen wir noch zeigen, daß das sichere Ereignis Ω die Wahrscheinlichkeit 1 besitzt:

$$\begin{aligned} \sum_{(o_j)} \mathbf{P}[(o_j)] &= 1 \Rightarrow \\ \sum_{(o_j)} p[(o_j)] &= \text{Tr } \mathbf{R} = \sum_{n,k} \lambda_k \langle \psi_k | n \rangle \langle n | \psi_k \rangle = \sum_k \lambda_k = 1. \end{aligned} \quad (1.145)$$

Aus der Definition (1.134) des statistischen Operators und (1.91) geht unmittelbar hervor, daß der statistische Operator *kovariant konstant* ist, also

$$\dot{\mathbf{R}} = \frac{1}{i} [\mathbf{R}, \mathbf{H}] + \partial_t \mathbf{R} = 0 \quad (1.146)$$

ist. Die mathematische Zeitentwicklung ist gemäß (1.102) durch die unitäre Ähnlichkeitstransformation

$$\mathbf{R}(t) = \mathbf{C}(t, t_0) \mathbf{R}(t_0) \mathbf{C}^\dagger(t, t_0) \quad (1.147)$$

gegeben.

Aus (1.146) läßt sich unmittelbar die wichtige Folgerung ablesen, daß der statistische Operator genau dann nicht explizit von der Zeit abhängt, wenn er eine Funktion der *Erhaltungsgrößen des Systems* ist, also mit dem Hamiltonoperator \mathbf{H} kommutiert. Ein solcher Operator beschreibt ein *statistisch statisches System*. Man sagt auch, solch ein System befinde sich im *Gleichgewicht*. Wir werden unten sehen, daß sich daraus die vollständige *Gleichgewichtsthermodynamik* ableiten läßt.

Zusammenfassend haben wir also folgende Beschreibung von quantenstatistischen Ensembles (Gemischen) gefunden

- (S1) Ein Ensemble von voneinander unabhängigen gleich präparierten Quantensystemen (*Gemisch*) wird durch einen *statistischen Operator* \mathbf{R} beschrieben.
- (S2) Der statistische Operator ist *selbstadjungiert und positiv semidefinit* mit $\text{Tr } \mathbf{R} = 1$.
- (S3) Ein Zufallsexperiment an dem System ist durch die Messung einer oder mehrerer kompatibler Observabler (O_i) definiert. Die Wahrscheinlichkeit, die simultanen Eigenwerte (o_i) zu messen, ist $p[(o_i)] = \text{Tr}\{\mathbf{P}[(o_i)]\mathbf{R}$.
- (S4) Der Erwartungswert einer beliebigen Observablen \mathbf{O} ist $\langle O \rangle = \text{Tr}(\mathbf{O}\mathbf{R})$.

1.12 Die Entropie und das Jaynessche Prinzip

In den seltensten Fällen ist uns das Gemisch in der eben beschriebenen Weise durch Präparation als Gemisch aus Quellen von Systemen exakt bekannter Zustände $[\psi_j]$ gegeben. Die vollständige Kenntnis des Zustandes eines Systems erfordert ja die Messung eines vollständigen Satzes kompatibler Observablen. Nehmen wir etwa den Fall an, daß wir ein Mol eines Gases aus spinlosen Teilchen vorliegen haben. Dann müßten wir etwa die Orte von $\approx 10^{24}$ Teilchen bestimmen. Das ist praktisch unmöglich. Man wird zum Anfangszeitpunkt vielmehr die Mittelwerte einiger weniger globaler Systemparameter bestimmen, z.B. die Gesamtenergie und -impuls oder die *makroskopische* Teilchendichteverteilung etc. Dann kennen wir weder den reinen Zustand des Systems noch den Dichteoperator, der ein Ensemble aus solcherart gleich präparierten Systemen

statistisch beschreibt. Vielmehr muß der Dichteoperator aufgrund der vorliegenden Information *geschätzt* werden. Wir wollen uns wieder des *Jaynesschen Prinzips des geringsten Vorurteils* bedienen, um diese Schätzung vorzunehmen.

Die Entropie entnehmen wir direkt der Definition (1.75). Dazu schreiben wir \mathbf{R} in seinem *Eigenystem* $|\mathbf{R}, n\rangle$. Da \mathbf{R} selbstadjungiert und positiv definit ist, gilt

$$\mathbf{R} = \sum_{n=1}^{\infty} p_n |\mathbf{R}, n\rangle \langle \mathbf{R}, n|, \quad (1.148)$$

wobei nun im Gegensatz zu (1.134) die $p_n \geq 0$ die Wahrscheinlichkeiten für das Auftreten der Zustände $|\mathbf{R}, n\rangle$ sind. Da die Eigenvektoren orthogonal aufeinander stehen, spezifizieren sie einander ausschließende Ereignisse, und wir können das *Maß für die fehlende Information* durch (1.75) definieren.

$$I[\mathbf{R}] = - \sum_{n=1}^{\infty} p_n \ln p_n := S[\mathbf{R}]. \quad (1.149)$$

Jetzt schreiben wir für dieses Shannonsche Unkenntnismaß S , denn wie wir bald sehen werden, stimmt es mit der *thermodynamischen Entropie* überein.

Sie läßt sich unabhängig von der Eigenbasis von \mathbf{R} offensichtlich zu

$$S[\mathbf{R}] = - \langle \ln R \rangle = - \text{Tr}(\mathbf{R} \ln \mathbf{R}) \quad (1.150)$$

bestimmen. Dies stimmt mit der Definition der Entropie, die zuerst John von Neumann formuliert hat, überein. Nun können wir das *Jaynessche Prinzip des geringsten Vorurteils* für die Quantenstatistik formulieren: Der statistische Operator des Systems wird so bestimmt, daß die *von Neumann-Entropie* (1.150) unter den gegebenen Informationen *maximal* wird.

Seien nun die *Erwartungswerte* von Observablen O_k mit $k \in \{1, \dots, n\}$ als *einzige Information* über das System gegeben. Dabei spielt es keine Rolle, ob die Observablen O_k miteinander verträglich sind oder nicht, denn nicht miteinander verträgliche Observablen können ja nur nicht simultan *definierte scharfe Werte* haben, es kann aber durchaus sein, daß die Erwartungswerte der Observablen für ein Ensemble von Systemen simultan bekannt sind.

Wir müssen also den statistischen Operator \mathbf{R} von allen möglichen statistischen Operatoren finden, der die von Neumann-Entropie (1.150) unter den Nebenbedingungen

$$\mathcal{O}_k = \langle O_k \rangle = \text{Tr}(\mathbf{R} O_k), \quad \text{Tr} \mathbf{R} = 1 \quad (1.151)$$

maximal macht. Dabei haben wir die selbstverständliche Normierungsbedingung mit unter die Nebenbedingungen aufgenommen.

Dieses Problem kann nun aber wie jedes Extremwertproblem mit Nebenbedingungen gelöst werden. Statt die von Neumann-Entropie unter den Nebenbedingungen zu maximieren, suchen wir einfach den stationären Punkt des Funktionals

$$S'[\mathbf{R}, \Omega, \lambda] = - \text{Tr} \left[\mathbf{R} \left(\ln \mathbf{R} + \sum_{k=1}^n \lambda_k O_k + \Omega - 1 \right) \right]. \quad (1.152)$$

Dabei sind die λ_k und $\Omega - 1$ sogenannte *Lagrangeparameter* zur Erfüllung der Nebenbedingungen (1.151). Wir brauchen nun nur das Funktional (1.152) ohne Nebenbedingungen zu maximieren und können dann die Lagrangemeter so bestimmen, daß die Nebenbedingungen erfüllt sind.

Das liefert

$$\delta S'[\mathbf{R}, \Omega, \lambda] = -\text{Tr} \left[\delta \mathbf{R} \left(\ln \mathbf{R} + \sum_{k=1}^n \lambda_k \mathbf{O}_k + \Omega \right) \right]. \quad (1.153)$$

Da \mathbf{R} beliebig variiert werden darf, muß die innere Klammer für sich verschwinden, so daß sogleich

$$\mathbf{R} = \exp \left(-\Omega \mathbf{1} - \sum_{k=1}^n \lambda_k \mathbf{O}_k \right) \quad (1.154)$$

resultiert. Wir werden weiter unten zeigen, daß dies wirklich der eindeutig bestimmte statistische Operator ist, der das Jaynessche Prinzip erfüllt, vorausgesetzt die Bedingungen (1.151) sind überhaupt miteinander verträglich, d.h. sie widersprechen einander nicht, und es existieren *reelle Lagrangeparameter* derart, daß (1.154) die Nebenbedingungen (1.151) erfüllt. Die Lagrangeparameter müssen nämlich reell sein, damit \mathbf{R} selbstadjungiert ist. Dann ist er aber auch zwangsläufig positiv definit. Wir stellen weiter fest, daß die (verallgemeinerten) Eigenvektoren von (1.154) die (verallgemeinerten) Eigenvektoren des Operators

$$\mathbf{Q} = \sum_{k=1}^n \lambda_k \mathbf{O}_k \quad (1.155)$$

sind.

Besonders wichtig für das folgende wird der Lagrangeparameter Ω sein, welcher sich aus der Bedingung $\text{Tr} \mathbf{R} = 1$ zu

$$Z = \exp \Omega = \text{Tr} \left[\exp \left(- \sum_{k=1}^n \lambda_k \mathbf{O}_k \right) \right] \quad (1.156)$$

bestimmt. Man nennt Z die *Zustandssumme*. Wir werden weiter unten zeigen, daß *auch bei nicht miteinander kommutierenden Operatoren* \mathbf{O}_k der Erwartungswert durch die „naiverweise“ geltende Formel

$$\langle O_k \rangle := \mathcal{O}_k = \text{Tr}[\mathbf{O}_k \mathbf{R}] = -\frac{1}{Z} \frac{\partial Z}{\partial \lambda_k} = -\frac{\partial \Omega}{\partial \lambda_k} \quad (1.157)$$

berechnet werden kann.

1.13 Beispiel: Ort und Impuls

Als ein weiteres Beispiel wollen wir nach dem Jaynesschen Prinzip den statistischen Operator bestimmen, der zur Kenntnis des mittleren Ortes und des mittleren Impulses eines Teilchens und deren Standardabweichungen gehört. Die Operatoren \mathbf{O}_k in (1.151) sind in diesem Falle also durch

$$\mathbf{O}_1 = \mathbf{x}, \quad \mathbf{O}_2 = \mathbf{p}, \quad \mathbf{O}_3 = \mathbf{x}^2 - \mathcal{X}^2, \quad \mathbf{O}_4 = \mathbf{p}^2 - \mathcal{P}^2 \quad \text{mit } \mathcal{X} = \langle x \rangle, \quad \mathcal{P} = \langle p \rangle \quad (1.158)$$

gegeben. Dies können wir unmittelbar in (1.154) einsetzen. Ohne Beschränkung der Allgemeinheit können wir den statistischen Operator maximaler Unkenntnis bei gegebener Information (1.158) wie folgt parametrisieren:

$$\mathbf{R} = \exp(-\Omega) \exp \left[-\frac{\lambda_1}{2} (\mathbf{x} - x_0)^2 - \frac{\lambda_2}{2} (\mathbf{p} - p_0)^2 \right]. \quad (1.159)$$

Dabei ersetzen x_0, p_0, λ_1 und λ_2 die Lagrangeparameter in (1.154).

Wir können dieses Problem analytisch lösen, denn die Operatoren

$$\boldsymbol{\xi} = \mathbf{x} - x_0, \quad \boldsymbol{\pi} = \mathbf{p} - p_0 \quad (1.160)$$

erfüllen die Kommutatorrelationen der Heisenbergalgebra

$$\frac{1}{i} [\boldsymbol{\xi}, \boldsymbol{\pi}] = 1. \quad (1.161)$$

Der statistische Operator schreibt sich in diesen Koordinaten wie folgt:

$$\mathbf{R} = \exp(-\Omega) \exp\left(-\frac{\lambda_1}{2} \boldsymbol{\xi}^2 - \frac{\lambda_2}{2} \boldsymbol{\pi}^2\right). \quad (1.162)$$

Das Argument der Exponentialfunktion sieht formal aus wie der negative Hamiltonoperator eines harmonischen Oszillators, und die Lösung des dazugehörigen Eigenwertproblems kann in jedem elementaren Lehrbuch der Quantentheorie gefunden werden:

Dazu definieren wir den Vernichtungsoperator für ein Oszillatorquant:

$$\mathbf{a} = (\lambda_1 \lambda_2)^{1/4} \left(\frac{1}{\sqrt{2\lambda_2}} \boldsymbol{\xi} + i \frac{1}{\sqrt{2\lambda_1}} \boldsymbol{\pi} \right). \quad (1.163)$$

Die Heisenbergalgebra (1.161) liefert die Kommutatorrelation

$$[\mathbf{a}^\dagger, \mathbf{a}] = 1. \quad (1.164)$$

Mit Hilfe dieser Relation finden wir das Eigensystem des Hamiltonoperators als das Eigensystem des *Oszillatorquantenzahloperators*:

$$\mathbf{n} = \mathbf{a}^\dagger \mathbf{a}, \quad \mathbf{n} |n\rangle = n |n\rangle \quad \text{mit } n \in \mathbb{N} := \{0, 1, 2, \dots\}. \quad (1.165)$$

Diese Eigenvektoren können aus dem „Vakuumzustand“ $|0\rangle$ durch sukzessive Anwendung des Erzeugungsoperators gemäß

$$\mathbf{a}^\dagger |n\rangle = \sqrt{n+1} |n+1\rangle \quad (1.166)$$

gebildet werden. Dabei haben wir eine für die Physik unerhebliche Phase willkürlich so gewählt, daß der Koeffizient in (1.166) reell ist. Die Vertauschungsregel (1.166) ergibt wegen $\mathbf{a}^\dagger \mathbf{a} |n\rangle = 0$

$$\mathbf{a} |0\rangle = 0, \quad \mathbf{a} |n\rangle = \sqrt{n} |n-1\rangle \quad \text{für } n \in \mathbb{N}_{>0}. \quad (1.167)$$

Der statistische Operator kann dann in der Form

$$\mathbf{R} = \exp(-\Omega) \exp\left[-\sqrt{\lambda_1 \lambda_2} \left(\mathbf{n} + \frac{1}{2}\right)\right] \quad (1.168)$$

Die Zustandssumme kann leicht mit Hilfe der vollständigen Orthonormalbasis $\{|n\rangle\}_{n \in \mathbb{N}_{\geq 0}}$ berechnet werden:

$$\begin{aligned} Z = \exp(\Omega) &= \exp\left(-\frac{\sqrt{\lambda_1 \lambda_2}}{2}\right) \sum_{n=0}^{\infty} \exp(-n \sqrt{\lambda_1 \lambda_2}) = \frac{\exp\left(-\frac{\sqrt{\lambda_1 \lambda_2}}{2}\right)}{1 - \exp(-\sqrt{\lambda_1 \lambda_2})} \\ &= \frac{1}{2 \sinh\left(\frac{\sqrt{\lambda_1 \lambda_2}}{2}\right)}. \end{aligned} \quad (1.169)$$

Es ist klar, daß die Zustandssumme nur existiert, wenn $\lambda_1, \lambda_2 > 0$ sind. Die Existenz der Zustandssumme ist aber notwendig, damit $\text{Tr } \mathbf{R} = 1$ erfüllt werden kann, so daß diese Einschränkung an die Lagrangeparameter hier Voraussetzung für die Existenz des statistischen Operators minimaler Unschärfe ist.

Die Zustandssumme hängt nicht explizit von x_0 und p_0 ab. Man rechnet jedoch sofort direkt nach, daß

$$\langle \xi \rangle = \text{Tr } \boldsymbol{\xi} \mathbf{R} = 0, \quad \langle \pi \rangle = \text{Tr } \boldsymbol{\pi} \mathbf{R} = 0 \quad (1.170)$$

und folglich

$$\langle x \rangle = \text{Tr } \mathbf{x} \mathbf{R} = x_0, \quad \langle p \rangle = \text{Tr } \mathbf{p} \mathbf{R} = p_0. \quad (1.171)$$

Die Lagrangeparameter x_0 und p_0 sind also direkt die Erwartungswerte $\langle x \rangle$ bzw. $\langle p \rangle$.

Die Standardabweichungen berechnen sich demnach nach (1.157) als Ableitungen von Z

$$\begin{aligned} \langle \xi^2 \rangle = \Delta x^2 &= -\frac{2}{Z} \frac{\partial Z}{\partial \lambda_1} = \sqrt{\frac{\lambda_1}{4\lambda_2}} \coth \left(\frac{\sqrt{\lambda_1 \lambda_2}}{2} \right), \\ \langle \pi^2 \rangle = \Delta p^2 &= -\frac{2}{Z} \frac{\partial Z}{\partial \lambda_2} = \sqrt{\frac{\lambda_2}{4\lambda_1}} \coth \left(\frac{\sqrt{\lambda_1 \lambda_2}}{2} \right). \end{aligned} \quad (1.172)$$

Es ist also stets

$$\Delta x \Delta p = \frac{1}{2} \coth^2 \left(\frac{\sqrt{\lambda_1 \lambda_2}}{2} \right) \geq 1/2, \quad (1.173)$$

was besagt, daß es keinen statistischen Operator gibt, der das Jaynesschen Prinzip erfüllt und die *Heisenbergsche Unschärferelation* verletzt. In der Tat folgt aus der Unschärferelation für reine Zustände, wie sie in [Hee98] bewiesen wurde sofort, daß selbstverständlich auch in einem Gemisch die Unschärferelation nicht verletzt werden kann.

1.14 Anhang: Beweis zweier wichtiger Theoreme

In diesem Anhang zum gegenwärtigen Kapitel wollen wir zwei wichtige Theoreme beweisen.

Theorem 2 (Eindeutigkeit des statistischen Operators). *Sei $\{\mathbf{O}_j\}_{j \in \{1, \dots, n\}}$ eine Menge von nicht notwendig kompatiblen Operatoren. Dann ist der statistische Operator*

$$\mathbf{R} = \exp(-\Omega - \sum_{j=1}^n \lambda_j \mathbf{O}_j) \quad (1.174)$$

der einzige, der die Nebenbedingungen

$$\langle \mathbf{O} \rangle_j = \mathcal{O}_j \quad (1.175)$$

erfüllt und die von Neumann-Entropie

$$S[\rho] = -\langle \ln \rho \rangle = -\text{Tr}(\boldsymbol{\rho} \ln \boldsymbol{\rho}) \quad (1.176)$$

maximiert, vorausgesetzt es existiert ein Satz von Lagrangeparametern λ_j mit $j \in \{1, \dots, n\}$, so daß die Nebenbedingungen (1.175) erfüllt sind.

Beweis. Sei ρ ein beliebiger statistischer Operator, welcher die Nebenbedingungen (1.175) erfüllt. wir wollen zeigen, daß

$$S[\rho] - S[\mathbf{R}] \leq 0. \quad (1.177)$$

Dazu zeigen wir zunächst

$$\forall x \in \mathbb{R}_{>0} : \ln x \geq x - 1 \text{ und } x - 1 = \ln x \Leftrightarrow x = 1. \quad (1.178)$$

Dazu müssen wir lediglich die Funktion $f : \mathbb{R}_{>0} \rightarrow \mathbb{R}$ mit $f(x) = x - 1 - \ln x$ diskutieren: Die Ableitung dieser Funktion ist $f'(x) = 1 - 1/x$. Diese ist strikt monoton fallend und hat als einzige Nullstelle $x = 1$. Also hat f ein globales Minimum bei $x = 1$. Wegen $f(1) = 0$ folgt die Behauptung (1.178).

Mit Hilfe der Eigenbasis von \mathbf{R} folgt

$$\begin{aligned} \text{Tr}[\mathbf{R} \ln \mathbf{R}] &= - \sum_n R(n) \left(\sum_j \lambda_j o_{nn}^{(j)} + \Omega \right) = -\Omega - \sum_j \lambda_j \mathcal{O}_j \\ &= - \text{Tr} \left[\rho \left(-\Omega - \sum_j \lambda_j \mathbf{O}_j \right) \right] = \text{Tr}(\rho \ln \mathbf{R}) \end{aligned} \quad (1.179)$$

Dabei haben wir die Tatsache benutzt, daß nach Voraussetzung ρ und \mathbf{R} beide die Nebenbedingungen (1.175) erfüllen, wobei wir bei letzterem annehmen müssen, daß ein entsprechender Satz Lagrangeparameter λ_j existiert.

Jetzt benutzen wir die Eigenbasis $\{|\alpha\rangle\}_{\alpha \in \mathbb{N}}$ des statistischen Operators ρ :

$$S[\rho] - S[\mathbf{R}] = \sum_{\alpha} \rho_{\alpha} \ln \left(\frac{R_{\alpha\alpha}}{\rho_{\alpha}} \right) \leq \sum_{\alpha} (\rho_{\alpha} - R_{\alpha\alpha}) = 0 \quad (1.180)$$

Dabei wurde die Ungleichung (1.178) und $\text{Tr} \rho = \text{Tr} \mathbf{R} = 1$ benutzt. Benutzen wir nun die Äquivalenzrelation in (1.178) ergibt sich, daß $S[\rho] = S[\mathbf{R}]$ zu $\rho = \mathbf{R}$ äquivalent ist und daß alle anderen ρ eine kleinere von Neumann-Entropie als \mathbf{R} besitzen. \square

Theorem 3 (Ableitung von Exponentialfunktionen). *Sei $\mathbf{Y} : D \rightarrow \text{Hom}(\mathcal{H})$, wobei D eine offene Teilmenge von \mathbb{R} ist. Dann definieren wir*

$$\partial_a \mathbf{Y}(a) = \lim_{\Delta a \rightarrow 0} \frac{\mathbf{Y}(a + \Delta a) - \mathbf{Y}(a)}{\Delta a}, \quad (1.181)$$

falls der Limes existiert. Dann heißt \mathbf{Y} in a differenzierbar und (1.181) die Ableitung.

Wenn dies der Fall ist, dann ist auch der Operator $\exp[\mathbf{Y}(a)]$ differenzierbar, und es gilt

$$\partial_a \exp[\mathbf{Y}(a)] = \exp[\mathbf{Y}(a)] \int_0^1 d\tau \exp[-\tau \mathbf{Y}(a)] \partial_a \mathbf{Y}(a) \exp[\tau \mathbf{Y}(a)]. \quad (1.182)$$

Beweis. Sei $\mathbf{X} : D \rightarrow \text{Hom}(\mathcal{H})$ (wobei $D \subseteq \mathbb{R}$ offen sei) und $\mathbf{X}(t) = \mathbf{X}_1(t) + \mathbf{X}_2(t)$. Weiter definieren wir die Operatoren \mathbf{U} und \mathbf{U}_j mit ($j \in \{1, 2\}$) durch die Anfangswertprobleme

$$\begin{aligned} \partial_t \mathbf{U}(t, t_0) &= \mathbf{X}(t) \mathbf{U}(t, t_0), \quad \mathbf{U}(t_0, t_0) = 1; \\ \partial_t \mathbf{U}_j(t, t_0) &= \mathbf{X}_j(t) \mathbf{U}_j(t, t_0), \quad \mathbf{U}_j(t_0, t_0) = 1 \end{aligned} \quad (1.183)$$

Diese können genauso gelöst werden wie in Abschnitt 2.2 für die Zeitentwicklungsoperatoren gezeigt. Ein weiterer Operator \mathbf{V} sei wie folgt festgelegt:

$$\mathbf{V}(t, t_0) = \mathbf{U}_1(t, t_0) + \int_{t_0}^t dt' \mathbf{U}(t, t') \mathbf{X}_2(t') \mathbf{U}_1(t, t_0). \quad (1.184)$$

Leiten wir dies nach t ab und verwenden (1.183) sowie $\mathbf{U}(t, t) = 1$ für $t \in D$, so folgt

$$\partial_t \mathbf{V}(t, t_0) = \mathbf{X}(t) \mathbf{V}(t, t_0). \quad (1.185)$$

Aus (1.183) ersehen wir sofort, daß $\mathbf{V}(t_0, t_0) = 1$. Also ist $\mathbf{V} \equiv \mathbf{U}$, und somit gilt die Identität

$$\mathbf{U}(t, t_0) = \mathbf{U}_1(t, t_0) + \int_{t_0}^t dt' \mathbf{U}(t, t') \mathbf{X}_2(t') \mathbf{U}_1(t, t_0). \quad (1.186)$$

Angewandt auf den Fall, daß \mathbf{X}_1 und \mathbf{X}_2 zeitunabhängig sind, ergibt sich

$$\exp[(t - t_0)\mathbf{X}] = \exp[(t - t_0)\mathbf{X}_1] + \int_{t_0}^t dt' \exp[(t - t')\mathbf{X}] \mathbf{X}_2 \exp[(t' - t_0)\mathbf{X}_1]. \quad (1.187)$$

Für $t_0 = 0$ und $t = 1$ ist also

$$\exp[\mathbf{X}_1 + \mathbf{X}_2] = \exp[\mathbf{X}_1] + \exp(\mathbf{X}) \int_0^1 d\tau \exp[\tau(\mathbf{X}_1 + \mathbf{X}_2)] \mathbf{X}_2 \exp(\tau \mathbf{X}_1). \quad (1.188)$$

Wählen wir schließlich $\mathbf{X}_1 = \mathbf{Y}(a)$ und $\mathbf{X}_2 = \Delta a \partial_a \mathbf{Y}(a)$ ergibt sich daraus für $\Delta a \rightarrow 0$ sogleich die Behauptung. \square

Theorem 4 (Berechnung von Erwartungswerten). *Die Zustandssumme für einen statistischen Operator, der unter den Nebenbedingungen (1.175) das Jaynessche Prinzip erfüllt, ist durch*

$$Z = \exp \Omega = \text{Tr} \exp \left(- \sum_{j=1}^n \lambda_j \mathbf{O}_j \right) \quad (1.189)$$

definiert.

Die Erwartungswerte der Observablen O_j lassen sich dann als Ableitung von Z nach den Lagrangeparametern berechnen:

$$\langle O_j \rangle = - \frac{1}{Z} \frac{\partial Z}{\partial \lambda_j} = - \frac{\partial \Omega}{\partial \lambda_j}. \quad (1.190)$$

Beweis. Unter Verwendung von

$$\mathbf{Y} = - \sum_{j=1}^n \lambda_j \mathbf{O}_j \text{ and } a = \lambda_k \text{ mit } k = 1, \dots, n \quad (1.191)$$

in (1.182) ergibt sich

$$\frac{\partial}{\partial \lambda_k} \exp \left[- \sum_{j=1}^n \lambda_j \mathbf{O}_j \right] = - \exp \left[- \sum_{j=1}^n \lambda_j \mathbf{O}_j \right] \int_0^1 d\tau \exp \left[- \tau \sum_{j=1}^n \lambda_j \mathbf{O}_j \right] \mathbf{O}_k \exp \left[\tau \sum_{j=1}^n \lambda_j \mathbf{O}_j \right]. \quad (1.192)$$

Existiert nun die Spur über das Produkt zweier Operatoren \mathbf{X} und \mathbf{Y} , so gilt

$$\mathrm{Tr}(\mathbf{XY}) = \mathrm{Tr}(\mathbf{YX}). \quad (1.193)$$

Dies ist unmittelbar einsichtig, wenn man die Spur mit Hilfe eines beliebigen VONS $\{|n\rangle\}_{n \in \mathbb{N}_{>0}}$ bildet:

$$\mathrm{Tr}(\mathbf{XY}) = \sum_{n_1, n_2=1}^{\infty} \langle n_1 | \mathbf{X} | n_2 \rangle \langle n_2 | \mathbf{Y} | n_1 \rangle = \sum_{n_1, n_2=1}^{\infty} \langle n_2 | \mathbf{Y} | n_1 \rangle \langle n_1 | \mathbf{X} | n_2 \rangle = \mathrm{Tr}(\mathbf{YX}). \quad (1.194)$$

Nehmen wir also die Spur von (1.192), können wir den ganz rechts stehenden Exponentialoperator nach links tauschen. Wegen

$$[X, Y] = 0 \Rightarrow \exp(\mathbf{X} + \mathbf{Y}) = \exp(\mathbf{X}) \exp(\mathbf{Y}) \quad (1.195)$$

was ganz analog wie bei der Exponentialfunktion für komplexe Zahlen aus der definierenden Potenzreihe bewiesen werden kann, wird das Integral trivial, und wir finden

$$\mathrm{Tr} \left[\frac{\partial}{\partial \lambda_k} \exp \left(- \sum_{j=1}^n \lambda_j \mathbf{O}_j \right) \right] = - \mathrm{Tr} \left[\mathbf{O}_k \exp \left(\sum_{j=1}^n \lambda_j \mathbf{O}_j \right) \right]. \quad (1.196)$$

Jetzt müssen wir lediglich noch voraussetzen, daß die Spurbildung mit der Ableitung vertauscht werden kann. Dies erfordert, daß die die Spur definierende Formel

$$\partial_a \mathrm{Tr} \mathbf{X}(a) = \partial_a \sum_n \langle n | \mathbf{X}(a) | n \rangle \quad (1.197)$$

gliedweise abgeleitet werden darf, d.h. daß die Reihe in einer Umgebung von a gleichmäßig konvergiert und an wenigstens einer Stelle die Reihe, die durch gliedweise Differentiation der ursprünglichen Reihe hervorgeht, konvergiert. Wir gehen auf die genauen Forderungen an den Operator \mathbf{X} nicht ein. \square

Kapitel 2

Thermodynamisches Gleichgewicht

In diesem Abschnitt wollen wir zunächst allgemeine Aspekte des thermodynamischen Gleichgewichts betrachten und dann ideale Gase behandeln, wobei wir besonderes Augenmerk auf den thermodynamischen Limes haben werden.

2.1 Adiabatische Änderungen im thermodynamischen Gleichgewicht

Wir zeigen zuerst das *Adiabatentheorem*, welches uns gestatten wird, aus der Gleichgewichtsquantenstatistik die Hauptsätze der phänomenologischen Thermodynamik herzuleiten:

Theorem 5. *Der Hamiltonoperator $\mathbf{H}(\chi)$ eines Quantensystems sei von einem externen reellen Parameter χ abhängig¹. Wir nehmen weiter an, das Spektrum des Hamiltonoperators sei nicht entartet.*

Nun möge sich der Parameter mit der Zeit wie folgt ändern:

$$\chi(t) = \chi_1 + \frac{t}{\tau}(\chi_2 - \chi_1). \quad (2.1)$$

Im Grenzfall „unendlich langsamer Änderung“, also für $\tau \rightarrow \infty$ geht dann eine erhaltene Observable $O(\chi_1)$ durch die Zeitentwicklung in eine erhaltene Observable χ_2 über.

Beweis. Wir rechnen im folgenden im Heisenbergbild, so daß von eventueller expliziter Zeitabhängigkeit abgesehen nur die Observablen-Operatoren zeitabhängig sind, die Zustände sind zeitunabhängig. Die Eigenzustände des Hamiltonoperators $|\alpha, \chi\rangle$ sind allerdings zeitabhängig durch die vermöge (2.1) eingebrachte explizite Zeitabhängigkeit des externen Parameters χ . Die Eigenwertgleichung für den Hamiltonoperator schreiben wir wie folgt:

$$\mathbf{H}(\chi) |\alpha, \chi\rangle = E_\alpha(\chi) |\alpha, \chi\rangle. \quad (2.2)$$

Die Bewegungsgleichung im Heisenbergbild ist

$$-i \frac{d}{dt} |\alpha, \chi\rangle = \mathbf{H}(\chi) |\alpha, \chi\rangle - i \frac{\chi_2 - \chi_1}{\tau} \frac{\partial}{\partial \chi} |\alpha, \chi\rangle, \quad (2.3)$$

¹Das kann ein äußeres Feld oder Parameter wie das Volumen, in das ein Gas eingeschlossen wird, sein.

woraus sich die Schrödingergleichung in der Energiedarstellung

$$i \frac{d}{dt} \langle \alpha, \chi | \psi \rangle = \langle \alpha, \chi | \mathbf{H}(\chi) | \psi \rangle + i \frac{\chi_2 - \chi_1}{\tau} \frac{\partial}{\partial \chi} \langle \alpha, \chi | \psi \rangle. \quad (2.4)$$

ergibt. Unter Verwendung des Operators

$$\mathbf{G}(\chi) = i \sum_{\alpha} \left(\frac{\partial}{\partial \chi} | \alpha, \chi \rangle \right) \langle \alpha, \chi | = -i \sum_{\alpha} | \alpha, \chi \rangle \frac{\partial}{\partial \chi} \langle \alpha, \chi | \quad (2.5)$$

finden wir

$$i \frac{d}{dt} \langle \alpha, \chi | \psi \rangle = E_{\alpha}(\chi) \langle \alpha, \chi | \psi \rangle - \frac{\chi_2 - \chi_1}{\tau} \sum_{\alpha'} G_{\alpha\alpha'}(\chi) \langle \alpha', \chi | \psi \rangle \quad (2.6)$$

mit dem Matricelement

$$G_{\alpha\alpha'}(\chi) = \langle \alpha, \chi | \mathbf{G}(\chi) | \alpha', \chi \rangle. \quad (2.7)$$

Definieren wir jetzt

$$\langle \alpha, \chi | \psi \rangle = \exp \left[-i \int_0^t dt' E_{\alpha}(\chi(t')) \right] \phi_{\alpha}(t) \quad (2.8)$$

und substituieren dies in (2.6) unter Verwendung von

$$\int_0^t dt' E_{\alpha}(\chi(t')) = \frac{\tau}{\chi_2 - \chi_1} \int_{\chi_1}^{\chi(t)} d\chi E_{\alpha}(\chi) \quad (2.9)$$

ergibt nach einigen einfachen Umformungen

$$i \frac{d\phi_{\alpha}(t)}{dt} = -\frac{\chi_2 - \chi_1}{\tau} \sum_{\alpha'} G_{\alpha\alpha'}(\chi) \exp \left[\frac{i\tau}{\chi_2 - \chi_1} \int_{\chi_1}^{\chi(t)} d\chi [E_{\alpha}(\chi) - E_{\alpha'}(\chi)] \right]. \quad (2.10)$$

Nun definieren wir die Phase der Energieeigenvektoren wie folgt um:

$$| \alpha', \chi \rangle = \exp[i\sigma_{\alpha}(\chi)] | \alpha, \chi \rangle. \quad (2.11)$$

Dann wird \mathbf{G} gemäß (2.5) zu

$$\mathbf{G}'(\chi) = \mathbf{G}\chi - \sum_{\alpha} \langle \alpha, \chi | \frac{d\sigma_{\alpha}(\chi)}{d\chi} | \alpha, \chi \rangle, \quad (2.12)$$

also

$$G'_{\alpha\alpha}(\chi) = \langle \alpha', \chi | \mathbf{G}(\chi) | \alpha', \chi \rangle - \frac{d\sigma_{\alpha}(\chi)}{d\chi}. \quad (2.13)$$

Das bedeutet, daß wir durch geeignete Wahl der Phasen σ_{α} die Diagonalelemente von \mathbf{G} zum Verschwinden bringen können. Wir nehmen o.B.d.A. im folgenden an, dies sei bereits für \mathbf{G} der Fall.

Integrieren wir (2.10) bzgl. der Zeit, erhalten wir

$$i[\phi_{\alpha}(\tau) - \phi_{\alpha}(0)] = -\frac{\chi_2 - \chi_1}{\tau} \sum_{\alpha'} \int_0^{\tau} dt G_{\alpha\alpha'}(\chi) \phi_{\alpha'} \exp \left[\frac{i\tau}{\chi_2 - \chi_1} \int_{\chi_1}^{\chi(t)} d\chi [E_{\alpha}(\chi) - E_{\alpha'}(\chi)] \right]. \quad (2.14)$$

Wegen unserer Phasenwahl ist $G_{\alpha\alpha} = 0$ für alle α und folglich tauchen in (2.10) nur nichtverschwindende Energiedifferenzen auf, weil wir angenommen haben, der Hamiltonoperator besitze

ein nicht ausgeartetes Spektrum. Das Integral im Argument der Exponentialfunktion oszilliert also für $\tau \rightarrow \infty$ immer mehr, während alle übrigen Funktionen sich nur wenig mit der Zeit ändern. Die Integration über t mittelt also die Summe komplett weg, so daß wir die folgende Asymptotik haben:

$$\phi_\alpha(\tau) \underset{\tau \rightarrow \infty}{\cong} \phi_\alpha(0). \quad (2.15)$$

Sei nun $\mathbf{A}(\chi)$ eine Konstante der Bewegung bzgl. des parameterabhängigen Hamiltonoperators $\mathbf{H}(\chi)$. Dann ist dieser Operator in der Eigenbasis des Hamiltonoperators diagonal:

$$\mathbf{A} = \sum_{\alpha} a_{\alpha\alpha}(\chi) |\alpha, \chi\rangle \langle \alpha, \chi|, \quad (2.16)$$

und aus (2.15) folgt, daß die adiabatische Änderung des Parameters χ mit der Zeit „deformiert“ $\mathbf{A}(\chi_1)$, welches eine Konstante der Bewegung bzgl. $\mathbf{H}(\chi_1)$ ist, zu einer Konstante der Bewegung $\mathbf{A}(\chi_2)$ bzgl. $\mathbf{H}(\chi_2)$. \square

2.2 Phänomenologische Thermodynamik

In diesem Abschnitt zeigen wir kurz, daß die phänomenologische Thermodynamik, soweit sie die Gleichgewichtsthermodynamik betrifft, mikroskopisch aus der Quantenstatistik und dem eben bewiesenen Adiabatentheorem hergeleitet werden kann.

Wir betrachten also im folgenden Gleichgewichtszustände, also den Fall, daß der statistische Operator \mathbf{R} des Systems nicht explizit zeitabhängig ist.

Gemäß (1.146) bedeutet dies, daß der statistische Operator mit dem Hamiltonoperator vertauscht,

$$\frac{1}{i} [\mathbf{R}, \mathbf{H}] = 0, \quad (2.17)$$

und also für den Operator, welcher das Jaynessche Prinzip erfüllt gemäß (1.154), daß dieser nur von den Konstanten der Bewegung, also *Erhaltungsgrößen* abhängen darf. Statistische Operatoren für thermodynamisches Gleichgewicht sind also solche, für den als gegebene Information einzig und allein Erwartungswerte von Erhaltungsgrößen des Systems vorgegeben sind. Eine andere Möglichkeit sind sog. *mikrokanonische statistische Operatoren*, für die die Energie oder die Energie und andere Erhaltungsgrößen des Systems mit Sicherheit in bestimmten Intervallen befindet. Darauf gehen wir jedoch hier nicht ein. Sei also ein Satz linear unabhängiger erhaltener Größen $\{O_i\}_{i \in \{1, \dots, n\}}$ gegeben. Die gegebene Information seien die Erwartungswerte dieser Operatoren, so daß gemäß (1.154) der statistische Operator die Form

$$\mathbf{R} = \exp \left(-\Omega - \sum_{i=1}^n \lambda_i O_i \right) \quad (2.18)$$

besitzt.

Die wichtigsten Fälle sind der *kanonische statistische Operator*, wo nur die mittlere Energie des Systems vorgegeben ist, d.h.

$$\mathbf{R} = \exp(-\Omega - \beta \mathbf{H}), \quad (2.19)$$

und der *großkanonische statistische Operator*, wo die mittlere Energie und die mittlere Teilchenzahl gegeben sind. Hierbei ist allerdings zu beachten, daß dies nur im Rahmen der nichtrelativistischen Quantentheorie sinnvoll ist, da im relativistischen Fall die Teilchenzahl i.a. nicht

erhalten ist. Dann tritt an die Stelle der Teilchenzahl eine erhaltene Noetherladung, z.B. die Nettoleptonenzahl (die Differenz aus Leptonen- und Antileptonen):

$$\mathbf{R} = \exp(-\Omega - \beta\mathbf{H} - \alpha\mathbf{N}). \quad (2.20)$$

Wegen Theorem 4 können wir die Erwartungswerte der O_i durch Bildung der Ableitungen von Ω nach den entsprechenden Lagrangeparametern λ_i gewinnen:

$$\langle O_i \rangle = -\frac{\partial \Omega}{\partial \lambda_i}. \quad (2.21)$$

Aufgrund des Adiabaten-Theorems im vorigen Abschnitt, verbleibt der statistische Operator ein Gleichgewichtsoperator, wenn wir eine *adiabatische* Änderung äußerer Parameter (wie z.B. das Volumen, in dem ein Gas eingesperrt ist) vornehmen. Wegen des Ehrenfest'schen Theorems bleibt während einer solchen adiabatischen Zustandsänderung auch die Entropie $S = -\langle \ln \mathbf{R} \rangle$ unverändert.

Auf der anderen Seite ist

$$S = -\langle \ln \mathbf{R} \rangle = \Omega + \sum_j \lambda_j \mathcal{O}_j \quad \text{mit} \quad \mathcal{O}_j = \langle \mathbf{O}_j \rangle. \quad (2.22)$$

Dies zeigt, daß Ω eine *Legendretransformierte* der Entropie, also eine sog. *Massiefunktion* ist. Eine adiabatische Änderung eines äußeren Parameters bedeutet also, daß das System während dieser Änderung ständig in einem Gleichgewichtszustand verbleibt, und für ein abgeschlossenes System

$$dS = \left(\frac{\partial S}{\partial \chi} \right)_{\mathcal{O}_j = \text{const.}} d\chi + \sum_j \lambda_j d\mathcal{O}_j = 0. \quad (2.23)$$

Dies nennen wir *thermodynamisches Adiabatentheorem*.

Es folgt daraus und wegen $\partial \Omega / \partial \lambda_j = -\mathcal{O}_j$

$$dS = \sum_{\chi} \left(\frac{\partial \Omega}{\partial \chi} \right)_{\lambda_j = \text{const.}} d\chi + \sum_j \lambda_j d\mathcal{O}_j = 0, \quad (2.24)$$

denn

$$\left(\frac{\partial S}{\partial \chi} \right)_{\mathcal{O}_j = \text{const.}} = \left(\frac{\partial \Omega}{\partial \chi} \right)_{\lambda_j = \text{const.}}. \quad (2.25)$$

Im allgemeinen können wir aber nicht die Durchschnittswerte aller erhaltenen Größen eines Systems kennen, weil dies für ein makroskopisches System eine viel zu große Zahl von Größen beinhaltet. Gewöhnlich betrachten wir daher in der Thermodynamik *offene Systeme*, d.h. wir haben ein großes makroskopisches System, und wir betrachten ein kleines Untersystem von ebenfalls makroskopischen Ausmaßen, von dem wir nur die mittlere Energie und die mittlere Teilchenzahl kennen.

Wir nehmen weiter an, daß sich das System als ganzes im *thermischen und chemischen* Gleichgewicht befindet. Der statistische Operator des Gesamtsystems ist also in diesem Falle

$$\mathbf{R} = \exp(-\Omega - \beta\mathbf{H} - \alpha\mathbf{N}), \quad (2.26)$$

wo \mathbf{N} der Operator der totalen Teilchenzahl ist.

Jetzt betrachten wir eine adiabatische Änderung an dem kleinen Untersystem und nehmen dabei an, daß das Gesamtsystem so groß ist, daß die Änderung der Entropie des großen Restsystems vernachlässigbar klein ist.

Aus (2.23) erhalten wir unter diesen Voraussetzungen

$$dS_1 = \beta d\mathcal{E}_1 + \alpha d\mathcal{N}_1 + \beta \sum_{\chi} d\chi \left(\frac{\partial \Omega}{\partial \chi} \right)_{\lambda_j = \text{const.}} = -dS_2 = 0. \quad (2.27)$$

Dabei beziehen sich hier und im folgenden alle Größen mit Index 1 auf das kleine Untersystem und solche mit 2 auf den Rest.

Betrachten wir nun speziell das Volumen V des kleinen Untersystems als einen äußeren Parameter, so erhalten wir mit (2.27):

$$d\mathcal{E}_1 = \frac{1}{\beta} dS_1 - \frac{1}{\beta} \left(\frac{\partial \Omega}{\partial V} \right)_{\alpha, \beta} dV - \frac{\alpha}{\beta} d\mathcal{N}_1. \quad (2.28)$$

Wir können damit die Lagrangeparameter und die Ableitung der Massiefunktion Ω mit den *phänomenologischen Zustandsgrößen Entropie S , Temperatur T , Druck p und chemisches Potential μ* des Untersystems identifizieren:

$$\begin{aligned} S &= -\text{Tr}(\mathbf{R} \ln \mathbf{R}) = \Omega + \beta \mathcal{E} + \alpha \mathcal{N}, & T &= \frac{1}{T}, \\ p &= \frac{1}{\beta} \left(\frac{\partial \Omega}{\partial V} \right)_{\alpha, \beta} = \frac{T}{V} \Omega, & \mu &= -\frac{\alpha}{\beta}, \\ \mathcal{E} &= - \left(\frac{\partial \Omega}{\partial \beta} \right)_{\alpha, V}, & \mathcal{N} &= - \left(\frac{\partial \Omega}{\partial \alpha} \right)_{\beta, V}. \end{aligned} \quad (2.29)$$

so daß (2.28) die Form des *ersten Hauptsatzes der Thermodynamik* annimmt:

$$d\mathcal{E} = TdS - pdV + \mu d\mathcal{N}. \quad (2.30)$$

Es besagt nichts anderes als die Energieerhaltung für das Untersystem: Die Änderung der mittleren Energie ist durch die Wärme TdS , die es aus seiner Umgebung bezieht, die mechanische Arbeit $-pdV$, die es an die Umgebung abgibt und schließlich die für chemische Reaktionen erforderliche chemische Arbeit $\mu d\mathcal{N}$.

Das Shannon-von Neumannsche Informationsmaß, welches das Jaynessche Prinzip des geringsten Vorurteils erfüllt, für ein offenes System, das in thermischem und chemischem Kontakt mit einem größeren System (oft als *Wärmebad* bezeichnet) steht, ist also identisch mit der *makroskopischen Entropie*, wie sie von Clausius und Planck im Rahmen der phänomenologischen Thermodynamik eingeführt wurde.

Die phänomenologische Thermodynamik kann somit auf die allgemeinen Prinzipien der Quantenstatistik und die Annahme des Vorliegens thermodynamischen Gleichgewichts sowie die Beschränkung der Betrachtung auf adiabatische Zustandsänderungen, also solche Änderungen des Systems, die so langsam vor sich gehen, daß man das System die ganze Zeit als im thermischen Gleichgewicht befindlich ansehen kann.

Es sei hier nur noch erwähnt, daß in der Quantenstatistik einige Probleme wie das Gibbsche Paradoxon der klassischen Statistik gar nicht erst auftreten.

Kapitel 3

Ideale Gase im großkanonischen Ensemble

In diesem Abschnitt wollen wir die verschiedenen idealen Gase aus Fermionen und Bosonen nichtrelativistisch und relativistisch zusammenstellen. Sie sind eines der wenigen Systeme, die sich geschlossen auswerten lassen.

Wir werden gleich die allgemeinen Methoden der Vielteilchentheorie, also die Quantenfeldtheorie benutzen, weil sie die gemeinsame Basis für die Behandlung der nichtrelativistischen wie der relativistischen Quantentheorie darstellt.

3.1 Nichtrelativistische ideale Gase

Wir werden in diesem Skript weitgehend den Operatorformalismus der Pfadintegralformulierung vorziehen. Dies ist umso mehr gerechtfertigt als eine kovariante Quantisierung für die relativistische Feldtheorie nicht mehr die großen Vorteile bringt, die sie im Vakuumfalle besitzt. Wir kommen darauf am gegebenen Platze noch zurück. Beginnen wir jedoch zunächst mit dem *idealen nichtrelativistischen Gas*.

Die Wirkung für die Einteilchenschrödingergleichung lautet bekanntlich

$$A[\psi^*, \psi] = \int d^4x \underbrace{\left[\frac{i}{2} \psi^* \overleftrightarrow{\partial}_t \psi - \frac{1}{2m} (\nabla \psi)^* (\nabla \psi) \right]}_{\mathcal{L}} \quad \text{mit } f \overleftrightarrow{\partial}_t g = f \partial_t g - (\partial_t f) g. \quad (3.1)$$

Hier und im folgenden benutzen wir die Abkürzung $x = (t, \vec{x})$ im Integral. Wie wir sehen werden, entfaltet diese Notation allerdings erst im relativistischen Falle ihre volle abkürzende Wirkung. Die Bewegungsgleichungen sind durch das *Hamiltonsche Prinzip der kleinsten Wirkung* gegeben. Demnach haben wir die Felder so zu bestimmen, daß die Wirkung minimal wird. Das Resultat sind die *Euler-Lagrangegleichungen*

$$\frac{\delta A}{\delta \psi^*} = \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\psi}^*} + \nabla \frac{\partial \mathcal{L}}{\partial (\nabla \psi^*)} - \frac{\partial \mathcal{L}}{\partial \psi^*} = 0. \quad (3.2)$$

Die Variation bzgl. ψ führt auf die dazu konjugiert komplexe Gleichung. Benutzen wir die Lagrangedichte aus (3.1) ist (3.2) gerade die *Schrödingergleichung für ein freies Teilchen*:

$$i \partial_t \psi = - \frac{\Delta \psi}{2m}. \quad (3.3)$$

Der Vorteil, die Schrödingergleichung wie eine klassische Feldgleichung aus dem Wirkungsprinzip herzuleiten, liegt vor allem darin, daß wir die Ausdrücke für die Erhaltungsgrößen wie Energie, Impuls, Drehimpuls und Teilchenzahl aus dem Noethertheorem (vgl. dazu wieder [Hee98]).

Wir quantisieren nun diese Feldgleichungen „kanonisch“. Dazu gehen wir als erstes zum Hamiltonformalismus über, indem wir die *kanonischen Feldimpulse* einführen:

$$\Pi = \frac{\partial \mathcal{L}}{\partial \dot{\psi}} = \frac{i}{2} \psi^*, \quad \Pi^* = \frac{\partial \mathcal{L}}{\partial \dot{\psi}^*} = -\frac{i}{2} \psi. \quad (3.4)$$

Die Hamiltondichte ist durch ihre Definition gegeben:

$$\mathcal{H} = \Pi \dot{\psi} + \Pi^* \dot{\psi}^* - \mathcal{L} = \frac{1}{2m} (\nabla \psi^*) (\nabla \psi) = -\frac{i}{2m} [(\nabla \Pi) (\nabla \psi) - (\nabla \Pi^*) (\nabla \psi^*)]. \quad (3.5)$$

Wir haben dabei Lagrange- und Hamiltondichte so gewählt, daß reelle Ausdrücke entstehen. Es ist leicht zu sehen, daß auch das erweiterte Hamiltonsche Prinzip, wonach die Wirkung

$$A[\psi, \Pi; \psi^*, \Pi^*] = \int d^4x [\Pi \dot{\psi} + \Pi^* \dot{\psi}^* - \mathcal{H}] \quad (3.6)$$

unter unabhängigen Variationen nach den Feld- und Feldimpulsfreiheitsgraden zu minimieren ist, wieder auf die Schrödingergleichung und ihre konjugiert komplexe Gleichung führt. Daß die vier Gleichungen in diesem Falle teilweise redundant sind, liegt darin begründet, daß die Schrödingergleichung nur linear in der Zeit ist und folglich die Lagrangedichte auch nur linear von der Zeitableitung der Felder abhängt.

Die Lagrangedichte in (3.5) ist nun auch invariant unter *globalen Phasentransformationen*

$$\psi'(x) = \exp(-i\alpha) \psi(x), \quad \psi'^*(x) = \exp(+i\alpha) \psi^*(x). \quad (3.7)$$

Aus dem Noethertheorem ergibt sich als dazugehörige Erhaltungsgröße die Normierungskonstante

$$\mathbf{N} = \int d^3\vec{x} \psi^*(x) \psi(x) \text{ with } x = (t, \vec{x}). \quad (3.8)$$

Dies entspricht der Tatsache, daß die Zeitentwicklung des freien Einteilchensystems unitär ist. Die zeitliche Konstanz von (3.8) für Felder, die die Schrödingergleichung erfüllen, sollte der Leser als kleine Übung explizit nachrechnen!

Kommen wir nun zur Feldquantisierung. Man kann zeigen, daß für Teilchen in Konfigurationsräumen mit Dimension ≥ 3 Vielteilchensysteme aus *identischen Teilchen* genau zwei Quantisierungsmöglichkeiten bestehen, nämlich als *Bosonen* oder als *Fermionen*. Einen recht anschaulichen Beweis im Rahmen der Pfadintegralquantisierungsmethode findet sich in [LD70]. Im Rahmen der kanonischen Feldquantisierung werden die Poissonklammern des Hamiltonformalismus für Felder im Falle von Fermionen (Bosonen) einfach durch Antikommutator- bzw. Kommutatorregeln ersetzt:

$$[\psi(t, \vec{x}), \psi(t, \vec{y})]_{\pm} = [\mathbf{\Pi}(t, \vec{x}), \mathbf{\Pi}(t, \vec{y})]_{\pm} = 0, \quad \frac{1}{i} [\psi(t, \vec{x}), \mathbf{\Pi}(t, \vec{y})]_{\pm} = \delta^{(3)}(\vec{x} - \vec{y}). \quad (3.9)$$

Hier und im folgenden stehen die oberen Vorzeichen stets für Fermionen, die unteren für Bosonen. Man beachte, daß jetzt die fett gedruckten Symbole *Feldoperatoren* im *Heisenbergbild* sind.

Wir wollen nun sehen, wie dieser Formalismus mit dem *Teilchenbild* zusammenhängt. Um die Probleme mit kontinuierlichen Impulsvariablen zu umgehen, denken wir uns in einem unendlich

ausgedehnten „Wärmebad“ unseres Gases einen Würfel der Kantenlänge L abgegrenzt. Dies ist das sog. *Quantisierungsvolumen*. Wir werden noch sehen, daß diese Idee zu großen mathematischen Vorteilen führt. Wir werden an gegebener Stelle zum sog. *thermodynamischen Limes* übergehen, also die Kantenlänge L in einem bestimmten, unten noch genau zu spezifizierenden Sinne im Limes $L \rightarrow \infty$ betrachten. Wir können den Würfel als das im vorigen Abschnitt besprochene kleine Teilsystem des Gesamtwärmebads ansehen.

Da die Abgrenzung unseres Würfels lediglich eine gedachte Abteilung eines endlichen Teilvolumens in unserem homogenen Wärmebad ist, können wir an das Feld *räumlich periodische* Randbedingungen stellen, was die Rechnungen weiter erheblich vereinfacht, denn dann sind die üblichen Impulsoperatoren in der Ortsdarstellung selbstadjungiert, ohne daß wir komplizierte selbstadjungierte Erweiterungen vorzunehmen haben, wie dies für Randbedingungen einer „undurchlässigen Gefäßwand“ notwendig würde:

$$\psi(t, \vec{x} + L\vec{e}_i) = \psi(t, \vec{x}) \text{ for } i = 1, 2, 3. \quad (3.10)$$

Nun können wir den Feldoperator, der gemäß dem Heisenbergbild der freien Schrödingergleichung mit periodischen Randbedingungen (3.10) genügt, in seine *Fouriermoden* zerlegen:

$$\begin{aligned} \psi(t, \vec{x}) &= \sum_{\vec{n} \in \mathbb{Z}^3} \frac{1}{\sqrt{V}} \mathbf{a}(\vec{n}) \exp[-i\omega(\vec{n})t + i\vec{p}(\vec{n})\vec{x}] \\ \mathbf{a}(\vec{n}) &= \int_V \frac{d^3\vec{x}}{\sqrt{V}} \psi(t, \vec{x}) \exp[i\omega(\vec{n})t - i\vec{p}(\vec{n})\vec{x}] \\ &\text{with } \vec{p}(\vec{n}) = \frac{2\pi}{L}\vec{n}, \vec{n} \in \mathbb{Z}^3. \end{aligned} \quad (3.11)$$

Nun bestimmen wir die $\omega(\vec{n})$ aus den Bewegungsgleichungen des Heisenbergbildes für die Feldoperatoren, wobei wir von den Kommutatorrelationen (3.9) Gebrauch machen. Zunächst folgt aus (3.11):

$$\begin{aligned} [\mathbf{a}(\vec{n}), \mathbf{a}^\dagger(\vec{n}')]_{\pm} &= \int_V \frac{d^3\vec{x}'}{\sqrt{V}} \int_V \frac{d^3\vec{x}}{\sqrt{V}} [\psi(t, \vec{x}), \psi^\dagger(t, \vec{x}')]_{\pm} \times \\ &\times \exp\{-i[\omega(\vec{n}) - \omega(\vec{n}')t + i[\vec{p}(\vec{n}) - \vec{p}(\vec{n}')]\cdot\vec{x}]\} = \\ &= \int_V \frac{d^3\vec{x}}{V} \exp\{i[\vec{p}(\vec{n}) - \vec{p}(\vec{n}')]\cdot\vec{x}\} = \delta^{(3)}(\vec{n} - \vec{n}') := \begin{cases} 0 & \text{for } \vec{n} \neq \vec{n}' \\ 1 & \text{for } \vec{n} = \vec{n}'. \end{cases} \end{aligned} \quad (3.12)$$

Es ist nun einfach, die benötigten Größen \mathbf{H} und \mathbf{N} mittels der Operatoren $\mathbf{a}(\vec{n})$ auszudrücken. Dazu müssen wir nur (3.11) in (3.5) und (3.11) einsetzen, wobei wir sogleich die sog. *Normalordnungsvorschrift* verwenden, die wir weiter unten noch näher begründen werden. Diese besagt, daß in den Fällen, in denen bei der „Übersetzung klassischer Ausdrücke“ in den Operatorenformalismus der kanonischen Quantisierung Operatorordnungsprobleme auftauchen, alle Operatoren $\mathbf{a}(\vec{n})$ stets rechts von allen Operatoren $\mathbf{a}^\dagger(\vec{n})$ anzuordnen sein werden. Damit finden wir für den Hamilton- und den Teilchenzahloperator

$$\mathbf{H} = \frac{1}{2m} \sum_{\vec{n} \in \mathbb{Z}^3} \vec{p}^2(\vec{n}) \mathbf{N}(\vec{n}), \quad \mathbf{N} = \sum_{\vec{n} \in \mathbb{Z}^3} \mathbf{N}(\vec{n}) \text{ mit } \mathbf{N}(\vec{n}) = \mathbf{a}^\dagger(\vec{n})\mathbf{a}(\vec{n}). \quad (3.13)$$

Um die physikalische Bedeutung der Operatoren \mathbf{a} und \mathbf{a}^\dagger zu erhalten, lösen wir das Eigenwertproblem für die $\mathbf{N}(\vec{n})$. Als erstes bemerken wir, daß sie per definitionem selbstadjungiert

sind. Weiter zeigen wir, daß alle $\mathbf{N}(\vec{n})$ miteinander kommutieren, also kompatible Observablen beschreiben:

$$\begin{aligned}
 [\mathbf{N}(\vec{n}), \mathbf{N}(\vec{n}')] &= [\mathbf{a}^\dagger(\vec{n})\mathbf{a}(\vec{n}), \mathbf{a}^\dagger(\vec{n}')\mathbf{a}(\vec{n}')] = \\
 &= \mathbf{a}^\dagger(\vec{n}) [\mathbf{a}(\vec{n}), \mathbf{a}^\dagger(\vec{n}')\mathbf{a}(\vec{n}')] + [\mathbf{a}^\dagger(\vec{n}), \mathbf{a}^\dagger(\vec{n}')\mathbf{a}(\vec{n}')] \mathbf{a}(\vec{n}) = \\
 &= \mathbf{a}^\dagger(\vec{n}) \left\{ [\mathbf{a}(\vec{n}), \mathbf{a}^\dagger(\vec{n}')]_{\pm} \mathbf{a}(\vec{n}') \mp \mathbf{a}^\dagger(\vec{n}') [\mathbf{a}(\vec{n}), \mathbf{a}(\vec{n}')]_{\pm} \right\} + \\
 &+ \left\{ [\mathbf{a}^\dagger(\vec{n}), \mathbf{a}^\dagger(\vec{n}')]_{\pm} \mathbf{a}(\vec{n}') \mp \mathbf{a}^\dagger(\vec{n}') [\mathbf{a}^\dagger(\vec{n}), \mathbf{a}(\vec{n}')]_{\pm} \right\} \mathbf{a}(\vec{n}) = 0.
 \end{aligned} \tag{3.14}$$

Dabei haben wir die Identitäten

$$[\mathbf{A}, \mathbf{BC}]_{\pm} = [\mathbf{A}, \mathbf{B}]_{\pm} \mathbf{C} \mp \mathbf{B} [\mathbf{A}, \mathbf{C}]_{\pm} \quad \text{and} \quad [\mathbf{AB}, \mathbf{C}]_{\pm} = \mathbf{A} [\mathbf{B}, \mathbf{C}]_{\pm} \mp [\mathbf{A}, \mathbf{B}]_{\pm} \mathbf{C} \tag{3.15}$$

benutzt. Die $\mathbf{N}(\vec{n})$ kommutieren also allesamt miteinander, und folglich kann können die simultanen Eigenvektoren der Operatoren $\mathbf{N}(\vec{n})$ als Basis für den Hilbertraum dienen.

Der besseren Übersicht halber behandeln wir nun Fermionen und Bosonen zunächst getrennt. Beginnen wir mit dem Fall von Fermionen. In diesem Fall ist die Lösung des Eigenwertproblems besonders einfach, denn \mathbf{N} ist hier ein Projektionsoperator:

$$\mathbf{N}(\vec{n})^2 = \mathbf{a}(\vec{n})^\dagger \mathbf{a}(\vec{n}) \mathbf{a}(\vec{n})^\dagger \mathbf{a}(\vec{n}) = \left\{ \left\{ \mathbf{a}^\dagger(\vec{n}), \mathbf{a}(\vec{n}) \right\} - \mathbf{a}(\vec{n}) \mathbf{a}(\vec{n})^\dagger \right\} \mathbf{a}(\vec{n})^\dagger \mathbf{a}(\vec{n}) = \mathbf{N}(\vec{n}), \tag{3.16}$$

wobei wir benutzt haben, daß gemäß (3.12) für Fermionen stets $[\mathbf{a}^\dagger(\vec{n})]^2 = 0$. Wir nehmen nun an, $|\alpha\rangle$ sei Eigenvektor von \mathbf{N} mit Eigenwert α . Dann folgt aus sofort (3.16) sofort $\alpha^2 = \alpha$ und damit $\alpha = 0$ oder $\alpha = 1$. Wir nehmen nun an, daß die Eigenräume von \mathbf{N} nicht entartet sind. Wie wir unten sehen werden, ist diese Annahme dazu äquivalent, daß die Wirkung der Feldoperatoren auf dem Hilbertraume irreduzibel ist.

Jetzt untersuchen wir die Wirkung der Operatoren $\mathbf{a}(\vec{n})$ und $\mathbf{a}^\dagger(\vec{n})$ auf $|\alpha\rangle$. Die Kommutatoren mit \mathbf{N} folgen mit Hilfe von (3.15) aus (3.11):

$$[\mathbf{N}(\vec{n}), \mathbf{a}^\dagger(\vec{n})] = \mathbf{a}^\dagger(\vec{n}), \quad [\mathbf{N}(\vec{n}), \mathbf{a}(\vec{n})] = -\mathbf{a}(\vec{n}). \tag{3.17}$$

Daraus ergibt sich

$$\mathbf{N}(\vec{n})\mathbf{a}(\vec{n})|\alpha\rangle = \{[\mathbf{N}(\vec{n}), \mathbf{a}(\vec{n})] + \mathbf{a}(\vec{n})\mathbf{N}(\vec{n})\}|\alpha\rangle = (\alpha - 1)\mathbf{a}(\vec{n})|\alpha\rangle, \tag{3.18}$$

d.h. $\mathbf{a}(\vec{n})|\alpha\rangle$ ist entweder Eigenvektor von $\mathbf{N}(\vec{n})$ zum Eigenwert $\alpha - 1$ oder 0. Da die Eigenräume durch unsere obige Annahme jeweils nicht entartet sind, folgt daraus bis auf eine Phase

$$\mathbf{a}(\vec{n})|1\rangle = |0\rangle, \quad \mathbf{a}(\vec{n})|0\rangle = 0, \tag{3.19}$$

denn $\mathbf{a}(\vec{n})|0\rangle$ kann nicht von 0 verschieden sein, weil sonst $\mathbf{N}(\vec{n})$ den Eigenwert -1 haben müßte, was jedoch wegen (3.16) ausgeschlossen ist. Damit muß aber $\mathbf{a}(\vec{n})|1\rangle \neq 0$ sein, weil sonst $\mathbf{a}(\vec{n})$ überhaupt der Nulloperator wäre, was aber nicht mit den Antivertauschungsregeln der $\mathbf{a}(\vec{n})$ und $\mathbf{a}^\dagger(\vec{n})$ verträglich ist. Damit ist bis auf einen unerheblichen Phasenfaktor (3.19) bewiesen.

Daß die so definierten Beziehungen (3.19) korrekt normiert sind, wenn wir voraussetzen, daß beide $|\alpha\rangle$ normiert sind, folgt sofort aus

$$\langle \mathbf{a}(\vec{n})1 | \mathbf{a}(\vec{n})1 \rangle = \langle 1 | \mathbf{a}^\dagger(\vec{n})\mathbf{a}(\vec{n}) | 1 \rangle = \langle 1 | \mathbf{N}(\vec{n}) | 1 \rangle = \langle 1 | 1 \rangle = 1. \tag{3.20}$$

Genau analoge Rechnungen führen auf

$$\mathbf{a}^\dagger(\vec{n}) |0\rangle = |1\rangle, \quad \mathbf{a}^\dagger(\vec{n}) |1\rangle = 0. \quad (3.21)$$

Kehren wir nun wieder zu unserer ausführlichen Schreibweise zurück, folgt daraus, daß die $\mathbf{a}(\vec{n})$ Vernichtungsoperatoren und die $\mathbf{a}^\dagger(\vec{n})$ Erzeugungsoperatoren sind.

Es ist nämlich klar, daß der gesamte Hilbertraum durch fortgesetzte Anwendung der Operatoren $\mathbf{a}^\dagger(\vec{n})$ auf den Vakuumzustand $|\Omega\rangle$ aufgespannt wird, der dadurch definiert ist, daß für alle $\vec{n} \in \mathbb{Z}^3$ gilt

$$\mathbf{a}(\vec{n}) |\Omega\rangle = 0. \quad (3.22)$$

Der gesamte Hilbertraum wird durch die folgenden Zustände aufgespannt:

$$|N(\vec{n})\rangle = \prod_{\vec{n} \in \mathbb{Z}^3} [\mathbf{a}^\dagger(\vec{n})]^{N(\vec{n})} |\Omega\rangle, \quad (3.23)$$

wobei die $N(\vec{n}) \in \{0, 1\}$ alle möglichen Kombinationen von $\vec{n} \in \mathbb{Z}^3$ durchlaufen, wo nur endlich viele $N(\vec{n}) \neq 0$ sind. Die Phasenwahl muß dabei noch durch eine beliebige feste Standardanordnung der Operatoren $\mathbf{a}^\dagger(\vec{n})$ im Produkt erfolgen. Die Zustände zu gleichen Tupeln \vec{n} in (3.23) unterscheiden sich nämlich höchstens durch ein Vorzeichen und sind nur einmal zu nehmen, denn wir wollen ja ein VONS konstruieren. Es ist klar, daß die Zustände (3.23) unter Vertauschung zweier beliebiger \mathbf{a}^\dagger das Vorzeichen wechseln. Dies zeigt, daß die Antivertauschungsregeln beim kanonischen Quantisieren auf Fermionen führen. Man kann nämlich (3.23) als *antisymmetrisiertes direktes Produkt von Einteilchenzuständen* ansehen. Man nennt diese Konstruktion auch den *fermionischen Fockraum*. Das bedeutet daß diese Zustände den Fall beschreiben, daß gerade $N(\vec{n}) \in \{0, 1\}$ Teilchen mit jeweils einem wohlbestimmten Impuls $\frac{2\pi}{L}\vec{n}$ vorliegen hat. Die Antisymmetrie der Zustände verbietet dabei von selbst, daß mehr als ein Teilchen einen Einteilchenzustand besetzt. Dies ist das sog. *Paulische Ausschließungsprinzip*. Es ist aus den Kommutatorregeln klar, daß die Anwendung von $\mathbf{a}(\vec{n}')$ auf (3.23) den Zustand entweder annihiliert (wenn $N(\vec{n}') = 0$) ist oder für $N(\vec{n}') = 1$ den Zustand ergibt, der genau ein Teilchen mit dem Impuls $\frac{2\pi}{L}\vec{n}'$ weniger enthält als (3.23). Daher sind die \mathbf{a} Vernichtungsoperatoren. Entsprechend verhält es sich mit $\mathbf{a}^\dagger(\vec{n}')$, wobei allerdings der Zustand (3.23) gemäß (3.23) annihiliert wird, wenn $N(\vec{n}') = 1$ ist: Das Paulische Ausschließungsprinzip verbietet es eben, daß dem System ein Teilchen in einem Zustand, welcher bereits besetzt ist, hinzugefügt werden kann!

Die gleiche Konstruktion ergibt sich für Bosonen, nur daß hier die Erzeugungsoperatoren den *bosonischen Fockraum* aufspannen, der durch die *symmetrisierten* Produkte aus Einteilchenoperatoren aufgespannt wird:

$$|N(\vec{n})\rangle = \prod_{\vec{n} \in \mathbb{Z}^3} \frac{1}{\sqrt{N(\vec{n})!}} [\mathbf{a}^\dagger(\vec{n})]^{N(\vec{n})} |0\rangle \quad \text{with } N(\vec{n}) \in \mathbb{N} := \{0, 1, 2, \dots\}. \quad (3.24)$$

Dabei ist klar, daß auch hier wieder nur solche Produkte infrage kommen, für die die Gesamtteilchenzahl $\sum_{\vec{n}} N(\vec{n})$ endlich ist.

Aus der Heisenbergschen Bewegungsgleichung

$$\partial_t \psi(t, \vec{x}) = i [\mathbf{H}, \psi(t, \vec{x})] \quad (3.25)$$

können wir nun $\omega(\vec{n})$ in (3.11) bestimmen. Auf der rechten Seite setzen wir dazu $|\psi\rangle$ in der Form (3.11) ein und finden unter Benutzung der Kommutator- bzw. Antikommutatorrelationen (3.12)

$$[\mathbf{H}, \psi(t, \vec{x})] = -\frac{1}{2m} \vec{p}^2(\vec{n}) \psi(t, \vec{x}). \quad (3.26)$$

Drücken wir auch die linke Seite von Gl. (3.25) mittels der Fourierreihe (3.11) aus, ergibt sich die *Dispersionsrelation des freien Schrödingerfeldes*:

$$\omega(\vec{n}) = \frac{\vec{p}^2(\vec{n})}{2m}. \quad (3.27)$$

Daraus folgt auch, daß unser Ansatz mit zeitunabhängigen Vernichtungsoperatoren in (3.11) im hier verwendeten Heisenbergbild korrekt ist.

Jetzt wollen wir die makroskopischen Eigenschaften des Gases wechselwirkungsfreier Teilchen im thermodynamischen Gleichgewicht berechnen. Dazu benutzen wir das *großkanonische Ensemble*, welches im Sinne des Jaynesschen Prinzips dadurch gegeben ist, daß als gegebene Information der Erwartungswert der Energie $\mathcal{E} = \langle \mathbf{H} \rangle$ und der Gesamtteilchenzahl $\mathcal{N} = \langle \mathbf{N} \rangle$ gegeben ist. Wegen (1.154) folgt für den statistischen Operator

$$\mathbf{R} = \exp[-\Omega - \beta \mathbf{H} - \alpha \mathbf{N}]. \quad (3.28)$$

Wie wir in Abschnitt 3.4 gezeigt haben, können wir die alle thermodynamischen Größen aus der Zustandssumme bestimmen, die wir unter Zuhilfenahme der Teilchenzahlbasis (3.23) bzw. (3.24) bestimmen können. Für Fermionen und Bosonen gilt zunächst gleichermaßen:

$$Z = \text{Tr} [\exp(-\beta \mathbf{H} - \alpha \mathbf{N})] = \prod_{\vec{n} \in \mathbb{Z}^3} \sum_{N(\vec{n})} \exp \left[\left(-\beta \frac{\vec{p}^2(\vec{n})}{2m} - \alpha \right) N(\vec{n}) \right]. \quad (3.29)$$

Die Summe über $N(\vec{n})$ ist einfach zu bestimmen: Im Falle von Fermionen, läuft jedes $N(\vec{n})$ lediglich über die Werte 0 und 1. Diese Summe existiert also *für Fermionen für beliebige Werte von $\alpha \in \mathbb{R}$ und $\beta \geq 0$* . Für Bosonen hingegen handelt es sich um eine geometrische Reihe, weil jedes $N(\vec{n})$ von 0 bis ∞ läuft, so daß $\alpha \in \mathbb{R}_{>0}$ sein muß. Wir werden später sehen, daß dieser Unterschied zu gänzlich von Fermionen verschiedenem Verhalten führt.

Nehmen wir an, daß für Bosonen das Konvergenzkriterium erfüllt ist, erhalten wir für die Zustandssumme

$$Z = \prod_{\vec{n} \in \mathbb{Z}^3} \left[1 \pm \exp \left(-\beta \frac{\vec{p}^2(\vec{n})}{2m} - \alpha \right) \right]^{\pm 1}, \quad (3.30)$$

wobei das obere Vorzeichen für Fermionen, das untere für Bosonen steht.

Die Massiefunktion ist einfacher zu diskutieren:

$$\Omega = \ln Z = \pm \sum_{\vec{n} \in \mathbb{Z}^3} \ln \left[1 \pm \exp \left(-\beta \frac{\vec{p}^2(\vec{n})}{2m} - \alpha \right) \right]. \quad (3.31)$$

Zur Vollständigkeit der obigen Betrachtungen fügen wir noch Spinfreiheitsgrade hinzu. Der einzige Unterschied zu der obigen Rechnung ist, daß die Wellenfunktionen noch einen Spinorindex σ tragen, der für ein Teilchen mit Spin $s \in \frac{1}{2}\mathbb{N}_{\geq 0}$ die Werte $\{-s, -s+1, \dots, s-1, s\}$ annimmt. Für Elektronen ist z.B. $s = 1/2$. Bei der Besprechung der relativistischen Theorie werden wir sehen, daß unter sehr allgemeinen Annahmen Teilchen mit ganzzahliger Spinquantenzahl s zwingend Bosonen, solche mit halbzahligem s Fermionen sein müssen.

Es ist klar, daß wir in (3.31) lediglich noch zusätzlich über σ summieren müssen. Für freie Teilchen, also ideale Gase, multipliziert sich daher Ω lediglich um den "Entartungsfaktor" $g = 2s + 1$. Die vollständige Zustandssumme unter Berücksichtigung der Spinfreiheitsgrade ist also

$$\Omega = \ln Z = \pm g \sum_{\vec{n} \in \mathbb{Z}^3} \ln \left[1 \pm \exp \left(-\beta \frac{\vec{p}^2(\vec{n})}{2m} - \alpha \right) \right]. \quad (3.32)$$

3.1.1 Das ideale Fermigas

Wie oben bereits bemerkt, ist im Falle von Fermionen die Summe (3.32) für $\alpha \in \mathbb{R}$ und $\beta \in \mathbb{R}_{>0}$ wohldefiniert. Wir können ohne größere Schwierigkeiten zum Grenzfall eines unendlich großen Quantisierungsvolumens übergehen. Für große, jedoch endliche L , haben wir sehr viele Einteilchenimpulseigenwerte in einem kleinen Volumenelement $d^3\vec{p}$ im Impulsraum:

$$\vec{p}(\vec{n}) = \frac{2\pi}{L}\vec{n}, \quad \vec{n} \in \mathbb{Z}^3, \quad dn_i = \frac{L}{2\pi}dp_i \Rightarrow D(\vec{p})d^3\vec{p} = \frac{V}{(2\pi)^3}d^3\vec{p}, \quad (3.33)$$

wo $D(\vec{p})$ die Dichte der Einteilchenquantenzustände im Volumenelement um den Impuls \vec{p} ist. Im Grenzfall eines sehr großen Quantisierungsvolumens können wir daher die Summe in (3.32) durch ein Integral

$$\Omega = g \int d^3\vec{p} \frac{V}{(2\pi)^3} \ln \left[1 + \exp \left(-\beta \frac{\vec{p}^2}{2m} - \alpha \right) \right] \quad (3.34)$$

ersetzen. Durch Einführung sphärischer Koordinaten erhalten wir

$$\Omega = \frac{gV}{2\pi^2} \int_0^\infty dp p^2 \ln \left[1 + \exp \left(-\beta \frac{p^2}{2m} - \alpha \right) \right]. \quad (3.35)$$

Dieses Integral kann nicht geschlossen ausgewertet werden. Wir wollen daher eine Reihenentwicklung angeben und verschiedene Näherungen betrachten. Zunächst können wir uns des 'Logarithmus' entledigen, indem wir (3.35) partiell integrieren:

$$\Omega = \frac{gV\beta}{6\pi^2 m} \int_0^\infty dp \frac{p^4}{1 + \exp \left(\beta \frac{p^2}{2m} + \alpha \right)}. \quad (3.36)$$

Für $\alpha > 0$ können wir den Integranden ohne weitere Probleme entwickeln:

$$\begin{aligned} \Omega &= \frac{gV\beta}{6\pi^2 m} \sum_{k=0}^{\infty} (-1)^k \int_0^\infty dp p^4 \exp \left[-(k+1) \left(\frac{p^2}{2m} + \alpha \right) \right] \\ &= gV \left(\frac{m}{2\pi\beta} \right)^{3/2} \sum_{k=0}^{\infty} \frac{(-1)^k \exp[-\alpha(k+1)]}{(1+k)^{5/2}} \quad \text{für } \alpha > 0. \end{aligned} \quad (3.37)$$

Für $\alpha < 0$ müssen wir den Integrationsbereich in einen Teil $p^2 < -2m\alpha$ und den Rest umwandeln, was auf sehr komplizierte Integrale führt. Wir wollen dies nicht näher ausführen und uns mit den folgenden Näherungen begnügen.

3.1.2 Der klassische Limes

Die wichtigste Quanteneigenschaft des Fermigases ist durch Paulis Ausschließungsprinzip bedingt: Es verbietet, daß zwei Teilchen den gleichen Einteilchenzustand besetzen. Daraus folgt, daß das Gas klassisches Verhalten zeigen sollte, wenn im Mittel nur wenig Teilchen in ein Volumenelement des Phasenraums fallen, d.h. für kleine Dichten und kleine β , also hohe Temperaturen.

Formal erhalten wir den klassischen Limes für $\exp \alpha \gg 1$. Dann genügt es in (3.37) nur das Reihenglied mit $k = 0$ mitzunehmen:

$$\Omega = gV \left(\frac{m}{2\pi\beta} \right)^{3/2} \exp(-\alpha). \quad (3.38)$$

Daraus ergibt sich die mittlere Teilchenzahl und die mittlere Energie durch Ableiten nach den dazugehörigen Lagrangeparametern:

$$\mathcal{N} = -\frac{\partial\Omega}{\partial\alpha} = gV \left(\frac{m}{2\pi\beta}\right)^{3/2} \exp(-\alpha), \quad \mathcal{E} = -\frac{\partial\Omega}{\partial\beta} = \frac{3gV}{2} \left(\frac{m}{2\pi\beta}\right)^{3/2} \frac{1}{\beta} \exp(-\alpha). \quad (3.39)$$

Die erste Gleichung zeigt, daß unsere Betrachtungen nur für den Limes niedriger Dichten gilt, denn es ist nach Voraussetzung $\exp(-\alpha) \ll 1$.

Kombinieren wir beide Gleichungen (3.39) folgt die klassische *kalorische Zustandsgleichung des idealen Gases*:

$$\mathcal{E} = \frac{3}{2} \frac{\mathcal{N}}{\beta}. \quad (3.40)$$

Wegen (2.29) sind die statistischen Größen α und β mit den thermodynamischen Größen durch

$$T = \frac{1}{\beta}, \quad \alpha = -\mu\beta \quad (3.41)$$

verknüpft, wobei T die Temperatur, gemessen in Energieeinheiten, und μ das chemische Potential des Gases sind.

Aus (2.22) erhalten wir für die Entropie des idealen Gases

$$S = \Omega + \beta\mathcal{E} + \alpha\mathcal{N}. \quad (3.42)$$

Wegen (2.23) sind die natürlichen unabhängigen Zustandsgrößen für die Entropie \mathcal{E} , \mathcal{N} und V . Ersetzen wir also die übrigen thermodynamischen Größen vermöge (3.39) durch diese, erhalten wir die *Sackur-Tetrode-Formel* für die Entropie eines klassischen idealen Gases:

$$S = \frac{5}{2}\mathcal{N} + \mathcal{N} \ln \left[\frac{gV}{\mathcal{N}} \left(\frac{m\mathcal{E}}{3\pi\mathcal{N}} \right)^{3/2} \right]. \quad (3.43)$$

Daraus folgt aus (2.29) bzw. (2.30)

$$p = \frac{1}{\beta} \left(\frac{\partial S}{\partial V} \right)_{\mathcal{N}, \mathcal{E} = \text{const.}} = \frac{1}{\beta} \left(\frac{\partial \Omega}{\partial V} \right)_{\alpha, \beta = \text{const.}} = \frac{\mathcal{N}}{V\beta}, \quad (3.44)$$

was vermöge

$$pV = \mathcal{N}T \quad (3.45)$$

in die übliche Form der *Zustandsgleichung des idealen Gases* übergeht.

3.1.3 Der Quantenlimes

Die dem klassischen Regime entgegengesetzte Näherung ist das „Quantenregime“. Man spricht auch oft vom *entarteten Fermigas*. Man erhält es im Limes $\beta \rightarrow \infty$, wobei $\alpha/\beta = -\mu$ konstant gehalten wird.

Diesen Limes erhalten wir aus (3.37) durch Entwicklung nach Potenzen von T/μ . Wir haben stets Integrale der Form

$$F[f] = \frac{gV}{2\pi^2} \int_0^\infty dP P^2 f(P) \frac{1}{1 + \exp \left[\beta \left(\frac{P^2}{2m} - \mu \right) \right]} \quad (3.46)$$

zu berechnen. Durch Substitution von $x = \beta[P^2/(2m) - \mu]$ erhalten wir

$$F[f] = \frac{gV}{4\pi^2} \left(\frac{2m}{\beta}\right)^{3/2} \int_{-\mu\beta}^{\infty} dx \sqrt{x + \mu} f \left(\sqrt{\frac{2m(x + \mu\beta)}{\beta}} \right) \frac{1}{1 + \exp x}. \quad (3.47)$$

Beginnen wir mit der Berechnung der Teilchenzahl:

$$\mathcal{N} = F[1] = \frac{gV}{4\pi^2} \left(\frac{2m}{\beta}\right)^{3/2} \int_{-\mu\beta}^{\infty} dx \frac{\sqrt{x + \mu}}{1 + \exp x}. \quad (3.48)$$

Eine partielle Integration liefert

$$\mathcal{N} = \frac{gV}{6\pi^2} \left(\frac{2m}{\beta}\right)^{3/2} \int_{-\mu\beta}^{\infty} dx \frac{\exp x}{(1 + \exp x)^2} (x + \mu\beta)^{3/2}. \quad (3.49)$$

Der erste Faktor ist eine gerade Funktion von x , die für $|x| \rightarrow \infty$ exponentiell gedämpft ist, während der zweite Faktor langsam veränderlich für große $\beta\mu$ ist. Wir können also für $\mu\beta \gg 1$ zum einen die untere Integrationsgrenze zu $-\infty$ setzen, zum anderen aber den zweiten Faktor in Potenzen von $1/(\mu\beta)$ entwickeln:

$$(x + \mu\beta)^{3/2} = (\mu\beta)^{3/2} \left(1 + \frac{3}{2} \frac{x}{\mu\beta} + \frac{3}{8} \frac{x^2}{\mu^2\beta^2} + \dots \right). \quad (3.50)$$

Die Anwendung der Formeln

$$\begin{aligned} \int_{-\infty}^{\infty} dx \frac{\exp x}{(1 + \exp x)^2} &= - \frac{1}{1 + \exp x} \Big|_{-\infty}^{\infty} = 1, \\ \int_{-\infty}^{\infty} dx x^2 \frac{\exp x}{(1 + \exp x)^2} &= \frac{\pi^2}{3}, \end{aligned} \quad (3.51)$$

welch letztere im Anhang noch bewiesen wird, in (3.49) ergibt

$$\mathcal{N} = \frac{gV}{6\pi^2} (2m\mu)^{3/2} \left[1 + \frac{\pi^2}{8} \frac{1}{\mu^2\beta^2} + O[(\mu\beta)^{-4}] \right]. \quad (3.52)$$

Die mittlere Energie berechnet sich in der gleichen Weise zu

$$\mathcal{E} = \frac{gV}{20m\pi^2} (2m\mu)^{5/2} \left[1 + \frac{5}{8} \left(\frac{\pi}{\mu\beta} \right)^2 + O[(\mu\beta)^{-4}] \right]. \quad (3.53)$$

Ein wichtiges Resultat für das freie Elektronengasmodell für Metalle ist die spezifische Wärme. Dazu müssen wir \mathcal{E} bei festgehaltenem \mathcal{N} und V nach der Temperatur ableiten. Dazu benötigen wir zunächst die Ableitung von μ nach T . Aus $\mathcal{N}, V = \text{const}$ folgt mit der Kettenregel aus (3.52)

$$0 = \frac{3}{2} \sqrt{\mu} \left(\frac{\partial \mu}{\partial T} \right)_{V,N} + \frac{\pi^2}{4\sqrt{\mu}} T + O(T^2) \Rightarrow \left(\frac{\partial \mu}{\partial T} \right)_{V,N} = - \frac{\pi^2 T}{6\mu} + O(T^2). \quad (3.54)$$

Wieder mit der Kettenregel erhalten wir für die spezifische Wärme

$$c_{V,N} = \frac{1}{V} \left(\frac{\partial \mathcal{E}}{\partial T} \right)_{V,N} = \frac{(2s+1)}{12\pi^2 \hbar^3} (2m\mu)^{3/2} \pi^2 \frac{T}{\mu} + O(T^2) = \frac{\pi^2 \mathcal{N} k_B^2 T}{2\mu} + O(T^2). \quad (3.55)$$

Dies ist eine direkte Konsequenz des Quantencharakters der Elektronen, die sich in diesem Falle bei niedrigen Temperaturen dadurch bemerkbar macht, daß die Elektronen zur spezifischen Wärme des Metalls bei niedrigen Temperaturen kaum beitragen.

Wir müssen noch die Güte der obigen Näherung abschätzen. Die mittlere Teilchenzahl \mathcal{N} ergibt sich nach Substitution von $x = \beta p^2/(2m)$ zu

$$\mathcal{N} = \frac{gV}{4\pi^2} \left(\frac{2m}{\beta}\right)^{3/2} \int_0^\infty dx \frac{\sqrt{x}}{1 + \exp(x + \alpha)}, \quad (3.56)$$

so daß die Größe

$$Q = \frac{4\pi^2 \mathcal{N}}{gV} \left(\frac{\beta}{2m}\right)^{3/2} = \int_0^\infty dx \frac{\sqrt{x}}{1 + \exp(x + \alpha)} \quad (3.57)$$

ein Maß für das Quantenverhalten des Gases darstellt, denn der klassische Limes ergibt sich für $\alpha \rightarrow \infty$, und dann ist Q klein, während es im Quantenlimes $\alpha \rightarrow -\infty$ groß wird. Für den Fall der Leitungselektronen in Metallen, die sich in erster Näherung als ideales Fermigas behandeln lassen, ergibt sich die Temperatur, wo das Gas klassisches Verhalten zeigt, also bei $Q \approx 1$ zu $\Theta \approx 10^5 \text{ K} \approx 8.6 \text{ eV}$. Demnach ist das **Elektronengas**, welches eine grobe Näherung für das Verhalten der **Leitungselektronen in Metallen** darstellt, bei Raumtemperaturen ($T_{\text{Raum}} \simeq 20^\circ\text{C} \simeq 293 \text{ K}$) entartet. Dies erklärt, warum die Leitungselektronen zur spezifischen Wärme von Metallen bei Raumtemperatur kaum beitragen. Die Erklärung dieses Effekts geht auf Sommerfeld zurück: Drude hatte nämlich aus der Idee, daß die Leitungselektronen in Metallen näherungsweise als ideales Boltzmann-Gas behandelt werden können, eine recht gute Erklärung für Temperaturabhängigkeit der elektrischen Leitfähigkeit und der Wärmeleitfähigkeit (Wiedemann-Franz'sches Gesetz) von Metallen gefunden. Allerdings ergab sein Modell einen viel zu hohen Beitrag seines klassischen Elektronengases zur spezifischen Wärme. Wie Sommerfeld dann gezeigt hat, ist dies auf die Gasentartung bei Raumtemperatur zurückzuführen. Wir werden diese Zusammenhänge aus moderner Sicht im weiteren Verlauf dieser Vorlesung noch besprechen.

3.1.4 Das ideale Bosegas

Die großkanonische Zustandssumme für das Bosegas ergab sich gemäß (3.32) zu

$$\Omega = - \sum_{\vec{n} \in \mathbb{Z}^3} \ln \left[1 - \exp \left(-\beta \frac{\vec{p}^2(\vec{n})}{2m} - \alpha \right) \right], \quad (3.58)$$

die nur für $\alpha = -\mu\beta > 0$ einen Sinn ergibt.¹ Wie wir gleich sehen werden, führt die Beschränkung $\mu < 0$ zur Vorhersage der *Bose-Einstein-Kondensation*.

Für $\alpha \gg 0$ ergibt sich wieder derselbe klassische Limes wie oben für Fermigase diskutiert, denn es folgt im Limes $V \rightarrow \infty$ dieselbe Boltzmannsche großkanonische Zustandssumme (3.38). Dies ist nicht weiter überraschend, da im klassischen Limes die Quanteneigenschaften von Fermionen oder Bosonen vernachlässigbar sind, denn die mittleren Besetzungszahlen der Einteilchenzustände sind in diesem Fall klein.

¹Das ist zumindest im Limes $V \rightarrow \infty$ der Fall, denn dann wird das Integral, in das die Summe in diesem Fall übergeht, für $\mu > 0$ singular bei $p^2/(2m) = \mu$.

Betrachten wir also sogleich den Quantenlimes des idealen Bosegases. Für die mittlere Teilchenzahl ergibt die Ableitung von (3.58) nach α

$$\mathcal{N} = -\frac{\partial\Omega}{\partial\alpha} = g \sum_{\vec{n} \in \mathbb{Z}^3} \frac{1}{\exp\left(\beta \frac{\vec{p}^2(\vec{n})}{2m} + \alpha\right) - 1}. \quad (3.59)$$

Für $T \rightarrow 0^+$, also $\beta \rightarrow \infty$ trägt in dieser Summe nur der Term mit $\vec{n} = (0, 0, 0)$ bei. Drücken wir hingegen die Summe in naiver Weise im Limes $V \rightarrow \infty$ als Integral aus, ergibt sich

$$\mathcal{N}_f = \frac{gV}{2\pi^2} \int_0^\infty \frac{P^2 dP}{\exp\left(\beta \frac{P^2}{2m} + \alpha\right) - 1}, \quad (3.60)$$

und der Beitrag für kleine P verschwindet. Wir interpretieren also (3.60) als die Zahl von Teilchen in angeregten Zuständen. Die maximale Zahl bei festgehaltener Temperatur wird offenbar für $\alpha = 0$ erreicht. Substituieren wir $x = \beta P^2/(2m)$, ergibt sich

$$\mathcal{N}_f = \frac{gV}{4\pi^2} \left(\frac{2m}{\beta}\right)^{3/2} \int_0^\infty dx \frac{\sqrt{x}}{\exp(x + \alpha) - 1}. \quad (3.61)$$

Das Integral ist wohldefiniert für $\alpha > 0$ und ist monoton fallend mit α . Das Integral existiert auch für $\alpha = 0^+$, denn in der Nähe von $x = 0$ verhält sich der Integrand wie $1/\sqrt{x}$. Daher ist die Singularität bei $x = 0$ integrierbar.

Um zu klären, wie die Teilchenzahl im Tieftemperaturbereich korrekt zu berechnen ist, betrachten wir den thermodynamischen Limes $V \rightarrow \infty$ des Summe (3.59) etwas genauer. Der Limes ist dabei so zu verstehen, daß wir die Teilchendichte \mathcal{N}/V konstant halten. Es gibt nun eine kritische Temperatur T_c , bei der diese Teilchendichte gerade noch durch das Integral (3.61) erreicht werden kann, nämlich für $\alpha = 0$. Diese kritische Temperatur ist also mit $\beta_c = 1/T_c$ durch

$$\frac{\mathcal{N}}{V} = \frac{g}{4\pi^2} \left(\frac{2m}{\beta_c}\right)^{3/2} \int_0^\infty dx \frac{\sqrt{x}}{\exp x - 1}. \quad (3.62)$$

definiert. Für $T < T_c$, also $\beta > \beta_c$ müssen wir den $\vec{n} = (0, 0, 0)$ -Term gesondert berücksichtigen, d.h. für endliches Volumen gilt dann gemäß (3.59)

$$\frac{\mathcal{N}}{V} = \frac{g}{V} \frac{1}{\exp \alpha - 1} + \frac{\mathcal{N}_f}{V} \quad \text{für } \beta > \beta_c. \quad (3.63)$$

Damit die linke Seite im Limes $V \rightarrow \infty$ konstant bleiben kann, muß in diesem Falle also $\alpha \rightarrow 0$ geführt werden, so daß

$$\frac{\mathcal{N}}{V} = \frac{g}{V} \frac{1}{\exp \alpha - 1} + \frac{\mathcal{N}_f(\beta, \alpha \rightarrow 0)}{V} = \text{const.} \quad (3.64)$$

Da der zweite Term unabhängig von V ist, ist also bei der Limesbildung $V \rightarrow \infty$

$$\frac{g}{V} \frac{1}{\exp \alpha - 1} = \frac{\mathcal{N} - \mathcal{N}_f(T, \alpha = 0)}{V} := \frac{\mathcal{N}_{\text{BEC}}}{V} = n_{\text{BEC}} = \text{const.} \quad (3.65)$$

zu setzen. Dabei ist \mathcal{N}_{BEC} die Anzahl der Teilchen, die den Grundzustand besetzen und das *Bose-Einsteinkondensat* bilden. Wir können diese Zahl mit Hilfe von β_c ausdrücken, denn definitionsgemäß ist $\mathcal{N}_f(T = T_c, \alpha = 0) = \mathcal{N}$, also gilt

$$\mathcal{N}_{\text{BEC}} = \mathcal{N} \left[1 - \left(\frac{\beta_c}{\beta}\right)^{3/2} \right]. \quad (3.66)$$

Wir können nun die übrigen thermodynamischen Größen für den Fall $\beta > \beta_c$ im Limes $V \rightarrow \infty$ behandeln, indem wir zu unserem thermodynamischen Potential (3.58) zurückkehren. Bei endlichem Volumen ist zunächst nach einer partiellen Integration

$$\frac{\Omega}{V} = -\frac{g}{V} \ln[1 - \exp(-\alpha)] + \frac{g}{6\pi^2} \left(\frac{2m}{\beta}\right)^{3/2} \int_0^\infty dx \frac{x^{3/2}}{\exp(x + \alpha) - 1} \quad \text{für } \beta > \beta_c. \quad (3.67)$$

Der Limes $V \rightarrow \infty$ ist nun unter der Nebenbedingung (3.65) zu führen. Es ergibt sich, daß der Kondensatanteil verschwindet, also

$$\Omega = \frac{gV}{6\pi^2} \left(\frac{2m}{\beta}\right)^{3/2} \int_0^\infty dx \frac{x^{3/2}}{\exp x - 1} \quad \text{für } \beta > \beta_c \quad (3.68)$$

zu setzen ist, d.h. daß gemäß (2.29) nur die angeregten Zustände zum Druck und der inneren Energie des Gases beitragen. Letzteres ist unmittelbar verständlich, da die Kondensatteilchen den Grundzustand besetzen und wir den Energienullpunkt bei $E_0 = 0$ gewählt haben. Für die Energie ergibt sich aus (2.21)

$$\mathcal{E} = -\frac{\partial \Omega}{\partial \beta} = \frac{gV}{4\pi^2} \left(\frac{2m}{\beta}\right)^{3/2} \frac{1}{\beta} \int_0^\infty dx \frac{x^{3/2}}{\exp x - 1} \quad \text{für } \beta > \beta_c = \frac{3}{2} \frac{\Omega}{\beta}. \quad (3.69)$$

Die Entropie ergibt sich schließlich im thermodynamischen Limes wegen $\alpha = 0$ aus (3.42) zu

$$S = \Omega + \beta \mathcal{E} = \frac{5}{2} \Omega \quad \text{für } \beta > \beta_c. \quad (3.70)$$

Betrachten wir nun noch den Fall $\beta < \beta_c$ etwas genauer. Drücken wir dazu

$$\begin{aligned} \Omega &= -\frac{gV}{4\pi^2} \left(\frac{2m}{\beta}\right)^{3/2} \int_0^\infty dx \sqrt{x} \ln[1 - \exp(-x - \alpha)] \\ &= \frac{gV}{6\pi^2} \left(\frac{2m}{\beta}\right)^{3/2} \int_0^\infty dx \frac{x^{3/2}}{\exp(x + \alpha) - 1} \quad \text{für } \beta < \beta_c \end{aligned} \quad (3.71)$$

mittels der *Fugazität* $\gamma = \exp(-\alpha)$ aus, können wir das Integral nach Potenzen von γ entwickeln:

$$\int_0^\infty dx \frac{x^{3/2}}{\exp(x + \alpha) - 1} = \int_0^\infty dx x^{3/2} \frac{\gamma \exp(-x)}{1 - \gamma \exp(-x)} = \sum_{k=1}^\infty \gamma^k \int_0^\infty x^{3/2} \exp(-kx). \quad (3.72)$$

Das Integral läßt sich durch Substitution von $x = y^2$ durch ein Gaußintegral ausdrücken:

$$\int_0^\infty dx \frac{x^{3/2}}{\exp(x + \alpha) - 1} = 2 \sum_{k=1}^\infty \gamma^k \int_0^\infty dy y^4 \exp(-ky^2) = \frac{3\sqrt{\pi}}{4} \sum_{k=1}^\infty \frac{\gamma^k}{k^{5/2}}. \quad (3.73)$$

Die Reihe ist eine sog. *polylogarithmische Funktion*

$$\text{Li}_j(\gamma) = \sum_{k=1}^\infty \frac{\gamma^k}{k^j}. \quad (3.74)$$

Insgesamt ergibt sich also für (3.71)

$$\Omega = gV \left(\frac{m}{2\pi\beta}\right)^{3/2} \text{Li}_{5/2}(\gamma) \quad \text{für } \beta < \beta_c. \quad (3.75)$$

In der Tat ergibt sich für $\gamma \rightarrow 0$, also $\alpha \rightarrow \infty$ der klassische Limes, da $\text{Li}_{5/2}(\gamma) = \gamma + O(\gamma^2)$.

Die übrigen thermodynamischen Größen folgen wieder durch Ableitung nach den entsprechenden Parametern

$$\begin{aligned}\mathcal{E} &= -\frac{\partial \Omega}{\partial \beta} = \frac{3}{2}gV \left(\frac{m}{2\pi\beta}\right)^{3/2} \frac{1}{\beta} \text{Li}_{5/2}(\gamma), \\ \mathcal{N} &= gV \left(\frac{m}{2\pi\beta}\right)^{3/2} \text{Li}_{3/2}(\gamma).\end{aligned}\tag{3.76}$$

Die Entropie ergibt sich wieder mit Hilfe von (3.42):

$$S = gV \left(\frac{m}{2\pi\beta}\right)^{3/2} \left[\frac{5}{2} \text{Li}_{5/2}(\gamma) - \ln \gamma \text{Li}_{3/2}(\gamma) \right].\tag{3.77}$$

Für $\gamma = 1$ gehen die Polylogarithmusfunktionen in die ζ -Funktion über,

$$\zeta(x) = \text{Li}_x(1),\tag{3.78}$$

so daß wir insgesamt die thermodynamischen Größen für das Bosegas wie folgt schreiben können:

$$\begin{aligned}\Omega &= \begin{cases} gV \left(\frac{m}{2\pi\beta}\right)^{3/2} \text{Li}_{5/2}(\gamma) & \text{für } \beta < \beta_c \\ gV \left(\frac{m}{2\pi\beta}\right)^{3/2} \zeta(5/2) & \text{für } \beta \geq \beta_c, \end{cases} \\ \mathcal{N} &= \begin{cases} gV \left(\frac{m}{2\pi\beta}\right)^{3/2} \text{Li}_{3/2}(\gamma) & \text{für } \beta < \beta_c \\ \mathcal{N}_{\text{BEC}} + gV \left(\frac{m}{2\pi\beta}\right)^{3/2} \zeta(3/2) & \text{für } \beta \geq \beta_c, \end{cases} \\ \mathcal{E} &= \begin{cases} \frac{3}{2}gV \left(\frac{m}{2\pi\beta}\right)^{3/2} \frac{1}{\beta} \text{Li}_{5/2}(\gamma) & \text{für } \beta < \beta_c \\ \frac{3}{2}gV \left(\frac{m}{2\pi\beta}\right)^{3/2} \zeta(5/2) & \text{für } \beta \geq \beta_c. \end{cases}\end{aligned}\tag{3.79}$$

Dabei ist für vorgegebenes \mathcal{N} die inverse kritische Temperatur β_c durch (3.62), also

$$\mathcal{N} = gV \left(\frac{m}{2\pi\beta_c}\right)^{3/2} \zeta(3/2)\tag{3.80}$$

gegeben.

Literaturverzeichnis

- [Bal98] L. E. Ballentine, *Quantum Mechanics*, World Scientific, Singapore, New Jersey, London, Hong Kong (1998).
- [Fic79] E. Fick, *Einführung in die Grundlagen der Quantentheorie*, 4. Aufl., Aula-Verlag, Wiesbaden (1979).
- [Hee98] H. v. Hees, *Prinzipien der Quantentheorie* (1998).
<http://cyclotron.tamu.edu/hees/index.html>
- [Hob87] A. Hobson, *Concepts in Statistical Mechanics*, 2. Aufl., Gordon and Breach Science Publishers (1987).
- [LD70] M. G. G. Laidlaw, C. M. DeWitt, *Feynman Functional Integrals for Systems of Indistinguishable Particles*, Phys. Rev. D **3** (1970) 1375.
<http://link.aps.org/abstract/PRD/v3/i6/p1375>