

1. Grundlagen

1.1 Einleitung

Die statistische Mechanik behandelt die physikalischen Eigenschaften von Systemen, die aus sehr vielen Teilchen bestehen, d.h. Vielteilchensystemen, aufgrund der mikroskopischen Naturgesetze. Beispiele derartiger Vielteilchensysteme sind Gase, Flüssigkeiten, Festkörper in ihren verschiedenen Formen (kristallin, amorph), flüssige Kristalle, biologische Systeme, Sternmaterie, das Strahlungsfeld etc. Zu den interessierenden physikalischen Eigenschaften gehören Gleichgewichtseigenschaften (spezifische Wärme, thermische Ausdehnung, Elastizitätsmodul, magnetische Suszeptibilität, etc.) und Transporteigenschaften (Wärmeleitfähigkeit, elektrische Leitfähigkeit, etc.).

Schon lange vor ihrer Fundierung durch die statistische Mechanik wurde die Thermodynamik entwickelt, die allgemeine Beziehungen zwischen den makroskopischen Parametern des Systems lieferte. Der erste Hauptsatz der Thermodynamik wurde von Robert Mayer 1842 formuliert; dieser besagt, daß sich der Energieinhalt eines Körpers aus der Summe der an ihm geleisteten Arbeit und der ihm zugeführten Wärmemenge zusammensetzt:

$$dE = \delta Q + \delta A . \quad (1.1.1)$$

Daß Wärme eine Form von Energie ist oder präziser, daß Energie in Form von Wärme übertragen werden kann, wurde von Joule in den Jahren 1843–1849 experimentell überprüft (Reibungsversuche).

Der zweite Hauptsatz wurde von Clausius und von Lord Kelvin (W. Thomson¹) 1850 aufgestellt. Dieser geht von der Feststellung aus, daß ein und derselbe Zustand eines thermodynamischen Systems durch unterschiedliche Aufteilung der Energiezufuhr in Wärme und Arbeitsanteil erreicht werden kann, d.h. Wärme ist keine „Zustandsgröße“ (Zustandsgröße = physikalische Größe, die durch den Zustand des Systems bestimmt ist; später wird dieser Begriff mathematisch präzise festgelegt). Die wesentliche Erkenntnis des zweiten Hauptsatzes war, daß es eine Zustandsgröße S , die Entropie, gibt, die für reversible Veränderungen mit der Wärmezufuhr durch

¹ Geb. W. Thomson, der Name wurde später in Zusammenhang mit der Adelserhebung für hervorragende wissenschaftliche Verdienste angenommen.

$$\delta Q = TdS \quad (1.1.2)$$

zusammenhängt, während für irreversible Vorgänge $\delta Q < TdS$ ist. Der zweite Hauptsatz ist identisch mit der Feststellung, daß ein perpetuum mobile 2. Art unmöglich ist (p.m. 2. Art = periodisch arbeitende Maschine, die nur ein Wärmereservoir abkühlt und Arbeit leistet).

Die atomistische Fundierung der Thermodynamik wurde eingeleitet durch die kinetische Theorie verdünnter Gase. Die von Maxwell (1831–1879) gefundene Geschwindigkeitsverteilung erlaubt die Ableitung der kalorischen und thermischen Zustandsgleichung von idealen Gasen. Boltzmann (1844–1906) stellte im Jahre 1874 die nach ihm benannte grundlegende Transportgleichung auf. Er leitete daraus das Anwachsen der Entropie (H -Theorem) beim Streben ins Gleichgewicht her. Weiter erkannte Boltzmann, daß die Entropie mit der Zahl der Zustände $W(E, V, \dots)$, die mit den makroskopischen Angaben der Energie E , des Volumens V, \dots verträglich sind, durch

$$S \propto \log W(E, V, \dots) \quad (1.1.3)$$

zusammenhängt.² Es ist bemerkenswert, daß die atomistischen Grundlagen der Theorie der Gase zu einer Zeit geschaffen wurden, in der der atomare Aufbau der Materie nicht nur experimentell nicht gesichert war, sondern von angesehenen Physikern wie Mach (1828–1916) sogar zugunsten von Kontinuumstheorien erheblich angezweifelt wurde.

Die Beschreibung makroskopischer Systeme durch statistische Ensembles wurde von Boltzmann durch die Ergodenhypothese begründet. Grundlegende Beiträge zur Thermodynamik und statistischen Theorie der makroskopischen Systeme erfolgten von Gibbs (1839–1903) in den Jahren 1870–1900.

Erst durch die Quantentheorie (1925) war die korrekte Theorie im atomaren Bereich geschaffen. Im Unterschied zur klassischen statistischen Mechanik nennt man die auf der Quantentheorie basierende statistische Mechanik auch Quantenstatistik. Viele Phänomene wie z.B. elektronische Eigenschaften von Festkörpern, Supraleitung, Suprafluidität, Magnetismus, konnten erst auf der Basis der Quantenstatistik erklärt werden.

Auch heute gehört die statistische Mechanik noch zu den aktivsten Gebieten der theoretischen Physik: Theorie der Phasenübergänge, Theorie der Flüssigkeiten, ungeordnete Festkörper, Polymere, Membrane, biologische Systeme, granulare Medien, Oberflächen, Grenzflächen, Theorie der irreversiblen Prozesse, Systeme weit entfernt vom Gleichgewicht, nichtlineare Prozesse, Strukturbildung in offenen Systemen, biologische Vorgänge, und immer noch Magnetismus und Supraleitung.

Nach diesen Bemerkungen über den Problemkreis der statistischen Mechanik und über deren historische Entwicklung wollen wir nun einige charakteristische Probleme, die sich in der Theorie makroskopischer Systeme erge-

² Planck ergänzte diese Formel zu $S = k \log W$, wodurch die Ableitung $(\frac{\partial S}{\partial E})^{-1}$ gleich der absoluten Temperatur ist.

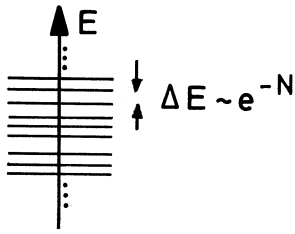


Abb. 1.1. Abstand der Energieniveaus für große Teilchenzahl N .

ben, aufzeigen. Konventionelle makroskopische Systeme wie Gase, Flüssigkeiten und Festkörper bei Zimmertemperatur bestehen aus $10^{19} - 10^{23}$ Teilchen pro cm^3 . Die Zahl der quantenmechanischen Eigenzustände wächst natürlich mit der Teilchenzahl. Wie wir später sehen werden, ist der Abstand der Energieniveaus von der Größe e^{-N} , d.h. die Energieniveaus liegen so dicht, daß schon die kleinste Störung das System von einem Zustand in einen anderen überführen kann, der faktisch die gleiche Energie besitzt. Soll man nun als Ziel die Berechnung des Bewegungsablaufs der $3N$ Koordinaten in der klassischen Physik anstreben, oder der Zeitabhängigkeit der Wellenfunktion in der Quantenmechanik, um daraus Zeitmittelwerte berechnen zu können? Beide derartige Unterfangen wären undurchführbar und sind auch überflüssig. Man kann weder Newtonsche Gleichungen noch die Schrödinger-Gleichung für $10^{19} - 10^{23}$ Teilchen lösen. Und selbst wenn wir die Lösung hätten, würden wir nicht alle Koordinaten und Geschwindigkeiten oder Quantenzahlen im quantenmechanischen Fall kennen, um die Anfangswerte festzulegen. Außerdem spielt die detaillierte Zeitentwicklung für die interessierenden, makroskopischen Eigenschaften keine Rolle. Darüber hinaus führt auch die schwächste Wechselwirkung (äußere Störung), die auch bei der besten Isolierung von der Umgebung immer noch vorhanden ist, zur Änderung des mikroskopischen Zustandes bei gleichbleibenden makroskopischen Eigenschaften.

Es sind für die weitere Diskussion zwei Begriffe zu definieren.

Mikrozustand: Dieser ist definiert durch die Wellenfunktion des Systems in der Quantenmechanik, bzw. durch alle Koordinaten und Impulse des Systems in der klassischen Physik.

Makrozustand: Dieser wird charakterisiert durch einige makroskopische Angaben (Energie, Volumen, ...).

Aus den vorhergehenden Überlegungen folgt, daß der Zustand eines makroskopischen Systems statistisch beschrieben werden muß. Der Umstand, daß das System während des Meßvorganges eine Verteilung von Mikrozuständen durchläuft, erfordert, daß der Makrozustand durch Angabe der Wahrscheinlichkeit für das Auftreten bestimmter Mikrozustände zu charakterisieren ist. Die Gesamtheit der mit ihrer Häufigkeit gewichteten Mikrozustände, die einen Makrozustand repräsentieren, nennt man ein *statistisches Ensemble*. Statt Ensemble ist auch der Ausdruck statistische Gesamtheit oder Gesamtheit gebräuchlich.

Obwohl der Zustand eines makroskopischen Systems durch ein statistisches Ensemble charakterisiert wird, sind die Vorhersagen für makroskopische Größen scharf. Mittelwerte und Schwankungsquadrate sind beide proportional zur Teilchenzahl N . Die relative Schwankung, das Verhältnis von Schwankung zu Mittelwert, geht im thermodynamischen Grenzfall gegen Null (siehe (1.2.21c)).

1.2 Exkurs über Wahrscheinlichkeitstheorie

Wir wollen an dieser Stelle einige grundlegende mathematische Definitionen aus der Wahrscheinlichkeitstheorie zusammenstellen, um dann den zentralen Grenzwertsatz abzuleiten.³

1.2.1 Wahrscheinlichkeitsdichte, charakteristische Funktion

Zunächst müssen wir den Begriff der *Zufallsvariable* erläutern. Darunter versteht man eine Größe X , die Werte x abhängig von den Elementen e einer „Ereignismenge“ E annehmen kann. Bei jeder einzelnen Beobachtung steht der Wert von X nicht fest, sondern es ist lediglich die Wahrscheinlichkeit für das Auftreten eines der möglichen Ergebnisse (Ereignisse) aus E bekannt. So ist bei einem idealen Würfel die Zufallsvariable die Augenzahl, die Werte von 1 bis 6 annehmen kann, wobei jedes dieser Ereignisse die Wahrscheinlichkeit $1/6$ besitzt. Hätten wir die genaue Kenntnis der Anfangslage und der durch den Wurf ausgeübten Kräfte, wäre das Ergebnis des Wurfes aus der klassischen Mechanik berechenbar. In Unkenntnis derartiger detaillierter Angaben kann man nur die oben genannte Wahrscheinlichkeitsaussage treffen. Sei $e \in E$ ein Ereignis aus der Menge E und P_e die zugehörige Wahrscheinlichkeit, dann ist bei einer großen Zahl N von Versuchen die Anzahl N_e , mit der das Ergebnis e auftritt, durch $\lim_{N \rightarrow \infty} \frac{N_e}{N} = P_e$ mit P_e verknüpft.

Es sei X eine *Zufallsvariable*. Die von X angenommenen Werte x seien kontinuierlich verteilt, und die *Wahrscheinlichkeitsdichte* der Zufallsvariablen sei $w(x)$. Das bedeutet, $w(x)dx$ ist die Wahrscheinlichkeit, daß X einen Wert im Intervall $[x, x + dx]$ annimmt. Die gesamte Wahrscheinlichkeit ist eins, d.h., $w(x)$ ist auf eins normiert:

$$\int_{-\infty}^{+\infty} dx w(x) = 1 . \quad (1.2.1)$$

Definition 1: Der Mittelwert von X ist durch

$$\langle X \rangle = \int_{-\infty}^{+\infty} dx w(x) x \quad (1.2.2)$$

³ Siehe z.B.: M. Fisz, *Wahrscheinlichkeitsrechnung und mathematische Statistik*, VEB Deutscher Verlag der Wissenschaften, Berlin, 1980.

definiert. Nun sei $F(X)$ eine Funktion der Zufallsvariablen X ; man nennt dann $F(X)$ Zufallsfunktion. Deren Mittelwert ist entsprechend (1.2.2) durch

$$\langle F(X) \rangle = \int dx w(x) F(x) \tag{1.2.2'}$$

definiert.⁴ Besondere Bedeutung haben die Potenzen von X , durch deren Mittelwerte die Momente der Wahrscheinlichkeitsdichte eingeführt werden.

Definition 2: Das n -te Moment der Wahrscheinlichkeitsdichte $w(x)$ ist durch

$$\mu_n = \langle X^n \rangle \tag{1.2.3}$$

definiert. (Das erste Moment von $w(x)$ ist durch den Mittelwert gegeben).

Definition 3: Das Schwankungsquadrat ist definiert durch

$$(\Delta x)^2 = \langle X^2 \rangle - \langle X \rangle^2 = \langle (X - \langle X \rangle)^2 \rangle. \tag{1.2.4}$$

Definition 4: Schließlich definieren wir die *charakteristische Funktion*:

$$\chi(k) = \int dx e^{-ikx} w(x) \equiv \langle e^{-ikX} \rangle. \tag{1.2.5}$$

Durch Umkehrung der Fouriertransformation kann $w(x)$ durch $\chi(k)$ ausgedrückt werden

$$w(x) = \int \frac{dk}{2\pi} e^{ikx} \chi(k). \tag{1.2.6}$$

Unter der Voraussetzung, daß alle Momente der Wahrscheinlichkeitsdichte $w(x)$ existieren, folgt aus Gl. (1.2.5) folgende Darstellung der charakteristischen Funktion

$$\chi(k) = \sum_n \frac{(-ik)^n}{n!} \langle X^n \rangle. \tag{1.2.7}$$

Falls das Wertespektrum von X diskret ist, also die Werte ξ_1, ξ_2, \dots mit Wahrscheinlichkeiten p_1, p_2, \dots angenommen werden, ist die Wahrscheinlichkeitsdichte von der Form

$$w(x) = p_1 \delta(x - \xi_1) + p_2 \delta(x - \xi_2) + \dots \tag{1.2.8}$$

Häufig wird die Wahrscheinlichkeitsdichte diskrete und kontinuierliche Bereiche haben.

Für mehrdimensionale (mehrkomponentige) Systeme $\mathbf{X} = (X_1, X_2, \dots)$ seien die von \mathbf{X} angenommenen Werte $\mathbf{x} = (x_1, x_2, \dots)$. Dann ist die Wahrscheinlichkeitsdichte $w(\mathbf{x})$ und besitzt folgende Bedeutung: Es ist $w(\mathbf{x}) d\mathbf{x} \equiv w(\mathbf{x}) dx_1 dx_2 \dots dx_N$ die Wahrscheinlichkeit, \mathbf{x} im Hyperkubus $\mathbf{x}, \mathbf{x} + d\mathbf{x}$ zu finden.

⁴ Falls die Integrationsgrenzen nicht angegeben sind, erstreckt sich die Integration von $-\infty$ bis $+\infty$. Eine analoge Vereinfachung der Notation verwenden wir auch bei mehrdimensionalen Integralen.

Definition 5: Der Mittelwert einer Funktion $F(\mathbf{X})$ der Zufallsvariablen \mathbf{X} ist durch

$$\langle F(\mathbf{X}) \rangle = \int d\mathbf{x} w(\mathbf{x}) F(\mathbf{x}) \quad (1.2.9)$$

definiert.

Theorem: Wahrscheinlichkeitsdichte von Zufallsfunktionen

Eine Funktion F der Zufallsvariablen \mathbf{X} ist selbst eine Zufallsvariable, die Werte f entsprechend einer Wahrscheinlichkeitsdichte $w_F(f)$ annimmt. Die Wahrscheinlichkeitsdichte $w_F(f)$ kann aus der Wahrscheinlichkeitsdichte $w(\mathbf{x})$ berechnet werden. Wir behaupten:

$$w_F(f) = \langle \delta(F(\mathbf{X}) - f) \rangle. \quad (1.2.10)$$

Beweis: Wir drücken die Wahrscheinlichkeitsdichte $w_F(f)$ durch ihre charakteristische Funktion aus

$$w_F(f) = \int \frac{dk}{2\pi} e^{ikf} \sum_n \frac{(-ik)^n}{n!} \langle F^n \rangle.$$

Setzen wir für $\langle F^n \rangle = \int d\mathbf{x} w(\mathbf{x}) F(\mathbf{x})^n$ ein, ergibt sich

$$w_F(f) = \int \frac{dk}{2\pi} e^{ikf} \int d\mathbf{x} w(\mathbf{x}) e^{-ikF(\mathbf{x})}$$

und nach Verwendung der Fourier-Darstellung der δ -Funktion $\delta(y) = \int \frac{dk}{2\pi} e^{iky}$ schließlich

$$w_F(f) = \int d\mathbf{x} w(\mathbf{x}) \delta(f - F(\mathbf{x})) = \langle \delta(F(\mathbf{X}) - f) \rangle,$$

also (1.2.10).

Definition 6: Für mehrdimensionale Verteilungen definieren wir *Korrelationen*

$$K_{ij} = \langle (X_i - \langle X_i \rangle)(X_j - \langle X_j \rangle) \rangle \quad (1.2.11)$$

der Zufallsvariablen X_i und X_j . Diese geben an, inwieweit Fluktuationen (Abweichungen vom Mittelwert) von X_i und X_j korreliert sind.

Falls die Variable x_i unabhängig von den übrigen auftritt, also die Wahrscheinlichkeitsdichte die Form

$$w(\mathbf{x}) = w_i(x_i) w'(\{x_k, k \neq i\})$$

hat, wobei $w'(\{x_k, k \neq i\})$ nicht von x_i abhängt, dann ist $K_{ij} = 0$ für $j \neq i$, also X_i und X_j unkorreliert.

Im Spezialfall

$$w(\mathbf{x}) = w_1(x_1) \cdots w_N(x_N)$$

sind die stochastischen Variablen X_1, \dots, X_N völlig unkorreliert.

Sei $P_n(x_1, \dots, x_{n-1}, x_n)$ die Wahrscheinlichkeitsdichte der Zufallsvariablen X_1, \dots, X_{n-1}, X_n . Dann ist die Wahrscheinlichkeitsdichte für eine Untermenge dieser Zufallsvariablen durch Integration von P_n über den Wertebereich der übrigen Zufallsvariablen gegeben; z.B. ist die Wahrscheinlichkeitsdichte $P_{n-1}(x_1, \dots, x_{n-1})$ für die Zufallsvariablen X_1, \dots, X_{n-1}

$$P_{n-1}(x_1, \dots, x_{n-1}) = \int dx_n P_n(x_1, \dots, x_{n-1}, x_n).$$

Schließlich führen wir noch den Begriff der *bedingten Wahrscheinlichkeit* und der bedingten Wahrscheinlichkeitsdichte ein.

Definition 7: Sei $P_n(x_1, \dots, x_n)$ die Wahrscheinlichkeit(sdichte). Die bedingte Wahrscheinlichkeit(sdichte)

$$P_{k|n-k}(x_1, \dots, x_k | x_{k+1}, \dots, x_n)$$

ist definiert durch die Wahrscheinlichkeit(sdichte) von x_1, \dots, x_k , wenn mit Sicherheit x_{k+1}, \dots, x_n vorliegen. Es gilt

$$P_{k|n-k}(x_1, \dots, x_k | x_{k+1}, \dots, x_n) = \frac{P_n(x_1, \dots, x_n)}{P_{n-k}(x_{k+1}, \dots, x_n)}, \quad (1.2.12)$$

wobei $P_{n-k}(x_{k+1}, \dots, x_n) = \int dx_1 \dots dx_k P_n(x_1, \dots, x_n)$.

Anmerkung zur bedingten Wahrscheinlichkeit: Formel (1.2.12) wird in der mathematischen Literatur meist als Definition eingeführt, sie kann aber folgendermaßen deduziert werden, wenn man die Wahrscheinlichkeiten mit den statistischen Häufigkeiten identifiziert. $P_n(x_1, \dots, x_k, x_{k+1}, \dots, x_n)$ bei festen x_{k+1}, \dots, x_n legt die Häufigkeit der x_1, \dots, x_k bei festen x_{k+1}, \dots, x_n fest. Die Wahrscheinlichkeitsdichte, die diesen Häufigkeiten entspricht, ist deshalb proportional zu $P_n(x_1, \dots, x_k, x_{k+1}, \dots, x_n)$. Da $\int dx_1 \dots dx_k P_n(x_1, \dots, x_k, x_{k+1}, \dots, x_n) = P_{n-k}(x_{k+1}, \dots, x_n)$, ist die auf 1 normierte bedingte Wahrscheinlichkeitsdichte

$$P_{k|n-k}(x_1, \dots, x_k | x_{k+1}, \dots, x_n) = \frac{P_n(x_1, \dots, x_n)}{P_{n-k}(x_{k+1}, \dots, x_n)}.$$

1.2.2 Zentraler Grenzwertsatz

Gegeben seien voneinander unabhängige Zufallsgrößen X_1, X_2, \dots, X_N , die durch gleiche aber unabhängige Wahrscheinlichkeitsverteilungen $w(x_1), w(x_2), \dots, w(x_N)$ charakterisiert seien. Der Mittelwert und das Schwankungsquadrat der X_1, \dots, X_N mögen existieren. Gesucht ist die Wahrscheinlichkeitsdichte für die Summe

$$Y = X_1 + X_2 + \dots + X_N \quad (1.2.13)$$

in der Grenze für $N \rightarrow \infty$. Es wird sich zeigen, daß die Wahrscheinlichkeitsdichte für Y durch eine Gauß-Verteilung gegeben ist.

Anwendungsbeispiele für diese Situation sind

- a) System von *nicht wechselwirkenden Teilchen*
 X_i = Energie des i -ten Teilchens, Y = Gesamtenergie des Systems
- b) *Random Walk* (Zufallsbewegung)
 X_i = Zuwachs beim i -ten Schritt, Y = Position nach N Schritten.

Zur übersichtlichen Berechnung der Wahrscheinlichkeitsdichte von Y ist es zweckmäßig, die Zufallsvariable Z einzuführen

$$Z = \sum_i (X_i - \langle X \rangle) / \sqrt{N} = (Y - N\langle X \rangle) / \sqrt{N}, \quad (1.2.14)$$

wo $\langle X \rangle \equiv \langle X_1 \rangle = \dots = \langle X_N \rangle$ definiert ist.

Nach (1.2.10) ist die Wahrscheinlichkeitsdichte $w_Z(z)$ der Zufallsvariablen Z durch

$$\begin{aligned} w_Z(z) &= \int dx_1 \dots dx_N w(x_1) \dots w(x_N) \delta\left(z - \frac{x_1 + \dots + x_N}{\sqrt{N}} + \sqrt{N}\langle X \rangle\right) \\ &= \int \frac{dk}{2\pi} e^{ikz} \int dx_1 \dots dx_N w(x_1) \dots w(x_N) e^{\frac{-ik(x_1 + \dots + x_N)}{\sqrt{N}} + ik\sqrt{N}\langle X \rangle} \\ &= \int \frac{dk}{2\pi} e^{ikz + ik\sqrt{N}\langle X \rangle} \left(\chi\left(\frac{k}{\sqrt{N}}\right)\right)^N \end{aligned} \quad (1.2.15)$$

gegeben, wobei $\chi(q)$ die charakteristische Funktion zu $w(x)$ ist.

Die Darstellung (1.2.7) der charakteristischen Funktion durch die Momente der Wahrscheinlichkeitsdichte kann man umformen, indem man den Logarithmus der Momentenentwicklung nimmt.

$$\chi(q) = \exp\left[-iq\langle X \rangle - \frac{1}{2}q^2(\Delta x)^2 + \dots q^3 + \dots\right], \quad (1.2.16)$$

d.h. allgemein

$$\chi(q) = \exp\left[\sum_{n=1}^{\infty} \frac{(-iq)^n}{n!} C_n\right]. \quad (1.2.16')$$

Im Unterschied zu (1.2.7) wird in (1.2.16') der Logarithmus der charakteristischen Funktion in eine Potenzreihe entwickelt. Die darin auftretenden Entwicklungskoeffizienten C_n heißen *Kumulanten n -ter Ordnung*. Sie sind durch die Momente (1.2.3) ausdrückbar; die drei niedrigsten haben die Gestalt:

$$\begin{aligned} C_1 &= \langle X \rangle = \mu_1 \\ C_2 &= (\Delta x)^2 = \langle X^2 \rangle - \langle X \rangle^2 = \mu_2 - \mu_1^2 \\ C_3 &= \langle X^3 \rangle - 3\langle X^2 \rangle \langle X \rangle + 2\langle X \rangle^3 = \mu_3 - 3\mu_1\mu_2 + 2\mu_1^3. \end{aligned} \quad (1.2.17)$$

Die Relationen (1.2.17) zwischen den Kumulanten und Momenten erhält man, indem man die Exponentialfunktion in (1.2.16) bzw. (1.2.16') entwickelt und die Koeffizienten der Taylor-Reihe mit (1.2.7) vergleicht.

Einsetzen von (1.2.16) in (1.2.15) ergibt

$$w_Z(z) = \int \frac{dk}{2\pi} e^{ikz - \frac{1}{2}k^2(\Delta x)^2 + \dots k^3 N^{-\frac{1}{2}} + \dots} . \quad (1.2.18)$$

Daraus folgt unter Vernachlässigung der Terme, die für große N wie $1/\sqrt{N}$ oder stärker gegen Null gehen

$$w_Z(z) = (2\pi(\Delta x)^2)^{-1/2} e^{-\frac{z^2}{2(\Delta x)^2}} \quad (1.2.19)$$

und schließlich unter Verwendung von $W_Y(y)dy = W_Z(z)dz$ für die Wahrscheinlichkeitsdichte der Zufallsvariablen Y

$$w_Y(y) = (2\pi N(\Delta x)^2)^{-1/2} e^{-\frac{(y - \langle X \rangle N)^2}{2(\Delta x)^2 N}} . \quad (1.2.20)$$

Dies ist der *zentrale Grenzwertsatz*. $w_Y(y)$ ist eine Gauß-Verteilung, obwohl keineswegs vorausgesetzt war, daß $w(x)$ eine solche Verteilung war.

$$\text{Mittelwert:} \quad \langle Y \rangle = N \langle X \rangle \quad (1.2.21a)$$

$$\text{Schwankungsbreite:} \quad \Delta y = \Delta x \sqrt{N} \quad (1.2.21b)$$

$$\text{Relative Schwankung:} \quad \frac{\Delta y}{\langle Y \rangle} = \frac{\Delta x \sqrt{N}}{N \langle X \rangle} = \frac{\Delta x}{\langle X \rangle \sqrt{N}} . \quad (1.2.21c)$$

Der zentrale Grenzwertsatz ist die mathematische Grundlage dafür, daß im Grenzfall großer N die Aussagen über Y scharf werden. Die relative Schwankung, das Verhältnis von Schwankung zu Mittelwert, geht nach (1.2.21c) im Grenzfall großer N gegen Null.

1.3 Ensemble in der klassischen Statistik

Obwohl die korrekte Theorie im atomistischen Bereich auf der Quantentheorie basiert und die klassische Statistik aus der Quantenstatistik abgeleitet werden kann, ist aus Gründen der anschaulichen Begriffsbildung zweckmäßig, die klassische Statistik schon jetzt parallel zur Quantenstatistik zu entwickeln. Wir werden später die klassische Verteilungsfunktion in ihrem Gültigkeitsgebiet aus der Quantenstatistik herleiten.

1.3.1 Phasenraum, Verteilungsfunktion

Wir betrachten N Teilchen in drei Dimensionen mit Koordinaten q_1, \dots, q_{3N} und Impulsen p_1, \dots, p_{3N} . Als *Phasenraum*, auch Γ -Raum genannt, definiert

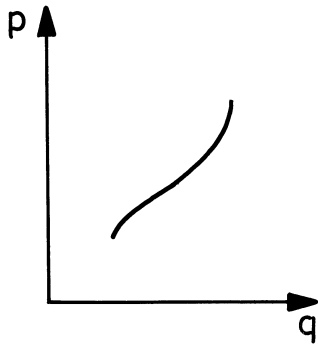


Abb. 1.2. Trajektorie im Phasenraum. Hier repräsentieren q und p die $6N$ Koordinaten und Impulse q_1, \dots, q_{3N} und p_1, \dots, p_{3N} .

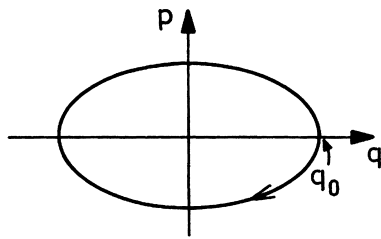


Abb. 1.3. Phasenbahn des eindimensionalen harmonischen Oszillators.

man den Raum, der durch die $6N$ Koordinaten und Impulse aufgespannt wird. Ein mikroskopischer Zustand wird durch einen Punkt im Γ -Raum und die Bewegung des Gesamtsystems durch eine Kurve im Phasenraum repräsentiert (Abb. 1.2), die man auch als Phasenbahn oder Phasentrajektorie bezeichnet.

Als Beispiel betrachten wir den *eindimensionalen harmonischen Oszillator*

$$\begin{aligned} q &= q_0 \cos \omega t \\ p &= -mq_0 \omega \sin \omega t, \end{aligned} \quad (1.3.1)$$

dessen Bahn im Phasenraum in Abb. 1.3 dargestellt ist.

Für große N ist der Phasenraum ein hochdimensionaler Raum. In aller Regel reicht die Kenntnis über ein derartiges System nicht aus, um seine Position im Phasenraum festzulegen. Wie schon im einleitenden Abschnitt 1.1 ausgeführt wurde, wird ein Makrozustand, charakterisiert durch die makroskopischen Angaben wie Energie E , Volumen V , Teilchenzahl N etc., durch sehr viele Mikrozustände, d.h. durch sehr viele Punkte im Phasenraum, gleichermaßen realisiert. Anstatt nun einen einzelnen dieser Mikrozustände willkürlich herauszugreifen, betrachten wir stattdessen eine Gesamtheit, d.h. ein Ensemble von Systemen, die alle ein und denselben Makrozustand repräsentieren, sich aber in allen möglichen dieser Mikrozustände befinden. Die Häufigkeit, mit der ein Punkt $(q, p) \equiv (q_1, \dots, q_{3N}, p_1, \dots, p_{3N})$ zur Zeit t auftritt, wird durch die Wahrscheinlichkeitsdichte $\rho(q, p, t)$ angegeben.

Die Einführung dieser Wahrscheinlichkeitsdichte ist nun keineswegs nur Ausdruck unserer Unkenntnis der detaillierten Form des Mikrozustandes, son-

den hat vielmehr den folgenden physikalischen Hintergrund. Jedes realistische makroskopische System spürt auch bei bester Isolierung noch eine Wechselwirkung mit der Umgebung. Diese Wechselwirkung ist zwar so schwach, daß die makroskopischen Eigenschaften des Systems nicht geändert werden, d.h. der Makrozustand ändert sich nicht, sie führt aber dazu, daß das System immer wieder seinen Mikrozustand ändert und somit beispielsweise während eines Meßvorganges eine Verteilung von Mikrozuständen durchläuft. Diese während eines kurzen Zeitintervalls durchlaufenen Zustände werden zur Verteilung $\rho(q, p)$ zusammengefaßt. Diese Verteilung beschreibt also nicht nur die statistischen Eigenschaften eines fiktiven Ensembles von sehr vielen Kopien des betrachteten Systems in den diversen Mikrozuständen, sondern das Einzelsystem. Statt bei zeitlichen Mittelwerten die sequentielle stochastische Folge dieser Mikrozustände zugrunde zu legen, betrachten wir die simultane Zeitentwicklung des Ensembles. Es wird die Hauptaufgabe des nächsten Kapitels sein, die bestimmten physikalischen Situationen entsprechenden Verteilungsfunktionen zu bestimmen. Dafür wird auch die Kenntnis der Bewegungsgleichung sehr wichtig sein, die wir im nächsten Abschnitt herleiten werden. Für große N kennen wir nur die Wahrscheinlichkeitsverteilung $\rho(q, p, t)$. Dabei ist

$$\rho(q, p, t) dq dp \equiv \rho(q_1, \dots, q_{3N}, p_1, \dots, p_{3N}, t) \prod_{i=1}^{3N} dq_i dp_i \quad (1.3.2)$$

gleich der Wahrscheinlichkeit zur Zeit t , ein System des Ensembles (oder das Einzelsystem im Laufe der Beobachtung) im Element $dq dp$ in der Umgebung des Punktes q, p im Γ -Raum zu finden. Man nennt $\rho(q, p, t)$ Verteilungsfunktion. Diese muß positiv, $\rho(q, p, t) \geq 0$ und normierbar sein. Hier stehen q, p für die Gesamtheit der Koordinaten und Impulse $q_1, \dots, q_{3N}, p_1, \dots, p_{3N}$.

1.3.2 Liouville-Gleichung

Wir wollen nun den Zeitverlauf von $\rho(q, p, t)$ ausgehend von der Anfangsverteilung $W(q_0, p_0)$ zur Zeit 0 aufgrund der Hamilton-Funktion $H \equiv H(p, q)$ bestimmen. Dabei setzen wir das System als abgeschlossen voraus. Die folgenden Resultate sind aber auch gültig, wenn H äußere zeitabhängige Kräfte enthält. Wir betrachten zunächst ein System, dessen Koordinaten im Phasenraum zur Anfangszeit 0 q_0, p_0 seien. Die zugehörige Trajektorie im Phasenraum, die aus den Hamiltonschen Bewegungsgleichungen folgt, sei $q(t; q_0, p_0), p(t; q_0, p_0)$, wobei wir in den folgenden Darlegungen die Anfangswerte der Trajektorien explizit anführen. Für eine einzelne Trajektorie ist die Wahrscheinlichkeitsdichte der Koordinaten q und Impulse p von der Gestalt

$$\delta(q - q(t; q_0, p_0)) \delta(p - p(t; q_0, p_0)) . \quad (1.3.3)$$

Hier ist $\delta(k) \equiv \delta(k_1) \dots \delta(k_{3N})$. Die Anfangswerte sind jedoch im allgemeinen nicht scharf vorgegeben, sondern es liegt eine Verteilung $W(q_0, p_0)$ vor. In

diesem Falle ergibt sich deshalb die Wahrscheinlichkeitsdichte im Phasenraum zur Zeit t durch Multiplikation von (1.3.3) mit $W(q_0, p_0)$ und Integration über die Anfangswerte

$$\rho(q, p, t) = \int dq_0 \int dp_0 W(q_0, p_0) \delta(q - q(t; q_0, p_0)) \delta(p - p(t; q_0, p_0)) . \quad (1.3.3')$$

Wir wollen eine Bewegungsgleichung für $\rho(q, p, t)$ ableiten. Dazu benützen wir die Hamiltonschen Bewegungsgleichungen

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i} . \quad (1.3.4)$$

Die Geschwindigkeit im Phasenraum

$$\mathbf{v} = (\dot{q}, \dot{p}) = \left(\frac{\partial H}{\partial p}, -\frac{\partial H}{\partial q} \right) \quad (1.3.4')$$

erfüllt

$$\operatorname{div} \mathbf{v} \equiv \sum_i \left(\frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i} \right) = \sum_i \left(\frac{\partial^2 H}{\partial q_i \partial p_i} - \frac{\partial^2 H}{\partial p_i \partial q_i} \right) = 0 . \quad (1.3.5)$$

D.h. die Bewegung im Phasenraum kann als „Strömung“ einer inkompressiblen „Flüssigkeit“ veranschaulicht werden.

Durch Ableiten von (1.3.3') nach der Zeit findet man

$$\begin{aligned} & \frac{\partial \rho(q, p, t)}{\partial t} \\ &= - \sum_i \int dq_0 dp_0 W(q_0, p_0) \left(\dot{q}_i(t; q_0, p_0) \frac{\partial}{\partial q_i} + \dot{p}_i(t; q_0, p_0) \frac{\partial}{\partial p_i} \right) \\ & \quad \times \delta(q - q(t; q_0, p_0)) \delta(p - p(t; q_0, p_0)) . \end{aligned} \quad (1.3.6)$$

Drückt man die Geschwindigkeit im Phasenraum durch (1.3.4) aus und verwendet die δ -Funktionen in (1.3.6), erhält man nach (1.3.3') und (1.3.5) die folgenden Darstellungen der Bewegungsgleichung für $\rho(q, p, t)$

$$\begin{aligned} \frac{\partial \rho}{\partial t} &= - \sum_i \left(\frac{\partial}{\partial q_i} \rho \dot{q}_i + \frac{\partial}{\partial p_i} \rho \dot{p}_i \right) \\ &= - \sum_i \left(\frac{\partial \rho}{\partial q_i} \dot{q}_i + \frac{\partial \rho}{\partial p_i} \dot{p}_i \right) \\ &= \sum_i \left(-\frac{\partial \rho}{\partial q_i} \frac{\partial H}{\partial p_i} + \frac{\partial \rho}{\partial p_i} \frac{\partial H}{\partial q_i} \right) . \end{aligned} \quad (1.3.7)$$

Unter Verwendung der Poisson Klammern⁵ kann die letzte Zeile von Gl. (1.3.7) auch in der Form

⁵ $\{u, v\} = \sum_i \left[\frac{\partial u}{\partial p_i} \frac{\partial v}{\partial q_i} - \frac{\partial u}{\partial q_i} \frac{\partial v}{\partial p_i} \right]$; die Vorzeichenkonvention ist in der Literatur nicht einheitlich.

$$\frac{\partial \rho}{\partial t} = -\{H, \rho\} \quad (1.3.8)$$

geschrieben werden. Dies ist die *Liouville-Gleichung*, die grundlegende Bewegungsgleichung für die klassische Verteilungsfunktion $\rho(q, p, t)$.

Ergänzungen:

Wir besprechen noch äquivalente Darstellungen der Liouville-Gleichung und Folgerungen daraus.

(i) Die erste Zeile der Gleichungskette (1.3.7) kann verkürzt als Kontinuitätsgleichung

$$\frac{\partial \rho}{\partial t} = -\operatorname{div} \mathbf{v} \rho \quad (1.3.9)$$

geschrieben werden. Man kann sich die Bewegung des Ensembles im Phasenraum als die Strömung einer Flüssigkeit veranschaulichen. Dann ist (1.3.9) die Kontinuitätsgleichung für die Dichte und die Gleichung (1.3.5) besagt, daß die „Flüssigkeit“ inkompressibel ist.

(ii) Wir wollen nochmals die Analogie der Bewegung im Phasenraum zur Hydrodynamik einer Flüssigkeit aufgreifen. In unserer bisherigen Darstellung haben wir die Dichte an einem festen Punkt q, p des Γ -Raums betrachtet. Wir können aber auch die Bewegung vom Standpunkt eines mit der „Strömung“ mitbewegten Beobachters betrachten, d.h. nach dem Zeitverlauf von $\rho(q(t), p(t), t)$ fragen, wobei wir in $q(t)$ und $p(t)$ der Kürze halber die Anfangskoordinaten q_0, p_0 weglassen. Die zweite Zeile von Gl. (1.3.7) besagt

$$\frac{d}{dt} \rho(q(t), p(t), t) = 0. \quad (1.3.10)$$

Die Verteilungsfunktion ist längs einer Phasenraumtrajektorie konstant.

(iii) Wir untersuchen nun die Veränderung eines Volumelements $d\Gamma$ im Phasenraum. Zur Anfangszeit sei in einem Volumelement $d\Gamma_0$ eine gleichmäßig verteilte Zahl dN von Repräsentanten des Ensembles vorhanden. Aufgrund der Bewegung im Phasenraum nehmen diese zur Zeit t ein Volumen $d\Gamma$ ein. Das bedeutet, daß die Dichte ρ zur Anfangszeit $\frac{dN}{d\Gamma_0}$ ist, während sie zur Zeit t durch $\frac{dN}{d\Gamma}$ gegeben ist. Aus (1.3.10) folgt die Gleichheit dieser Größen, woraus (Abb. 1.4) die Gleichheit der Volumina

$$d\Gamma = d\Gamma_0 \quad (1.3.11)$$

folgt. Gleichung (1.3.8) ist als Liouville-Theorem aus der Mechanik bekannt.⁶ Es wird dort mit Hilfe der Theorie kanonischer Transformationen aus der Jacobischen Determinante berechnet. Man kann in umgekehrter Richtung, ausgehend von (1.3.11), Gl. (1.3.10) und die Liouville-Gleichung (1.3.8) ableiten.

⁶ L.D. Landau und E.M. Lifschitz, *Lehrbuch der Theoretischen Physik I, Mechanik*, Gl. (46.5), Akademie Verlag, Berlin, 1969

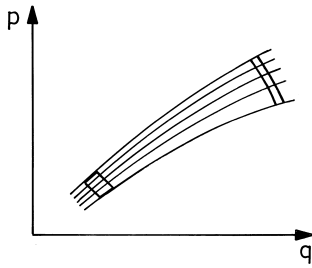


Abb. 1.4. Zeitliche Veränderung eines Volumenelements im Phasenraum, bei gleichbleibendem Volumsinhalt.

1.4 Quantenstatistik

1.4.1 Dichtematrix für reine und gemischte Gesamtheiten⁷

Die Dichtematrix ist von überragender Bedeutung für den Aufbau der Quantenstatistik; es sind dafür auch die Bezeichnungen statistischer Operator und Dichteoperator gebräuchlich.

Das System befinde sich im Zustand $|\psi\rangle$. Die Observable A hat in diesem Zustand den Mittelwert

$$\langle A \rangle = \langle \psi | A | \psi \rangle . \quad (1.4.1)$$

Die Struktur des Mittelwertes legt es nahe, die *Dichtematrix*

$$\rho = |\psi\rangle \langle \psi| \quad (1.4.2)$$

zu definieren. Es gilt:

$$\langle A \rangle = \text{Sp}(\rho A) \quad (1.4.3a)$$

$$\text{Sp } \rho = 1 , \rho^2 = \rho , \rho^\dagger = \rho . \quad (1.4.3b,c,d)$$

Dabei lautet die Definition der Spur (Sp)

$$\text{Sp } X = \sum_n \langle n | X | n \rangle , \quad (1.4.4)$$

wobei $\{|n\rangle\}$ ein beliebiges vollständiges Orthonormalsystem ist. Wegen

$$\begin{aligned} \text{Sp } X &= \sum_n \sum_m \langle n | m \rangle \langle m | X | n \rangle = \sum_m \sum_n \langle m | X | n \rangle \langle n | m \rangle \\ &= \sum_m \langle m | X | m \rangle \end{aligned}$$

ist die Spur unabhängig von der Basis.

⁷ Siehe z.B. F. Schwabl, *Quantenmechanik*, 6., erw. Aufl., Springer, Heidelberg, 2002, Kap. 20. Im folgenden wird dieses Lehrbuch mit QM I zitiert.

Anmerkung: Beweise von (1.4.3a–c):

$$\text{Sp } \rho A = \sum_n \langle n | \psi \rangle \langle \psi | A | n \rangle = \sum_n \langle \psi | A | n \rangle \langle n | \psi \rangle = \langle \psi | A | \psi \rangle \quad ,$$

$$\text{Sp } \rho = \text{Sp } \rho \mathbb{1} = \langle \psi | \mathbb{1} | \psi \rangle = 1 \quad , \quad \rho^2 = |\psi\rangle \langle \psi | \psi \rangle \langle \psi | = |\psi\rangle \langle \psi | = \rho \quad .$$

Wenn die untersuchten Systeme oder Objekte alle in ein und demselben Zustand $|\psi\rangle$ sind, spricht man von einer *reinen Gesamtheit* oder man sagt, die Systeme befinden sich in einem *reinen Zustand*.

Neben dem der Quantenmechanik inhärenten, dem einzelnen Zustand innewohnenden statistischen Charakter, kann in einem Ensemble darüber hinaus noch eine statistische Verteilung von Zuständen vorliegen. Liegt ein Ensemble mit unterschiedlichen Zuständen vor, nennt man dieses eine *gemischte Gesamtheit*, ein *Gemisch*, oder man spricht von einem *gemischten Zustand*. Wir nehmen an, daß der Zustand $|\psi_1\rangle$ mit der Wahrscheinlichkeit von p_1 , der Zustand $|\psi_i\rangle$ mit der Wahrscheinlichkeit p_i , u.s.w. realisiert ist, mit

$$\sum_i p_i = 1 \quad .$$

Der Mittelwert von A ist dann

$$\langle A \rangle = \sum_i p_i \langle \psi_i | A | \psi_i \rangle \quad . \quad (1.4.5)$$

Auch diesen Mittelwert kann man durch die nun folgendermaßen definierte *Dichtematrix*

$$\rho = \sum_i p_i |\psi_i\rangle \langle \psi_i| \quad (1.4.6)$$

darstellen. Es gilt:

$$\langle A \rangle = \text{Sp } \rho A \quad (1.4.7a)$$

$$\text{Sp } \rho = 1 \quad (1.4.7b)$$

$$\rho^2 \neq \rho \quad \text{und} \quad \text{Sp } \rho^2 < 1, \text{ falls } p_i \neq 0 \text{ ist für mehr als ein } i \quad (1.4.7c)$$

$$\rho^\dagger = \rho \quad . \quad (1.4.7d)$$

Die Ableitungen für diese Relationen und weitere Ergänzungen über Dichtematrizen von gemischten Gesamtheiten werden in Abschnitt 1.5.2 dargestellt.

1.4.2 Von Neumann-Gleichung

Aus der *Schrödinger-Gleichung* und ihrer Adjungierten

$$i\hbar \frac{\partial}{\partial t} |\psi, t\rangle = H |\psi, t\rangle, \quad -i\hbar \frac{\partial}{\partial t} \langle \psi, t| = \langle \psi, t| H$$

folgt

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} \rho &= i\hbar \sum_i p_i \left(|\dot{\psi}_i\rangle \langle \psi_i| + |\psi_i\rangle \langle \dot{\psi}_i| \right) \\ &= \sum_i p_i (H |\psi_i\rangle \langle \psi_i| - |\psi_i\rangle \langle \psi_i| H) . \end{aligned}$$

Daraus ergibt sich

$$\frac{\partial}{\partial t} \rho = -\frac{i}{\hbar} [H, \rho] \quad (1.4.8)$$

die *von Neumann-Gleichung*, die das quantenmechanische Pendant der Liouville-Gleichung darstellt. Diese beschreibt die Zeitentwicklung der Dichtematrix im Schrödinger-Bild. Sie gilt auch für zeitabhängige Hamilton-Operatoren H . Sie ist nicht zu verwechseln mit der Bewegungsgleichung für Heisenberg-Operatoren, in der auf der rechten Seite ein positives Zeichen auftritt.

Der Mittelwert einer Observablen A ist durch

$$\langle A \rangle_t = \text{Sp}(\rho(t)A) \quad (1.4.9)$$

gegeben, wobei sich $\rho(t)$ durch Lösung der von Neumann-Gleichung (1.4.8) ergibt. Die Zeitabhängigkeit des Mittelwertes wird durch den Index t hervorgehoben.

Die von Neumann-Gleichung wird uns im nächsten Kapitel bei der Aufstellung der Gleichgewichtsdichtematrizen wieder begegnen und ist natürlich für alle zeitabhängigen Vorgänge von fundamentaler Bedeutung.

Wir behandeln nun den Übergang zur *Heisenberg-Darstellung*. Die formale Lösung der Schrödinger-Gleichung ist von der Form

$$|\psi(t)\rangle = U(t, t_0) |\psi(t_0)\rangle , \quad (1.4.10)$$

wobei $U(t, t_0)$ ein unitärer Operator und $|\psi(t_0)\rangle$ der Anfangszustand zur Zeit t_0 ist. Daraus ergibt sich für die Zeitentwicklung der Dichtematrix

$$\rho(t) = U(t, t_0) \rho(t_0) U(t, t_0)^\dagger . \quad (1.4.11)$$

(Für zeitunabhängiges H ist $U(t, t_0) = e^{-iH(t-t_0)/\hbar}$.)

Der Mittelwert einer Observablen A kann sowohl im Schrödinger-Bild wie auch im Heisenberg-Bild berechnet werden

$$\langle A \rangle_t = \text{Sp}(\rho(t)A) = \text{Sp}(\rho(t_0)U(t, t_0)^\dagger A U(t, t_0)) = \text{Sp}(\rho(t_0)A_H(t)) . \quad (1.4.12)$$

Hier ist $A_H(t) = U^\dagger(t, t_0) A U(t, t_0)$ der Operator in der Heisenberg-Darstellung. Die Dichtematrix in der Heisenberg-Darstellung $\rho(t_0)$ ist zeitunabhängig.

*1.5 Ergänzungen

*1.5.1 Binomial- und Poisson-Verteilung

Wir besprechen nun zwei häufig auftretende Wahrscheinlichkeitsverteilungen. Gegeben sei ein Intervall der Länge L , das in zwei Teilintervalle $[0, a]$ und $[a, L]$ aufgeteilt werde. Es werden nun N unterscheidbare Objekte (Teilchen) völlig stochastisch auf die beiden Teilintervalle verteilt, so daß die Wahrscheinlichkeit eines Teilchens, sich im ersten oder zweiten Teilintervall zu befinden, durch $\frac{a}{L}$ oder $(1 - \frac{a}{L})$ gegeben ist. Die Wahrscheinlichkeit dafür, daß sich n Teilchen im Intervall $[0, a]$ befinden, ist dann durch die *Binomialverteilung*

$$w_n = \left(\frac{a}{L}\right)^n \left(1 - \frac{a}{L}\right)^{N-n} \binom{N}{n} \quad (1.5.1)$$

gegeben,⁸ wobei der kombinatorische Faktor $\binom{N}{n}$ die Zahl der Möglichkeiten gibt, aus N Objekten n auszuwählen. Der Mittelwert von n ist

$$\langle n \rangle = \sum_{n=0}^N n w_n = \frac{a}{L} N \quad (1.5.2a)$$

und das Schwankungsquadrat

$$(\Delta n)^2 = \frac{a}{L} \left(1 - \frac{a}{L}\right) N. \quad (1.5.2b)$$

Wir betrachten nun den Grenzfall $L \gg a$. Zunächst kann w_n mit $\binom{N}{n} = \frac{N \cdot (N-1) \cdots (N-n+1)}{n!}$ in der Form

$$\begin{aligned} w_n &= \left(\frac{aN}{L}\right)^n \left(1 - \frac{a}{L}\right)^{N-n} \frac{1}{n!} 1 \cdot \left(1 - \frac{1}{N}\right) \cdots \left(1 - \frac{n-1}{N}\right) \\ &= \bar{n}^n \frac{1}{n!} \left(1 - \frac{\bar{n}}{N}\right)^N \frac{1 \cdot \left(1 - \frac{1}{N}\right) \cdots \left(1 - \frac{n-1}{N}\right)}{\left(1 - \frac{a}{L}\right)^n} \end{aligned} \quad (1.5.3a)$$

geschrieben werden, wobei für den Mittelwert (1.5.2a) die Abkürzung $\bar{n} = \frac{aN}{L}$ eingeführt wurde. Im Grenzfall $\frac{a}{L} \rightarrow 0, N \rightarrow \infty$ bei endlichem \bar{n} geht der dritte Faktor in (1.5.3a) in $e^{-\bar{n}}$ über und der letzte Faktor wird 1, so daß sich für die Wahrscheinlichkeitsverteilung

$$w_n = \frac{\bar{n}^n}{n!} e^{-\bar{n}} \quad (1.5.3b)$$

ergibt. Dies ist die *Poisson-Verteilung*, die schematisch in Abb. 1.5 dargestellt ist. Die Poisson-Verteilung hat folgende Eigenschaften

⁸ Eine bestimmte Anordnung mit n Teilchen im Intervall a und $N - n$ in $L - a$, z.B. erstes Teilchen in a , zweites Teilchen in $L - a$, drittes Teilchen in $L - a$, u.s.w. hat die Wahrscheinlichkeit $\left(\frac{a}{L}\right)^n \left(1 - \frac{a}{L}\right)^{N-n}$. Daraus erhält man w_n durch Multiplikation mit der Zahl der Anordnungen $\binom{N}{n}$.

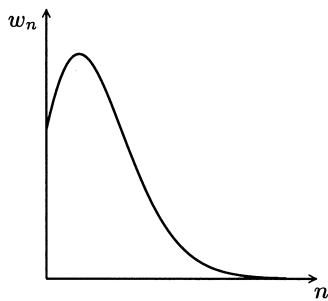


Abb. 1.5. Poisson-Verteilung

$$\sum_n w_n = 1, \quad \langle n \rangle = \bar{n}, \quad (\Delta n)^2 = \bar{n}. \quad (1.5.4a,b,c)$$

Die ersten beiden Relationen folgen unmittelbar aus der Herleitung der Poisson-Verteilung aus der Binomialverteilung. Sie werden in Aufgabe 1.5 zusammen mit 1.5.4c direkt aus 1.5.3b hergeleitet. Die relative Schwankung ist deshalb

$$\frac{\Delta n}{\bar{n}} = \frac{1}{\bar{n}^{1/2}}. \quad (1.5.5)$$

Für nicht sehr große Zahlen \bar{n} , z.B. $\bar{n} = 100$ ist $\Delta n = 10$ und $\frac{\Delta n}{\bar{n}} = \frac{1}{10}$. Für makroskopische Systeme, z.B. $\bar{n} = 10^{20}$ ist $\Delta n = 10^{10}$ und $\frac{\Delta n}{\bar{n}} = 10^{-10}$. Die relative Schwankung wird ungeheuer klein. Für große \bar{n} ist die Verteilung w_n sehr scharf um \bar{n} konzentriert. Die Wahrscheinlichkeit, daß sich überhaupt kein Teilchen im Teilsystem befindet, d.h. $w_0 = e^{-10^{20}}$ ist verschwindend klein. Die Teilchenzahl im Untersystem $[0, a]$ ist zwar nicht fest, aber dennoch ist die relative Schwankung sehr klein für makroskopische Untersysteme.

In der untenstehenden Abbildung 1.6a) wird die Binomialverteilung für $N = 5$ und $\frac{a}{L} = \frac{3}{10}$ (und somit $\bar{n} = 1.5$) dargestellt und mit der Poisson-

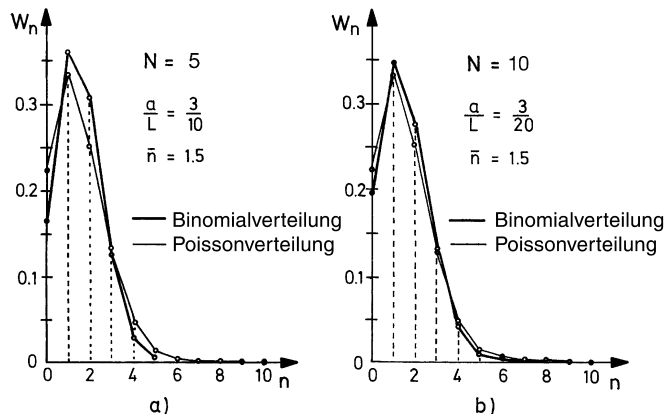


Abb. 1.6. Vergleich der Poisson-Verteilung mit der Binomialverteilung

Verteilung für $\bar{n} = 1.5$ verglichen, in b) das gleiche für $N \equiv 10$, $\frac{a}{L} = \frac{3}{20}$ (also wieder $\bar{n} = 1.5$). Selbst bei diesen kleinen N approximiert die Poisson-Verteilung die Binomialverteilung schon ziemlich gut. Schon bei $N = 100$ würden die graphischen Darstellungen der Binomial- und Poisson-Verteilung fast vollständig übereinstimmen.

***1.5.2 Gemischte Gesamtheiten und Dichtematrix von Teilsystemen**

(i) Beweise von (1.4.7a–d)

$$\text{Sp } \rho A = \sum_n \sum_i p_i \langle \psi_i | A | n \rangle \langle n | \psi_i \rangle = \sum_i p_i \langle \psi_i | A | \psi_i \rangle = \langle A \rangle .$$

Daraus folgt mit $A = 1$ auch (1.4.7b)

$$\rho^2 = \sum_i \sum_j p_i p_j |\psi_i\rangle \langle \psi_i | \psi_j\rangle \langle \psi_j| \neq \rho .$$

Für jedes $|\psi\rangle$ ist der Erwartungswert von ρ

$$\langle \psi | \rho | \psi \rangle = \sum_i p_i |\langle \psi | \psi_i \rangle|^2 \geq 0$$

nicht negativ. Da ρ hermitesch ist, sind die Eigenwerte P_m von ρ positiv reell:

$$\begin{aligned} \rho |m\rangle &= P_m |m\rangle \\ \rho &= \sum_{m=1}^{\infty} P_m |m\rangle \langle m| , \\ P_m &\geq 0, \quad \sum_{m=1}^{\infty} P_m = 1, \quad \langle m | m' \rangle = \delta_{mm'} . \end{aligned} \tag{1.5.6}$$

In dieser Basis ist $\rho^2 = \sum_m P_m^2 |m\rangle \langle m|$ und offensichtlich $\text{Sp } \rho^2 = \sum_m P_m^2 < 1$, wenn mehr als nur ein Zustand vorkommt. Man kann (1.4.7c) auch direkt aus (1.4.6) zeigen, wobei mindestens zwei verschiedene, aber nicht notwendigerweise orthogonale Zustände in (1.4.6) vorkommen müssen:

$$\begin{aligned} \text{Sp } \rho^2 &= \sum_n \sum_{i,j} p_i p_j \langle \psi_i | \psi_j \rangle \langle \psi_j | n \rangle \langle n | \psi_i \rangle \\ &= \sum_{i,j} p_i p_j |\langle \psi_i | \psi_j \rangle|^2 < \sum_i p_i \sum_j p_j = 1 . \end{aligned}$$

(ii) Das Kriterium für einen reinen oder einen gemischten Zustand ist nach Gl. (1.4.3c) und (1.4.7c) $\text{Sp } \rho^2 = 1$ bzw. $\text{Sp } \rho^2 < 1$.

(iii) Wir betrachten nun ein quantenmechanisches System, das aus zwei Teilsystemen 1 und 2 besteht. Deren Gesamtzustand sei

$$|\psi\rangle = \sum_n c_n |1n\rangle |2n\rangle, \quad (1.5.7)$$

wo mehr als ein c_n verschieden von Null sei. Die zugehörige Dichtematrix ist durch

$$\rho = |\psi\rangle \langle\psi| \quad (1.5.8)$$

gegeben. Wir führen nun Messungen durch, die nur das System 1 betreffen, d.h. die den Observablen A entsprechenden Operatoren wirken nur auf die Zustände $|1n\rangle$. Dann findet man für den Erwartungswert

$$\langle A \rangle = \text{Sp}_1 \text{Sp}_2 \rho A = \text{Sp}_1 [(\text{Sp}_2 \rho) A]. \quad (1.5.9)$$

Hier bedeutet Sp_i die Spurbildung über das Teilsystem i . Für diese Experimente ist somit die über das System 2 gemittelte Dichtematrix maßgeblich

$$\hat{\rho} = \text{Sp}_2 \rho = \sum_n |c_n|^2 |1n\rangle \langle 1n|. \quad (1.5.10)$$

Dies ist die Dichtematrix eines Gemisches, obwohl sich das Gesamtsystem in einem reinen Zustand befindet.

Der allgemeinste Zustand der Teilsysteme 1 und 2 ist von der Gestalt ⁹

$$|\psi\rangle = \sum_{n,m} c_{nm} |1n\rangle |2m\rangle. \quad (1.5.11)$$

Auch hier ist

$$\begin{aligned} \hat{\rho} \equiv \text{Sp}_2 |\psi\rangle \langle\psi| &= \sum_n \sum_{n'} \sum_m c_{nm} c_{n'm}^* |1n\rangle \langle 1n'| \\ &= \sum_m \left(\sum_n c_{nm} |1n\rangle \right) \left(\sum_{n'} c_{n'm}^* \langle 1n'| \right) \end{aligned} \quad (1.5.12)$$

im allgemeinen ein Gemisch. Da ein makroskopisches System immer mit irgendwelchen anderen Systemen in Verbindung war, befindet es sich auch in völliger Isolation niemals in einem reinen Zustand, sondern nur in einem gemischten.

Es mag auch noch instruktiv sein, den folgenden Spezialfall zu betrachten. Schreiben wir c_{nm} in der Form $c_{nm} = |c_{nm}| e^{i\varphi_{nm}}$. Falls die Phasen φ_{nm} stochastisch sind, wird aus $\hat{\rho}$ die Dichtematrix

$$\hat{\rho} = \prod_{\langle nm \rangle} \left(\int_0^{2\pi} d\frac{\varphi_{nm}}{2\pi} \right) \hat{\rho} = \sum_n \left(\sum_m |c_{nm}|^2 \right) |1n\rangle \langle 1n|.$$

⁹ Am Rande sei bemerkt, daß es möglich ist, eine biorthogonale Basis (Schmidt Basis) einzuführen, die den Zustand (1.5.11) in die Form (1.5.7) bringt, siehe QM I, Aufgabe 20.5.

Aufgaben zu Kapitel 1

1.1 Beweisen Sie die Stirlingsche Formel

$$x! \approx \sqrt{2\pi x} x^x e^{-x},$$

indem Sie von $N! = \int_0^\infty dx x^N e^{-x}$ ausgehen und den Integranden $f(x) \equiv x^N e^{-x}$ bis zur zweiten Ordnung an die Funktion $g(x) = A e^{-(x-N)^2/a^2}$ anpassen. $f(x)$ hat ein scharfes Maximum bei $x_0 = N$.

1.2 Man bestimme die Wahrscheinlichkeit $w(N, m)$ dafür, bei einem System von N Spins genau m mit der Einstellung „ \uparrow “ und dementsprechend $N - m$ mit der Einstellung „ \downarrow “ anzutreffen. Es sei kein äußeres Feld und keine Wechselwirkung unter den Spins vorhanden, so daß für jeden einzelnen Spin die Konfigurationen \uparrow und \downarrow gleich wahrscheinlich sind.

(a) Man prüfe

$$\sum_{m=0}^N w(N, m) = 1.$$

(b) Berechnen Sie den Mittelwert von m ,

$$\langle m \rangle = \sum_{m=0}^N w(N, m) m,$$

und die Schwankung $(\langle m^2 \rangle - \langle m \rangle^2)^{1/2}$.

Die dimensionslose Magnetisierung wird durch $M = 2m - N$ definiert; geben Sie dafür den Mittelwert und die Schwankung an.

(c) Berechnen Sie die Verteilung $w(N, M)$ für große N . Nehmen Sie $|M/N| \ll 1$ an.

1.3 Leiten Sie den zentralen Grenzwertsatz für $w_i(x_i)$ anstelle von $w(x_i)$ her. *Bemerkung:* Im Resultat haben Sie nur $N\langle X \rangle$ durch $\sum_i \langle X_i \rangle$ und $N(\Delta x)^2$ durch $\sum_i (\Delta x_i)^2$ zu ersetzen.

1.4 Random walk: Ein Teilchen geht bei jedem Schritt mit gleicher Wahrscheinlichkeit um die Distanz 1 nach links oder rechts.

(a) Berechnen Sie für $Y = N_+ - N_-$ $\langle Y \rangle, \langle Y^2 \rangle$ nach $N = N_+ + N_-$ Schritten exakt.

(b) Welches Resultat erhalten Sie aus dem zentralen Grenzwertsatz?

1.5 Bestätigen Sie für die Poisson-Verteilung (1.5.3b)

$$w(n) = e^{-\bar{n}} \frac{\bar{n}^n}{n!}, \quad n \geq 0 \text{ ganz,}$$

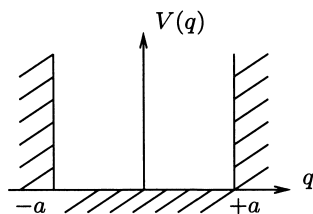
die Relationen (1.5.4a-c).

1.6 Die Verteilungsfunktion $\rho(E_1, \dots, E_N)$ habe die Form

$$\rho = \prod_{i=1}^N f(E_i).$$

Der Mittelwert und die Schwankung der einzelnen E_i seien e und $\langle (\Delta E_i)^2 \rangle^{1/2} = \Delta$. Berechnen Sie den Mittelwert und die Schwankung von $E = \sum_i E_i$.

1.7 Zeichnen Sie die Bahnkurve im Phasenraum für ein Teilchen,
(a) das sich mit Energie E in einem eindimensionalen, unendlich hohen Kastenpotential bewegt:



(b) das unter der Schwerkraft aus der Höhe h herabfällt, am Boden inelastisch reflektiert wird und wieder bis zur Höhe $9/10 h$ aufsteigt, usw.

1.8 Zeigen Sie für \mathbf{v} aus Gl. (1.3.4') und den Gradienten $\nabla H(p, q) = \left(\frac{\partial H}{\partial q}, \frac{\partial H}{\partial p} \right)$ die Relationen

$$|\mathbf{v}| = |\nabla H(p, q)| \quad \text{und} \quad \mathbf{v} \perp \nabla H(p, q).$$

1.9 Eine Ionenquelle emittiert Ionen der Energie $E_1 = 5.000 \pm 1.00$ eV von einer Fläche von 1mm^2 in einen Raumwinkel $\Omega_1 = 1$ sterad. Die Ionen werden mit elektrischen Feldern auf $E_2 = 10$ MeV beschleunigt und auf eine Fläche von 1cm^2 fokussiert. Berechnen Sie den Auffallöffnungswinkel des Ionenstrahls mit Hilfe des *Liouville-Satzes*.

Anleitung: Nehmen Sie an, daß die Beschleunigung so rasch erfolgt, daß die unterschiedlichen Geschwindigkeiten im Strahl zu keiner zusätzlichen Verbreiterung des Strahls führen, d.h. $dx_2 = dx_1$; und nehmen Sie an, daß die Breite der Energieverteilung sich nicht ändert, $dE_2 = dE_1$.

1.10 (a) Zeigen Sie $\text{Sp}(AB) = \text{Sp}(BA)$.

(b) Es seien die Operatoren ρ_ν Dichtematrizen, die also die Bedingungen (1.4.7b-d) erfüllen, und $p_\nu \geq 0$, $\sum_\nu p_\nu = 1$. Zeigen Sie, daß $\sum_\nu p_\nu \rho_\nu$ ebenfalls diese Bedingungen erfüllt.

1.11 Betrachten Sie einen in z -Richtung laufenden Lichtstrahl. Ein beliebiger reiner Polarisationszustand kann als Linearkombination geschrieben werden

$$a|\uparrow\rangle + b|\downarrow\rangle,$$

wobei $|\uparrow\rangle$ den in x -Richtung polarisierten Zustand und $|\downarrow\rangle$ den in y -Richtung polarisierten Zustand der Welle darstellen möge.

(a) Berechnen Sie die Dichtematrix: für einen allgemeinen reinen Zustand, für den in x -Richtung polarisierten Zustand, für den in 45° polarisierten Zustand, für den in 135° polarisierten Zustand.

(b) Wie sieht die Dichtematrix für einen gemischten Zustand aus, wenn z.B.: 50% längs 45° polarisiert und 50% längs 135° polarisiert oder 50% in x -Richtung polarisiert und 50% in y -Richtung polarisiert sind? Die Winkel stellen die Winkel zwischen x -Achse und Polarisationsrichtung dar.

1.12 Das Galtonsche Brett ist ein senkrecht aufgestelltes Nagelbrett, in das N horizontale, gleich lange Reihen gleichartiger Nägel eingeschlagen sind, so daß die Nägel aufeinanderfolgender Reihen genau in der Mitte zwischen denen der darüber befindlichen angeordnet sind. Oberhalb der ersten Nagelreihe in der Mitte befindet sich ein Trichter, aus dem Kugeln herausgelassen werden, die die Nagelreihen durchlaufen. Unter der letzten Reihe befinden sich Fächer, die Kugeln aufnehmen. Welche Kurve wird durch die unterschiedliche Füllung der Fächer beschrieben?

1.13 Ein Gefäß mit Volumen V enthalte N Teilchen. Betrachten Sie ein Teilvolumen v , und nehmen Sie an, daß die Wahrscheinlichkeit dafür, ein bestimmtes Teilchen in diesem Teilvolumen v zu finden, durch v/V gegeben ist.

- (a) Geben Sie die Wahrscheinlichkeit p_n an, n Teilchen in v zu finden.
- (b) Berechnen Sie den Mittelwert und das Schwankungsquadrat \bar{n} und $(n - \bar{n})^2$.
- (c) Zeigen Sie mit Hilfe der Stirlingschen Formel, daß p_n für große N und n näherungsweise durch eine Gauß-Verteilung gegeben ist.
- (d) Zeigen Sie, daß in der Grenze $\frac{v}{V} \rightarrow 0$, $V \rightarrow \infty$ bei $\frac{N}{V} = \text{const.}$ p_n in eine Poisson-Verteilung übergeht.

1.14 Gauß-Verteilung: Die Gauß-Verteilung ist definiert durch die kontinuierliche Wahrscheinlichkeitsdichte

$$w_G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-x_0)^2/2\sigma^2}.$$

Berechnen Sie für diese Verteilung $\langle X \rangle$, Δx , $\langle X^4 \rangle$ und $\langle X - \langle X^3 \rangle$.

1.15 Lognormal-Verteilung: Die Zufallsvariable X habe die Eigenschaft, daß $\log X$ einer Gauß-Verteilung genügt, mit $\langle \log X \rangle = \log x_0$.

(a) Zeigen Sie durch Transformation der Gauß-Verteilung, daß die Wahrscheinlichkeitsdichte für X die Form

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \frac{1}{x} e^{-\frac{(\log(x/x_0))^2}{2\sigma^2}}, \quad 0 < x < \infty$$

hat.

(b) Zeigen Sie

$$\langle X \rangle = x_0 e^{\sigma^2/2}$$

und

$$\langle \log X \rangle = \log x_0.$$

(c) Zeigen Sie, daß die Lognormal-Verteilung umgeschrieben werden kann in

$$P(x) = \frac{1}{x_0 \sqrt{2\pi\sigma^2}} (x/x_0)^{-1-\mu(x)}$$

mit

$$\mu(x) = \frac{1}{2\sigma^2} \log \frac{x}{x_0},$$

sie kann deshalb bei der Analyse von Daten mit einem Potenzgesetz verwechselt werden.