

KAPITEL VII

Gleichgewichtsverteilungen der Statistischen Physik

| | | |
|---------|---|----|
| VII.1 | Wahl der Wahrscheinlichkeitsverteilung | 83 |
| VII.1.1 | Thermodynamisches Gleichgewicht | 83 |
| VII.1.2 | Prinzip der maximalen Entropie | 84 |
| VII.2 | Gleichgewichtsverteilungen | 84 |
| VII.2.1 | Maximierung der von Neumann-Entropie | 85 |
| VII.2.2 | Zustandssumme | 86 |
| VII.2.3 | Gleichgewichtsentropie | 87 |
| VII.3 | Übliche statistische Ensembles | 88 |
| VII.3.1 | Mikrokanonisches Ensemble | 88 |
| VII.3.2 | Kanonisches Ensemble | 89 |
| VII.3.3 | Großkanonisches Ensemble | 90 |
| VII.3.4 | Vergleich der unterschiedlichen statistischen Ensembles | 92 |
| VII.3.5 | Gleichgewichtsverteilungen der klassischen statistischen Mechanik | 93 |

Im vorigen Kapitel wurde gezeigt, dass die Herleitung der Eigenschaften eines makroskopischen Systems aus mikroskopischen Gesetzen probabilistischer Natur sein soll. Die Frage bleibt offen, welche Wahrscheinlichkeitsverteilung unter gegebenen Umständen einzusetzen ist. In Abschn. VII.1 und VII.2 wird diese Frage allgemein beantwortet im Fall makroskopischer Systeme im thermodynamischen Gleichgewicht. Einige häufige Gleichgewichtsverteilungen werden dann in Abschn. VII.3 ausführlicher behandelt.

VII.1 Wahl der Wahrscheinlichkeitsverteilung

Der „Zustand“ eines makroskopischen Systems wird durch die Angabe einiger makroskopischer Größen bestimmt, die auch *Zustandsgrößen* oder *Zustandsvariablen* genannt werden. Genau genommen sind einige dieser makroskopischen Zustandsvariablen nur dann definiert, wenn sich das System im thermodynamischen Gleichgewicht (§ VII.1.1) befindet. In diesem Fall kann man eindeutig vom makroskopischen Zustand, kurz *Makrozustand*, des Systems sprechen. Die Wahrscheinlichkeitsverteilung für die Mikrozustände, die diesen Makrozustand bilden, wird dann durch die Bedingung bestimmt, dass die zugehörige statistische Entropie maximal ist (§ VII.1.2).

VII.1.1 Thermodynamisches Gleichgewicht

Wie im § VI.2.2 angemerkt wurde beruht der Einsatz von Wahrscheinlichkeiten in die Beschreibung eines physikalischen Systems auf der Interpretation der Erwartungswerte von Observablen als *Ensemblemittelwerte* über ein Ensemble von gleich präparierten Systemen.

Dagegen entspricht ein Messprozess in der Physik keiner Ensemblemittelung über viele Systeme, sondern einer *Zeitmittelung* über die zeitliche Reihenfolge der Mikrozustände eines einzigen Systems,

die während der Dauer der Messung realisiert sind. Somit liefert die Messung einer klassischen Observable O den *Zeitmittelwert*

$$\bar{O}_\tau \equiv \frac{1}{\tau} \int_t^{t+\tau} O(\{\mathbf{q}^a(t')\}, \{\mathbf{p}_a(t')\}) \rho(t', \{\mathbf{q}^a(t')\}, \{\mathbf{p}_a(t')\}) dt' \quad (\text{VII.1})$$

entlang der Trajektorie $\{\mathbf{q}^a(t'), \mathbf{p}_a(t')\}$ im Phasenraum des Systems, wobei die mögliche Abhängigkeit der Wahrscheinlichkeitsdichte von der Zeit berücksichtigt wurde. Beispielsweise unterliegt jedes Molekül eines Gases unter normalen Bedingungen in einer Messzeit von $1 \mu\text{s}$ ca. 10^3 Stöße, die den Impuls des Moleküls und somit den Mikrozustand des Systems ändern.

Um die zeitliche Änderung der Wahrscheinlichkeitsdichte zu vermeiden, kann man zunächst nur Systeme betrachten, deren makroskopischen Eigenschaften stationär sind, so dass das System im gleichen Makrozustand bleibt. Wenn es zusätzlich keinen (stationären) makroskopischen Strom im System gibt, dann spricht man von einem *thermodynamischen Gleichgewichtszustand* oder kurz von *thermodynamischem Gleichgewicht*.

Gemäß der *Quasi-Ergodenhypothese* werden im Lauf der spontanen Zeitentwicklung eines Systems im thermodynamischen Gleichgewicht alle erlaubten Mikrozustände beliebig nah⁽²²⁾ erreicht, und zwar mit derselben relativen Häufigkeit, vorausgesetzt die Zeitspanne der Entwicklung lang genug ist. Anders gesagt ist im Limes $\tau \rightarrow \infty$ der zeitliche Mittelwert (VII.1) für fast jede Observable gleich dem Ensemblemittelwert $\langle O \rangle$.

VII.1.2 Prinzip der maximalen Entropie

Die Wahrscheinlichkeitsverteilung der Mikrozustände für ein makroskopisches System im thermodynamischen Gleichgewicht wird durch das folgende *Prinzip der maximalen Entropie* charakterisiert [24, 25]:

Unter allen Wahrscheinlichkeitsverteilungen — Dichteoperatoren $\hat{\rho}$ bei quantenmechanischen Systemen oder Phasenraumwahrscheinlichkeitsdichten ρ bei klassischen Systemen —, die mit dem vorhandenen Wissen über das System verträglich sind, ist der eigentlich realisierte Makrozustand derjenige, dessen statistische Entropie maximal ist. (VII.2)

Dieses Prinzip stellt eine Verallgemeinerung des *Indifferenzprinzips* oder *Prinzips vom unzureichenden Grund* von Laplace^(ap) dar, laut dem den unterscheidbaren Ereignissen eines Zufallsexperiments in Abwesenheit von Information eine gleichförmige Wahrscheinlichkeitsverteilung zugeordnet werden soll.

Bemerkung: Das Prinzip der maximalen Entropie ist sehr allgemein und wird in allen Wissenschaftsbereichen benutzt — beispielsweise in der Bildverarbeitung [26, 27], der Biologie [28] oder der Linguistik [29] —, insbesondere wenn man eine stetige Verteilung aus einer endlichen Anzahl von Datenpunkten „ableiten“ möchte.

VII.2 Gleichgewichtsverteilungen

In diesem Abschnitt wird das Prinzip der maximalen Entropie angewandt, um die Gleichgewichtsverteilung im allgemeinen Fall zu bestimmen. Der Kürze halber werden die Berechnungen nur für den Fall eines quantenmechanischen Systems dargestellt — ein klassisches System bildet aber einen

⁽²²⁾Bei der noch stärkeren *Ergodenhypothese* soll die Phasenraumtrajektorie des Systems jeden erlaubten Mikrozustand genau durchlaufen.

^(ap)P.-S. LAPLACE, 1749–1827

Grenzwertfall davon. Somit wird in § VII.2.1 der statistische Operator $\hat{\rho}$ untersucht, der die maximale von Neumann-Entropie $S(\hat{\rho})$ unter allgemeinen Nebenbedingungen liefert. Der Paragraph VII.2.2 befasst sich dann mit einer wichtigen Funktion, der *Zustandssumme*, die in der Berechnung auftaucht und einen einfachen Zugriff auf messbare Größen gibt, insbesondere auf die Entropie des Gleichgewichtszustands (§ VII.2.3).

VII.2.1 Maximierung der von Neumann-Entropie

VII.2.1 a Berücksichtigung der Kenntnisse über das System

Eine erste Aufgabe, wenn man das Prinzip der maximalen Entropie anwenden will, ist das Feststellen und Ausdrücken der vorhandenen Kenntnisse über das System. Eine solche Kenntnis entspricht im Allgemeinen dem Erwartungswert einer Observable, die den Makrozustand des Systems im thermodynamischen Gleichgewicht charakterisiert. Unter diesen Observablen gibt es einerseits die Erhaltungsgrößen — Energie, Impuls, Drehimpuls, sowie in nicht-relativistischen Systemen die Teilchenzahl —, andererseits einige experimentell steuerbaren „Parameter“ wie das Volumen, ein externes (elektromagnetisches) Feld, usw.

Dazu sind die Kenntnisse über diese physikalischen Größen von zwei Arten. Einige werden *genau* oder fast genau bekannt. Beispielsweise ist die Teilchenzahl eines isolierten Systems genau bekannt, oder dessen Energie fast genau, d.h. die Energie gehört einem Intervall $[E, E + \delta E]$ mit $\delta E \ll E$ an.⁽²³⁾ Solches Wissen wird durch die Wahl des geeigneten Hilbert-Raums \mathcal{H} für die Beschreibung berücksichtigt, der als einen Unterraum des Hilbert-Raums aller möglichen Mikrozustände des Systems zu wählen ist.

Dagegen können die Kenntnisse über die physikalischen Größen statistischer Natur sein: z.B. kennt man nur den Erwartungswert der Energie $\langle E \rangle$ bzw. der Teilchenzahl $\langle N \rangle$. Für eine allgemeine Messgröße O_i , entsprechend der Observable \hat{O}_i , gibt eine solche Kenntnis eine Nebenbedingung

$$\langle O_i \rangle \equiv \langle \hat{O}_i \rangle = \text{Tr}(\hat{\rho} \hat{O}_i),$$

oder äquivalent

$$\text{Tr}(\hat{\rho} \hat{O}_i) - \langle O_i \rangle = 0. \quad (\text{VII.3a})$$

Mathematisch soll der Dichteoperator noch weitere Bedingungen erfüllen, und zwar zuerst die Normierungsbedingung $\text{Tr} \hat{\rho} = 1$ bzw.

$$\text{Tr} \hat{\rho} - 1 = 0, \quad (\text{VII.3b})$$

die die gleiche Form wie bei den „physikalischen“ Bedingungen (VII.3a) mit $\hat{O}_i = \hat{1}_{\mathcal{H}}$ annimmt. Dazu soll noch $\hat{\rho}$ hermitesch sein.

VII.2.1 b Herleitung der Gleichgewichtsverteilung

Zur Maximierung der von Neumann Entropie (VI.28) unter den Nebenbedingungen (VII.3) lässt sich die Methode der Lagrange^(aq)-Multiplikatoren (s. Anhang G) günstig einsetzen.

Dazu führt man die (Lagrange-)Hilfsfunktion

$$F(\hat{\rho}; \lambda_0, \{\lambda_i\}) \equiv -k_B \text{Tr}(\hat{\rho} \ln \hat{\rho}) - \lambda_0 k_B [\text{Tr} \hat{\rho} - 1] - \sum_i \lambda_i k_B [\text{Tr}(\hat{\rho} \hat{O}_i) - \langle \hat{O}_i \rangle] \quad (\text{VII.4})$$

mit Lagrange-Multiplikatoren λ_0 für die Normierungsbedingung (VII.3b) bzw. λ_i für jede Nebenbedingung (VII.3a) ein. Dabei wurden alle Nebenbedingungen mit k_B multipliziert.

Betrachte man jetzt eine infinitesimale Variation $\delta \hat{\rho}$ des Dichteoperators um einen Wert $\hat{\rho}_0$ des Operators, d.h. infinitesimale Variationen $\delta \rho_{jk}$ der Matrixelementen um solche von $\hat{\rho}_0$. Diese

⁽²³⁾Eine exakte Kenntnis der Energie eines Systems ist durch die Quantenmechanik verboten, weil sie gemäß der Energie-Zeit-Unschärferelation eine unendlich große Messzeit erfordert.

^(aq)J.-L. LAGRANGE, 1736–1813

Variation führt zu einer Variation der Hilfsfunktion um ihren Wert $F(\hat{\rho}_0; \lambda_0, \{\lambda_i\})$:

$$\delta F(\hat{\rho}_0; \lambda_0, \{\lambda_i\}) = -k_B \text{Tr} \left[\delta \hat{\rho} \left(\ln \hat{\rho}_0 + \hat{\mathbb{1}} + \lambda_0 \hat{\mathbb{1}} + \sum_i \lambda_i \hat{O}_i \right) \right], \quad (\text{VII.5})$$

wobei die Identität $\delta[\text{Tr} f(\hat{X})] = \text{Tr}[\delta \hat{X} f'(\hat{X})]$ für eine differenzierbare Funktion f des Operators \hat{X} benutzt wurde.

Diese Identität gilt nur dank der Invarianz der Spur unter zyklischen Transformationen! In der Tat kommutiert im Allgemeinen die Variation $\delta \hat{X}$ nicht mit \hat{X} . Die Reihenentwicklung von $f(\hat{X})$ lautet also

$$\delta f(\hat{X}) = \delta \left[\sum_n a_n \hat{X}^n \right] = \sum_n a_n \sum_{m=0}^{n-1} \hat{X}^m \delta \hat{X} \hat{X}^{n-m-1} \neq \delta \hat{X} \sum_n n a_n \hat{X}^{n-1} = \delta \hat{X} f'(\hat{X}).$$

Wenn man aber die Spur vom dritten Glied bildet, ist sie gleich der Spur des vierten.

Die Hilfsfunktion hat ein Extremum in $\hat{\rho}_0$ dann und nur dann, wenn $\delta F(\hat{\rho}_0; \lambda_0, \{\lambda_i\}) = 0$ für jede infinitesimale Variation $\delta \hat{\rho}$. Dies ist erfüllt, wenn der Term zwischen Klammern in Gl. (VII.5) verschwindet, d.h. wenn

$$\hat{\rho}_0 = \exp \left[-(\lambda_0 + 1) \hat{\mathbb{1}} - \sum_i \lambda_i \hat{O}_i \right].$$

Lässt man der Kürze halber den tiefgestellten Index von $\hat{\rho}_0$ weg und führt man die Bezeichnung $Z \equiv e^{\lambda_0 + 1}$ ein, so ergibt sich für den Dichteoperator, der die Hilfsfunktion F extremal macht

$$\hat{\rho} = \frac{1}{Z} e^{-\sum_i \lambda_i \hat{O}_i}. \quad (\text{VII.6})$$

Man muss noch zeigen, dass $\hat{\rho}$ ein Maximum ist. Da das zweifache Differential von $F(\hat{\rho})$ gleich dem von $S(\hat{\rho})$ ist, das wegen der Konkavität von S [Gl. (VI.31)] überall negativ ist, ist dies garantiert.

Bemerkung: Wenn eine der Nebenbedingungen den Erwartungswert des Hamilton-Operators $\langle \hat{H} \rangle$ bzw. der Energie betrifft, wird der zugehörige Lagrange-Parameter traditionell als β bezeichnet.

VII.2.2 Zustandssumme

Es bleibt noch, die Lagrange-Multiplikatoren λ_0 (oder äquivalent $Z \equiv e^{\lambda_0 + 1}$) und λ_i in Abhängigkeit von den Erwartungswerten $\langle \hat{O}_i \rangle$ zu bestimmen.

Die Normierungsbedingung (VII.3b) gibt

$$Z = \text{Tr} \left(e^{-\sum_i \lambda_i \hat{O}_i} \right), \quad (\text{VII.7})$$

wobei man sieht, dass Z tatsächlich von den anderen Lagrange-Multiplikatoren abhängt, was hier nach in der Schreibweise berücksichtigt wird. $Z(\{\lambda_i\})$ wird *Zustandssumme* genannt.

In der Praxis spielt die Zustandssumme eine wichtige Rolle, denn mit deren Hilfe können alle thermodynamischen Größen berechnet werden, die den Makrozustand charakterisieren. Beispielsweise lauten die in den Nebenbedingungen auftretenden Erwartungswerte

$$\langle \hat{O}_j \rangle = -\frac{\partial \ln Z}{\partial \lambda_j}(\{\lambda_i\}). \quad (\text{VII.8})$$

Dieses Ergebnis lässt sich wie folgt beweisen

$$\langle \hat{O}_j \rangle = \text{Tr}(\hat{\rho} \hat{O}_j) = \frac{1}{Z(\{\lambda_i\})} \text{Tr} \left(e^{-\sum_i \lambda_i \hat{O}_i} \hat{O}_j \right) = -\frac{1}{Z(\{\lambda_i\})} \frac{\partial}{\partial \lambda_j} \left[\text{Tr} \left(e^{-\sum_i \lambda_i \hat{O}_i} \right) \right].$$

Ähnlich liefert die zweifache Ableitung der Zustandssumme die Varianz einer Observable

$$\frac{\partial^2 \ln Z}{\partial \lambda_j^2}(\{\lambda_i\}) = \langle \hat{O}_j^2 \rangle - \langle \hat{O}_j \rangle^2, \quad (\text{VII.9})$$

oder allgemeiner die Kovarianz zweier kommutierenden Observablen

$$\frac{\partial^2 \ln Z}{\partial \lambda_j \partial \lambda_k}(\{\lambda_i\}) = \langle \hat{O}_j \hat{O}_k \rangle - \langle \hat{O}_j \rangle \langle \hat{O}_k \rangle. \quad (\text{VII.10})$$

Bemerkungen:

* Um die Lagrange-Multiplikatoren $\{\lambda_i\}$ zu erhalten, muss man die Gl. (VII.8) für alle j als ein System von Gleichungen betrachten und nach den $\{\lambda_i\}$ lösen. In der Praxis ist es nicht nötig, und ein Makrozustand wird durch die Angabe der Parameter $\{\lambda_i\}$ anstatt der Erwartungswerte $\langle O_i \rangle$ charakterisiert. Tatsächlich kann man den bisher als Hilfsparametern eingeführten $\{\lambda_i\}$ physikalische Deutungen zuordnen.

* Die Zustandssumme (VII.7) hängt nicht nur von den Parametern $\{\lambda_i\}$ ab, sondern auch implizit von den genau bekannten Größen, durch die Definition der Spur im Hilbert-Raum \mathcal{H} und durch die Form der Operatoren \hat{O}_i .

Faktorisierung der Zustandssummen unabhängiger Systeme

Sei jetzt ein System Σ bestehend aus nicht-wechselwirkenden Teilsystemen Σ_m . Der Hilbert-Raum \mathcal{H} für Σ lässt sich dann in das direkte Produkt $\mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_m \otimes \cdots$ von Teilräumen \mathcal{H}_m zerlegen, während sich der Hamilton-Operator \hat{H} als eine direkte Summe von Hamilton-Operatoren $\hat{H}^{(m)}$ für die individuellen Teilsysteme umschreiben lässt, $\hat{H} = \hat{H}^{(1)} \oplus \cdots \oplus \hat{H}^{(m)} \oplus \cdots$. Ähnlich können andere Observablen für das gesamte System als direkte Summen von Observablen für die Teilsysteme umgeschrieben werden:⁽²⁴⁾

$$\hat{O}_i = \bigoplus_m \hat{O}_i^{(m)}.$$

Die Zustandssumme für Σ lautet dann

$$Z(\{\lambda_i^{(m)}\}) = \text{Tr} \left[\exp \left(- \bigoplus_m \sum_i \lambda_i^{(m)} \hat{O}_i^{(m)} \right) \right] = \text{Tr} \left(\bigotimes_m e^{-\sum_i \lambda_i^{(m)} \hat{O}_i^{(m)}} \right).$$

Mit der in § VI.3.3 eingeführten reduzierten Spur Tr_{Σ_m} lässt sich diese Spur eines Tensorprodukts von Operatoren als das Produkt von reduzierten Spuren auf den jeweiligen Teilräumen umschreiben:

$$Z(\{\lambda_i^{(m)}\}) = \prod_m Z^{(m)}(\{\lambda_i^{(m)}\}) \quad \text{mit} \quad Z^{(m)}(\{\lambda_i^{(m)}\}) = \text{Tr}_{\Sigma_m} \left(e^{-\sum_i \lambda_i^{(m)} \hat{O}_i^{(m)}} \right). \quad (\text{VII.11})$$

Ebenso faktorisiert sich der Dichteoperator $\hat{\rho}$ für das gesamte System als Tensorprodukt von Dichteoperatoren $\hat{\rho}^{(m)}$ für die Teilsysteme.

Gemäß Gl. (VII.11) addieren sich die $\ln Z^{(m)}$ der Teilsysteme, d.h. man kann erwarten, dass $\ln Z$ sowie ihre Ableitungen nach den Lagrange-Parametern extensive Größen sind.

⁽²⁴⁾Genauer sollte man $\hat{O}_i = \hat{\mathbb{1}}_{\Sigma_1} \otimes \cdots \otimes \hat{\mathbb{1}}_{\Sigma_{m-1}} \otimes \hat{O}_i^{(m)} \otimes \hat{\mathbb{1}}_{\Sigma_{m+1}} \otimes \cdots$ mit $\hat{\mathbb{1}}_{\Sigma_{m'}}$ der Identität auf $\Sigma_{m'}$ schreiben.

ANHANG G

Methode der Lagrange-Multiplikatoren

Seien eine reelle Funktion f von M reellen Variablen x_1, \dots, x_M und $N < M$ unabhängige Nebenbedingungen

$$g_n(x_1, \dots, x_M) = 0 \quad \text{für } n = 1, \dots, N \quad (\text{G.1})$$

mit reellen Funktionen g_n . Dabei wird angenommen, dass die Funktionen f und $\{g_n\}$ differenzierbar sind. Das Ziel ist, die Funktion f unter diesen Nebenbedingungen zu optimieren.

Sei eine neue Funktion (*Lagrange-Hilfsfunktion*⁽³¹⁾)

$$F(x_1, \dots, x_M; \lambda_1, \dots, \lambda_N) \equiv f(x_1, \dots, x_M) - \sum_{n=1}^N \lambda_n g_n(x_1, \dots, x_M) \quad (\text{G.2})$$

mit N reellen Parametern $\lambda_1, \dots, \lambda_N$, den sog. *Lagrange-Multiplikatoren* oder *Lagrange-Parametern*, deren Festlegung hiernach diskutiert wird. Dann können die Extrema von f unter den Nebenbedingungen (G.1) durch die folgende Methode erhalten werden:

1. Man sucht die Stellen der Extrema von F als Funktion der Variablen x_1, \dots, x_M unter der Annahme, dass diese Variablen unabhängig voneinander sind. Somit löst man das System von M gekoppelten Gleichungen

$$\frac{\partial F}{\partial x_m}(x_1, \dots, x_M; \lambda_1, \dots, \lambda_N) = 0 \quad \text{für } m = 1, \dots, M \quad (\text{G.3})$$

nach den $\{x_m\}$. Eine Lösung (x_1^*, \dots, x_M^*) des Systems besteht aus Funktionen $x_m^*(\lambda_1, \dots, \lambda_N)$ der Lagrange-Multiplikatoren.

2. Die Werte dieser Multiplikatoren werden bestimmt, indem die Lösung (x_1^*, \dots, x_M^*) in die Nebenbedingungen (G.1) eingesetzt wird. Man erhält somit N Gleichungen für die N Multiplikatoren.
3. Nach dem Lösen dieser Gleichungen wird das Ergebnis in die $\{x_m^*\}$ eingesetzt, die dann ein Extremum von F bilden.

⁽³¹⁾ F wird auch „Lagrange-Funktion“ genannt, wie die Funktion $L(t, \{q^a\}, \{\dot{q}^a\})$ des Lagrange-Formalismus der analytischen Mechanik.