
Entropie in der statistischen Thermodynamik

*Dr. habil. Till Biskup**

Februar 2024

Zusammenfassung

Die folgende Darstellung greift nur einen sehr kleinen Teilaspekt der statistischen Thermodynamik heraus, namentlich die Definition und statistische Interpretation der Entropie nach Ludwig Boltzmann. Ausgehend von der klassischen, phänomenologischen Thermodynamik wird deren Definition der Entropie die Definition von Boltzmann gegenübergestellt und von dort ausgehend die relevanten Konzepte eingeführt. Die Kernthese ist dabei, dass Boltzmanns statistische bzw. wahrscheinlichkeitstheoretische Formulierung der Entropie wesentlich zu deren Verständnis beiträgt. Abschließend wird mit einem kurzen Ausblick auf Shannons Informationstheorie und seiner Definition der Information, die mathematisch äquivalent zu Boltzmanns Definition der Entropie ist, eine Verbindung zwischen Entropie und Informationsgehalt eines Systems hergestellt.

Inhaltsverzeichnis

1 Konfigurationen und Gewichte	3
2 Wahrscheinlichste Konfiguration: Boltzmann-Verteilung	5
3 Molekulare Zustandssumme	9
4 Statistische Interpretation der Entropie	11
5 Entropie und Shannon-Information	12

Einleitung

Die Entropie gilt weithin als das am Schwierigsten zu fassende Konzept der Thermodynamik. Als Begriff geprägt und in die Thermodynamik eingeführt wurde sie von Rudolf Clausius [1] im Kontext des zweiten Hauptsatzes der Thermodynamik. Bereits Clausius formuliert explizit zwei wesentliche Aspekte:

- Die Entropie kann niemals abnehmen, sondern nur zunehmen, die Entropie des gesamten Universums strebt also einem Maximum zu.
- Wenn dieser (Gleichgewichts-)Zustand einmal erreicht ist, ist das Universum tot, weil keine Veränderung mehr stattfindet.

* lehre@till-biskup.de

Zentrale Aspekte

- ☞ Die Entropie gibt die Richtung von Prozessen vor und führt das Konzept der Irreversibilität (und Zeit) ein.
- ☞ Die statistische Thermodynamik verknüpft Quantenmechanik und makroskopische thermodynamische Eigenschaften.
- ☞ Die Boltzmannverteilung gibt die wahrscheinlichste und weitaus dominierende Konfiguration eines Systems an.
- ☞ Boltzmanns statistische Formulierung der Entropie macht diese Größe anschaulich(er) und greifbar(er).
- ☞ Shannons Informationstheorie verknüpft die Entropie mit der über ein System verfügbaren Information.

Im Kontrast zum ersten Hauptsatz der Thermodynamik, der die Bedingungen des Gleichgewichts aufstellt (Energieerhaltung), formuliert der zweite Hauptsatz der Thermodynamik das Prinzip der Irreversibilität und definiert damit in letzter Konsequenz die Richtung des Zeitpfeils.

Ausgangspunkt der Betrachtung ist die direkte Gegenüberstellung der beiden Entropie-Definitionen in der klassischen und statistischen Thermodynamik.

Die Entropie S in der Definition der klassischen Thermodynamik, eingeführt von Clausius, ergibt sich zu:

$$dS = \frac{dq_{\text{rev}}}{T} \quad (1)$$

mit der reversiblen Änderung der inneren Energie dq_{rev} und der Temperatur T . Die Einheit der Entropie ergibt sich also zu $[S] = \text{J K}^{-1}$.

Ludwig Boltzmann gelingt in seinem Beitrag 1876 [2] die Verknüpfung der Entropie mit der Wahrscheinlichkeit und damit die Herleitung des Entropiebegriffs aus der statistischen Physik/Thermodynamik. Die berühmte Formulierung¹ geht allerdings nicht auf ihn, sondern auf Max Planck [3] zurück:

$$S = k_{\text{B}} \ln W \quad (2)$$

Hier ist k_{B} die Boltzmannkonstante und W das (statistische) Gewicht der wahrscheinlichsten Konfiguration eines Systems. Da der Logarithmus einer Größe immer einheitenlos ist, muss folglich die Einheit der Boltzmannkonstante mit jener der Entropie identisch sein, es gilt $[k_{\text{B}}] = \text{J K}^{-1}$.

Der entscheidende Unterschied zwischen den Betrachtungen von Clausius und Boltzmann ist, dass Clausius sich ein (makroskopisches) System anschaut und quantitativ messbare Größen ableitet, während Boltzmann aus der statistischen Physik kommend einzelne Teilchen betrachtet und messbare Größen als statistischen Mittelwert über die Werte für einzelne Teilchen interpretiert.

¹ Teil der Berühmtheit könnte darin begründet sein, dass sich die Formel in dieser Form auf dem Grabstein Boltzmanns in Wien findet.

Vereinfachend lässt sich also feststellen, dass Boltzmann die Entropie S als Funktion einer (thermodynamischen) Wahrscheinlichkeit W auffasst:

$$S = f(W)$$

Aus der klassischen Thermodynamik ist bekannt, dass die Entropie eine extensive Zustandsgröße ist. Das bedeutet insbesondere, dass sich die Entropien unterschiedlicher Systeme addieren:

$$S_{\text{ges}} = S_1 + S_2 + \dots = \sum_i S_i$$

Aus der Kombinatorik wiederum ist bekannt, dass sich für unabhängige Ereignisse die Einzelwahrscheinlichkeiten zu einer Gesamtwahrscheinlichkeit multiplizieren:

$$P_{\text{ges}} = P_1 \cdot P_2 \cdot \dots = \prod_i P_i$$

Sollte also, wie von Boltzmann postuliert, die Entropie tatsächlich mit einer Wahrscheinlichkeit verknüpft sein, suchen wir eine mathematische Funktion, bei der sich das Produkt der Argumente als Summe der Funktionswerte der Einzelargumente schreiben lässt:

$$f(a \cdot b) = f(a) + f(b)$$

Eine solche mathematische Funktion ist der Logarithmus:

$$\log(a \cdot b) = \log(a) + \log(b)$$

Das erklärt bereits, warum in Boltzmanns statistischer Formulierung der Entropie der Logarithmus auftaucht. Eine weitere Begründung wird später noch gegeben. Zunächst soll es erst einmal darum gehen herzuleiten, was genau hinter der Variablen W in Boltzmanns Formel steckt, dem (statistischen) Gewicht der wahrscheinlichsten Konfiguration eines Systems.

1 Konfigurationen und Gewichte

Als statistisches Gewicht wird die Zahl der möglichen Realisierungen für eine Konfiguration bezeichnet. Eine Konfiguration oder Realisierungsmöglichkeit ist dabei die Summe der Besetzungszahlen (Populationen) N_i für alle möglichen Zustände eines Systems aus N Teilchen. Jedem Zustand ist dabei die Energie ε_i zugeordnet. Eine grafische Darstellung unterschiedlicher Konfigurationen für ein einfaches System aus vier nicht unterscheidbaren Kugeln ist in Abb. 1 wiedergegeben.

Zwei wesentliche Voraussetzungen/Annahmen der nachfolgenden Betrachtung sind:

1. Die Teilchen des Systems sind voneinander unabhängig.

Die innere Energie des Systems ist folglich die Summe der Energien der einzelnen Teilchen.

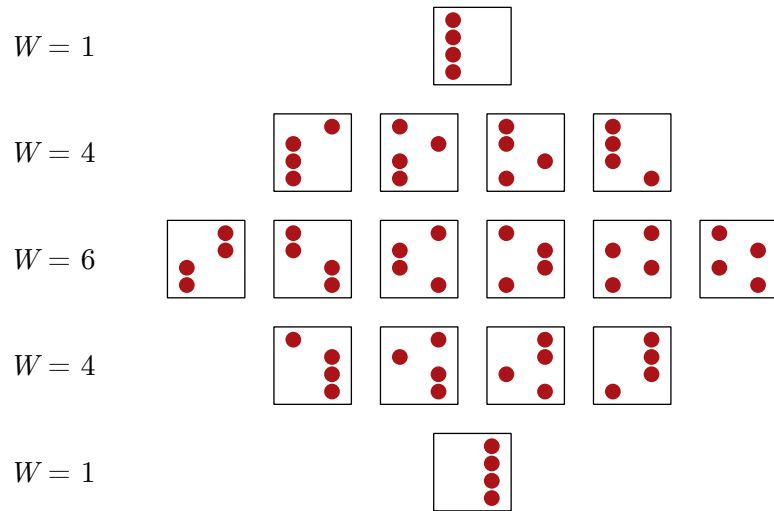


Abbildung 1: Grafische Veranschaulichung der Konfigurationen eines Systems. Dargestellt sind die Realisierungsmöglichkeiten (Konfigurationen) eines Systems mit vier Kugeln, die auf jeweils einer von zwei Seiten eines Behälters angeordnet sein können. Die Zeilen sind nach der Gesamtzahl der Kugeln auf je einer Seite sortiert. W gibt dabei die Anzahl der Möglichkeiten in der jeweiligen Kategorie, d.h. das Gewicht der Konfiguration, an.

2. Alle Realisierungsmöglichkeiten (Konfigurationen) besitzen die gleiche Wahrscheinlichkeit.

Für das Gewicht W einer Konfiguration ergibt sich damit unter Berücksichtigung der Tatsache, dass sich die Teilchen *nicht* unterscheiden lassen (daher der Nenner):

Anmerkung: $0! = 1$

$$W = \frac{N!}{N_0!N_1!N_2! \dots} \quad (3)$$

Aus den oben aufgeführten Forderungen der klassischen Thermodynamik, dass die Entropie eine extensive Größe (und damit additiv) sei und der Kombinatorik, dass die Wahrscheinlichkeiten unabhängiger Ereignisse multiplikativ sind, ergibt sich die Verwendung des Logarithmus. Logarithmieren von Gl. (3) ergibt entsprechend:

$$\begin{aligned} \ln W &= \ln \left(\frac{N!}{N_0!N_1!N_2! \dots} \right) \\ &\quad \downarrow \ln(x/y) = \ln x - \ln y \\ &= \ln N! - \ln(N_0!N_1!N_2! \dots) \\ &\quad \downarrow \ln(xy) = \ln x + \ln y \\ &= \ln N! - (\ln N_0! + \ln N_1! + \ln N_2! + \dots) \\ &= \ln N! - \sum_i \ln N_i! \end{aligned} \quad (4)$$

Ein Problem, mit dem wir hier konfrontiert sind, sind die Fakultäten in obiger Formel. Für große Zahlen – und realistische i sind, da es hier um Teilchen geht, typischerweise im Bereich der Avogadro-Konstanten

Stirlingformel

$N_A \approx 10^{23}$ – nehmen Fakultäten unhandlich große Zahlenwerte an, was die Rechnung mit ihnen für alle praktischen Belange unmöglich macht.

Als Näherungsformel für die Fakultät $x!$ einer Zahl x lässt sich die **Stirlingformel** in der Form

$$\ln x! \approx x \ln x - x \quad (5)$$

einsetzen. Es sei zumindest erwähnt, dass es sich hierbei um eine vereinfachte Form der Stirlingformel handelt, deren Abweichung vom korrekten Ergebnis aber aufgrund der betrachteten großen Werte für x nicht ins Gewicht fällt. Damit ergibt sich für das Gewicht einer Konfiguration:

$$\begin{aligned} \ln W &= (N \ln N - N) - \sum_i (N_i \ln N_i - N_i) \\ &= N \ln N - N - \sum_i (N_i \ln N_i) + \sum_i N_i \\ &\quad \downarrow \sum_i N_i = N \\ &= N \ln N - \sum_i N_i \ln N_i \end{aligned} \quad (6)$$

Die Anwendbarkeit der Stirlingformel ist eine weitere Begründung für die Verwendung des Logarithmus.

2 Wahrscheinlichste Konfiguration: Boltzmann-Verteilung

Wir hatten bei der Einführung von Boltzmanns statistischer Definition der Entropie festgehalten, dass das W das statistische Gewicht der wahrscheinlichsten Konfiguration sei. Für Systeme mit vielen Teilchen lässt sich zeigen, dass die wahrscheinlichste Konfiguration gleichzeitig die anderen Konfigurationen dominiert, weshalb sie auch als dominierende Konfiguration bezeichnet wird.



Definition: dominierende Konfiguration

Konfiguration, deren Gewicht so überwältigend groß ist, dass das System quasi immer in ihr vorliegt. \square

Die **Boltzmann-Verteilung** ermöglicht uns, die Populationen N_i in der Konfiguration mit dem größten Gewicht, in anderen Worten der wahrscheinlichsten oder dominierenden Konfiguration, zu berechnen.

$$\frac{N_i}{N} = \frac{\exp(-\beta \varepsilon_i)}{\sum_j \exp(-\beta \varepsilon_j)} \quad \text{mit} \quad \beta = \frac{1}{k_B T} \quad (7)$$

Zwei Annahmen sind wesentlich für die Boltzmann-Verteilung und ihre Herleitung – und ebenso für reale Systeme:

1. konstante Gesamtenergie E

$$\sum_i n_i \varepsilon_i = E \quad \sum_i \varepsilon_i dN_i = 0 \quad (8)$$

2. konstante Teilchenzahl N

$$\sum_i N_i = N \qquad \sum_i dN_i = 0 \qquad (9)$$

Wie aber bestimmt man nun die dominierende Konfiguration bzw. deren Gewicht? Es handelt sich um ein klassisches Optimierungsproblem: gesucht sind die N_i , für die W bzw. $\ln W$ maximal wird. Entsprechend variiert man N_i solange, bis sich keine Änderung in $\ln W$ mehr ergibt, also gilt: $d \ln W = 0$. Mathematisch lässt sich die Variation der Konfiguration wie folgt formulieren:

$$N_i \rightarrow N_i + dN_i \qquad \ln W \rightarrow \ln W + d \ln W$$

Für $d \ln W$ ergibt sich dabei:

$$d \ln W = \sum_i \left(\frac{\partial \ln W}{\partial N_i} \right) dN_i$$

Das ist gleichbedeutend damit, dass die Änderung von $\ln W$ die Summe der Beiträge ist, die von Änderungen für jeden Wert von N_i herrühren.

Die triviale Lösung, einfach den Term in Klammern Null zu setzen, kann durch die Forderung verhindert werden, die beiden Randbedingungen aus Gl. (8) und (9) bei der Optimierung von $\ln W$ mit zu berücksichtigen. Auf diesem Weg sind die dN_i nicht mehr unabhängig voneinander.

Um die Gleichung unter den gegebenen Randbedingungen zu lösen, bedient man sich der von Lagrange eingeführten Methode der unbestimmten Multiplikatoren. Ohne hier auf die Details der Methode einzugehen, sei angemerkt, dass die Randbedingungen jeweils mit einer Konstante multipliziert und anschließend zur zu variierenden Gleichung addiert werden. Die beiden Konstanten sind in unserem Fall α und $-\beta$, die Wahl des negativen Vorzeichens für β erklärt sich rückblickend und ist eine Frage der Bequemlichkeit. Entscheidend für das weitere Vorgehen ist, dass die Variablen nun alle als unabhängig voneinander betrachtet werden. Die Konstanten werden am Ende der Rechnung evaluiert.

$$\begin{aligned} d \ln W &= \sum_i \left(\frac{\partial \ln W}{\partial N_i} \right) dN_i + \alpha \sum_i dN_i - \beta \sum_i \varepsilon_i dN_i \\ &= \sum_i \left\{ \left(\frac{\partial \ln W}{\partial N_i} \right) + \alpha - \beta \varepsilon_i \right\} dN_i \end{aligned}$$

Da nun alle N_i als unabhängig voneinander betrachtet werden, gibt es nur noch eine Möglichkeit, wie die Bedingung $d \ln W = 0$ erfüllt werden kann. Es muss für alle i gelten, dass:

$$\frac{\partial \ln W}{\partial N_i} + \alpha - \beta \varepsilon_i = 0 \qquad (10)$$

Für $\ln W$ hatten wir in Gl. (6) erhalten:

$$\ln W = N \ln N - \sum_i N_i \ln N_i$$

Die Ableitung nach N_i ergibt folglich:

$$\frac{\partial \ln W}{\partial N_i} = \frac{\partial(N \ln N)}{\partial N_i} - \sum_j \frac{\partial(N_j \ln N_j)}{\partial N_i} \quad (11)$$

Man beachte hier den Wechsel der Indices in der Summe des zweiten Terms, da der Summationsindex nicht mit dem Index i der N_i , nach denen abgeleitet wird, übereinstimmt. Beide Terme auf der rechten Seite der Gleichung können nun zunächst unabhängig voneinander betrachtet werden.

Produktregel:

$$(u \cdot v)' = u' \cdot v + u \cdot v'$$

Durch Anwendung der Produktregel lässt sich der erste Term der Gl. (11) schreiben als:

$$\frac{\partial(N \ln N)}{\partial N_i} = \left(\frac{\partial N}{\partial N_i} \right) \ln N + N \left(\frac{\partial \ln N}{\partial N_i} \right)$$

Aus der Definition von N ergibt sich für die Ableitung von N nach N_i :

$$N = N_1 + N_2 + \dots + N_i = \sum_i N_i \quad \frac{\partial N}{\partial N_i} = 1$$

Des Weiteren ergibt die Ableitung des Logarithmus für den Term in der ganz rechten Klammer:

$$\frac{\partial \ln N}{\partial N_i} = \frac{1}{N} \frac{\partial N}{\partial N_i}$$

In aller Ausführlichkeit:

$$\begin{aligned} \frac{\partial(N \ln N)}{\partial N_i} &= \underbrace{\left(\frac{\partial N}{\partial N_i} \right)}_{=1} \ln N + N \left(\frac{\partial \ln N}{\partial N_i} \right) \\ &= \ln N + N \frac{1}{N} \underbrace{\frac{\partial N}{\partial N_i}}_{=1} \\ &= \ln N + 1 \end{aligned}$$

Für den zweiten Term von Gl (11) ergibt sich, wieder durch Ableitung des Logarithmus:

$$\frac{\partial \ln N_j}{\partial N_i} = \frac{1}{N_j} \left(\frac{\partial N_j}{\partial N_i} \right)$$

Wenn gilt, dass $i \neq j$, sind alle N_j unabhängig von N_i , und die (partiellen) Ableitungen verschwinden: $\partial N_j / \partial N_i = 0$. Im umgekehrten Fall, dass $i = j$, ergibt sich:

$$\frac{\partial N_j}{\partial N_i} = \frac{\partial N_j}{\partial N_j} = 1$$

und damit:

$$\frac{\partial N_j}{\partial N_i} = \delta_{ij} \quad \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

mit dem Kronecker-Delta δ_{ij} . Entsprechende erneute Anwendung der Produktregel und Ausnutzung der oben gezeigten Beziehungen ergibt:

$$\begin{aligned} \sum_j \frac{\partial(N_j \ln N_j)}{\partial N_i} &= \sum_j \left\{ \left(\frac{\partial N_j}{\partial N_i} \right) \ln N_j + N_j \left(\frac{\partial \ln N_j}{\partial N_i} \right) \right\} \\ &= \sum_j \left\{ \left(\frac{\partial N_j}{\partial N_i} \right) \ln N_j + \underbrace{\left(\frac{\partial N_j}{\partial N_i} \right)}_{=1} \right\} \\ &= \sum_j \underbrace{\left(\frac{\partial N_j}{\partial N_i} \right)}_{=\delta_{ij}} (\ln N_j + 1) \\ &= \sum_j \delta_{ij} (\ln N_j + 1) = \ln N_i + 1 \end{aligned}$$

Die Ergebnisse für die beiden Terme lassen sich nun in Gl. (11) einsetzen, und wir erhalten (unter Vertauschen der Summanden und Anwendung der Logarithmengesetze):

$$\frac{\partial \ln W}{\partial N_i} = -(\ln N_i + 1) + (\ln N + 1) = -\ln \frac{N_i}{N}$$

Setzen wir dieses Ergebnis jetzt in die oben aufgestellte Gleichung (10) für die Optimierungsbedingung von $d \ln W$ ein, erhalten wir:

$$-\ln \frac{N_i}{N} + \alpha - \beta \varepsilon_i = 0$$

Umstellen und anschließendes Exponentieren beider Seiten ergibt entsprechend:

$$\frac{N_i}{N} = \exp(\alpha - \beta \varepsilon_i)$$

Da N gerade der Summe aller N_i entspricht – die Bedingung konstanter Teilchenzahl, Gl. (9) – können wir entsprechend umschreiben. Dabei wird die obige Gleichung durch Multiplikation mit N nach N_i umgestellt und eingesetzt:

$$N = \sum_i N_i = \sum_i N \exp(\alpha - \beta \varepsilon_i) = N \exp(\alpha) \sum_i \exp(\beta \varepsilon_i)$$

Da sich N auf beiden Seiten dieser Gleichung gerade aufhebt, ergibt sich:

$$\exp(\alpha) = \frac{1}{\sum_i \exp(-\beta \varepsilon_i)}$$

Diese Gleichung können wir nun wieder in die darüber stehende einsetzen

– und stellen gleichzeitig fest, dass wir elegant darum herum gekommen sind, jemals die Konstante α zu bestimmen:

$$\begin{aligned}\frac{N_i}{N} &= \exp(\alpha - \beta\varepsilon_i) = \exp(\alpha) \exp(-\beta\varepsilon_i) = \frac{1}{\sum_j \exp(-\beta\varepsilon_j)} \exp(-\beta\varepsilon_i) \\ &= \frac{\exp(-\beta\varepsilon_i)}{\sum_j \exp(-\beta\varepsilon_j)}\end{aligned}$$

Das aber entspricht genau der Form der Boltzmann-Verteilung, wie sie in Gl. (7) angegeben ist. Jetzt wird auch die Wahl der Konstanten und ihrer Vorzeichen klar. Die Identifikation des Parameters β mit der (thermodynamischen) Temperatur in der Form $\beta = 1/(k_{\text{B}}T)$ lässt sich z.B. über die Formulierung der inneren Energie eines einatomigen idealen Gases mit Hilfe von dessen molekularer Zustandssumme begründen.²

3 Molekulare Zustandssumme

Die Boltzmannverteilung, Gl. (7), kann auch geschrieben werden als:

$$p_i = \frac{\exp(-\beta\varepsilon_i)}{q} \qquad p_i = \frac{N_i}{N} \qquad (12)$$

hierbei ist p_i der Bruchteil aller Moleküle, die sich im Zustand mit der Energie ε_i befinden, und q die molekulare Zustandssumme. Sie gibt die Zahl der thermisch zugänglichen Zustände einer Ansammlung von Molekülen bei einer Temperatur T an:

$$q = \sum_i \exp(-\beta\varepsilon_i) \quad \text{mit} \quad \beta = \frac{1}{k_{\text{B}}T} \qquad (13)$$

Hier ist k_{B} die Boltzmannkonstante und T die thermodynamische Temperatur. Im Falle der Entartung der Energieniveaus kann q auch geschrieben werden als:

$$q = \sum_{\text{Niveaus } I} g_I \exp(-\beta\varepsilon_I)$$

Für eine Interpretation der molekularen Zustandssumme ist es hilfreich, sich ihr Temperaturverhalten anzuschauen. Im Allgemeinen gibt es unbegrenzt viele Energieniveaus, was bedeutet, dass für beliebig hohe Temperaturen alle Energieniveaus thermisch zugänglich werden:

$$\lim_{T \rightarrow \infty} q = \infty \qquad \lim_{T \rightarrow \infty} \frac{1}{k_{\text{B}}T} = 0 \quad \exp(0) = 1$$

Die mathematische Begründung dafür ist im rechten Teil der vorangegangenen Gleichung angegeben: Für unbegrenzt hohe Temperaturen verschwindet der Exponent, weshalb in der Summe lauter Einsen stehen, die über die unbegrenzte Zahl der Energieniveaus dann ebenfalls unbegrenzt ist.

Die Betrachtung des umgekehrten Falls ist ebenso instruktiv. Für ver-

²Der geneigte Leser sei auf die einschlägige Literatur, z.B. [4, S. 607], verwiesen.

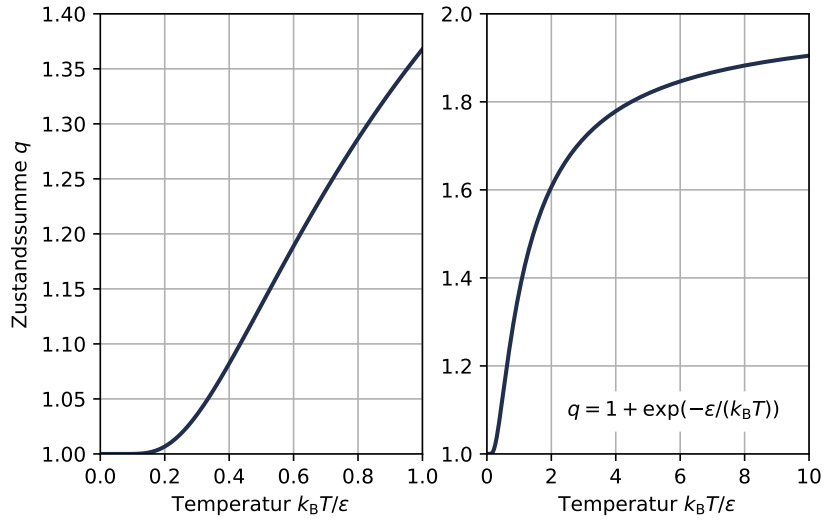


Abbildung 2: Temperaturabhängigkeit der Zustandssumme q eines Zweiniveau-Systems. Für sehr kleine Temperaturen ist die Zustandssumme $q \approx 1$ und strebt gegen beliebig hohe Temperaturen gegen $q = 2$. Konkret bedeutet das, dass für niedrige Temperaturen (fast) ausschließlich das untere Energieniveau thermisch zugänglich und besetzt ist, während für sehr hohe Temperaturen beide Energieniveaus zugänglich und quasi gleich besetzt sind. Die Temperaturabhängigkeit der Besetzungszahlen ist in Abb. 3 gezeigt.

schwindende Temperaturen wird die Zustandssumme $q = 1$, bzw. im Falle der Entartung des untersten Energieniveaus identisch mit dem Grad der Entartung:

$$\lim_{T \rightarrow 0} q = 1 \quad \begin{array}{l} i > 0 : \quad \lim_{T \rightarrow 0} \frac{1}{k_B T} = \infty \quad \exp(-\infty) = 0 \\ i = 0 : \quad \varepsilon_0 \equiv 0 \quad \frac{\varepsilon_0}{k_B T} \equiv 0 \quad \exp(0) = 1 \end{array}$$

Auch hier ist die mathematische Begründung wieder im rechten Teil angegeben, unterschieden nach den beiden Fällen $i > 0$ und $i = 0$. Für $i > 0$ geht β im Grenzfall gegen Unendlich, die Exponentialfunktion entsprechend gegen Null, und alle Summanden für $i > 0$ verschwinden. Im Falle von $i = 0$ ist $\varepsilon_0 \equiv 0$ und ebenso der Exponent, als Summand bleibt die Eins übrig.

Betrachten wir im Folgenden das einfachste denkbare System mit nur zwei Energieniveaus ε_0 und ε_1 mit einem Energieabstand zwischen den beiden Niveaus von ε :

$$\varepsilon_0 = 0 \qquad \varepsilon_1 = \varepsilon$$

Die Zustandssumme q für dieses Zweiniveau-System lässt sich direkt aufschreiben:

$$q = \sum_i \exp(-\beta \varepsilon_i) = \underbrace{\exp(-\beta \cdot 0)}_1 + \exp(-\beta \varepsilon) = 1 + \exp(-\beta \varepsilon)$$

Der Temperaturverlauf dieser Zustandssumme ist in Abb. 2 gezeigt. Für

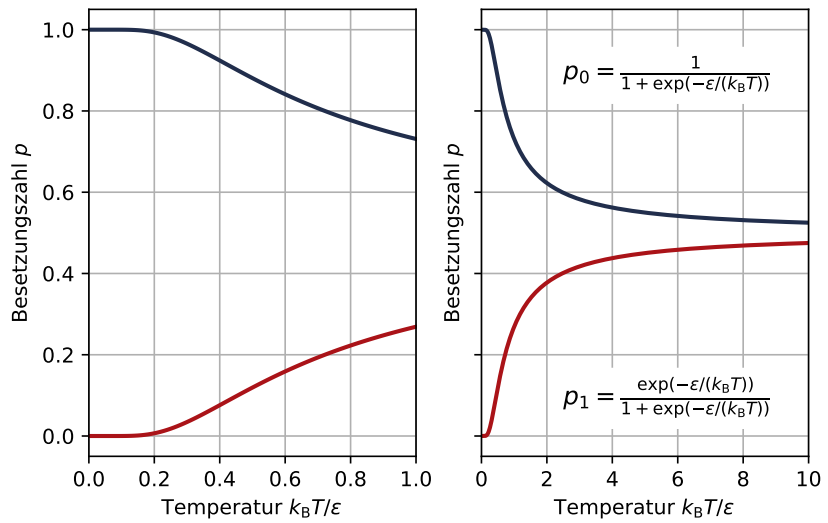


Abbildung 3: Temperaturabhängigkeit der Besetzungszahlen (Populationen) p_i eines Zweiniveau-Systems mit $i \in \{1, 2\}$. Für sehr kleine Temperaturen ist (fast) ausschließlich das untere Energieniveau thermisch zugänglich und besetzt, während für sehr hohe Temperaturen beide Energieniveaus zugänglich und quasi gleich besetzt sind.

sehr kleine Temperaturen ist die Zustandssumme $q \approx 1$ und strebt gegen beliebig hohe Temperaturen gegen $q = 2$. Konkret bedeutet das, dass für niedrige Temperaturen (fast) ausschließlich das untere Energieniveau thermisch zugänglich und besetzt ist, während für sehr hohe Temperaturen beide Energieniveaus zugänglich und quasi gleich besetzt sind. Die Temperaturabhängigkeit der Besetzungszahlen ist in Abb. 3 gezeigt.

Eine entscheidende Erkenntnis, die auch grafisch in Abb. 3 eindrucksvoll dargestellt ist: Für beliebig hohe Temperaturen nähern sich die Besetzungszahlen (Populationen) aller Energieniveaus einander an. Eine Besetzungsinversion – also eine größere Population eines energetisch höher liegenden Energieniveaus als die eines energetisch tiefer liegenden – ist folglich nicht auf thermischem Weg zu erreichen.

Die molekulare Zustandssumme enthält die gesamte Information über die thermodynamischen Eigenschaften eines Systems aus unabhängigen Molekülen im thermischen Gleichgewicht. Insofern ist ihre Rolle vergleichbar jener der Wellenfunktion in der Quantenmechanik: q ist eine Art thermische Wellenfunktion.

4 Statistische Interpretation der Entropie

Kehren wir nun noch einmal zu Boltzmanns Definition der Entropie zurück:

$$S = k_B \ln W$$

Hierbei ist W das Gewicht der dominierenden Konfiguration, oder, alternativ formuliert: die Anzahl der Mikrozustände des Systems. Die Konfiguration mit den meisten Mikrozuständen (sprich: Realisierungsmöglich-

keiten) ist die wahrscheinlichste.

Wir hatten bei der Einführung der Zustandssumme festgehalten, dass die molekulare Zustandssumme die gesamte Information über die thermodynamischen Eigenschaften eines Systems aus unabhängigen Molekülen im thermodynamischen Gleichgewicht enthält, ihr damit also eine ähnliche Bedeutung zukommt wie der Wellenfunktion in der Quantenmechanik. Wenn diese Aussage korrekt ist, muss sich die Entropie also auch als Funktion der molekularen Zustandssumme formulieren lassen. Dazu setzen wir zunächst die uns bekannte Definition für $\ln W$ ein, Gl. (6), die wir mit Hilfe der Stirlingformel erhalten hatten:

$$S = k_B \left\{ N \ln N - \sum_i N_i \ln N_i \right\}$$

↓ viele Schritte...

$$S = \frac{U - U(0)}{T} + N k_B \ln q$$

Für die genaue Herleitung sei der geneigte Leser auf die einschlägige Literatur verwiesen, z.B. [4, S. 608].

5 Entropie und Shannon-Information

Abschließend sei noch auf einen interessanten Randaspekt hingewiesen, der eher selten Eingang in die Behandlung der statistischen Thermodynamik findet: die Parallele zwischen Boltzmanns Entropiedefinition und Shannons Definition der Information (bzw. Entropie) im Kontext der Signalübertragung [5]. Zunächst sollen einmal beide Gleichungen gegenübergestellt werden.

Entropie nach Boltzmann:

$$S = k_B \ln W$$

Information nach Shannon:

$$I = K \ln P$$

Wie sich unschwer erkennen lässt, sind beide Gleichungen zunächst einmal mathematisch äquivalent. Bei Shannon steht P für den mathematischen Wahrscheinlichkeitsbegriff, mit $P \in [0; 1]$, bei Boltzmann steht W entsprechend für den thermodynamischen Wahrscheinlichkeitsbegriff, mit im Allgemeinen $W \in [1; \infty]$.

Der Faktor K in Shannons Gleichung bestimmt die Einheit von I . Häufig wird hier $-1/\ln(2)$ eingesetzt, so dass sich für die Information die Einheit *bit* („binary digit“) ergibt, also die Zahl notwendiger binärer Ja-Nein-Entscheidungen, die notwendig sind, um bei idealer Ratestrategie eine bestimmte Nachricht zu erraten. Für einen idealen Münzwurf wäre entsprechend $I = 1$.

$$P \leq 1 \Rightarrow \ln P \leq 0$$

$$I \geq 0 \Leftrightarrow K < 0$$

Fragen zum Verständnis

- ❓ Was ist die grundlegende Bedeutung der Entropie für den Ablauf von Prozessen?
- ❓ Welche zwei unabhängigen Begründungen gibt es für die Verwendung des Logarithmus in Boltzmanns Formulierung der Entropie?
- ❓ Welche wichtigen Einschränkungen sind für die Herleitung der Boltzmann-Verteilung essentiell? Was drückt die Boltzmann-Verteilung aus?
- ❓ Was besagt die molekulare Zustandssumme in einem Satz ausgedrückt?
- ❓ Was ist die Bedeutung der molekularen Zustandssumme? Welchem Konzept aus der Quantenmechanik kommt eine ähnliche Bedeutung (für die Quantenmechanik) zu?
- ❓ Was ist der niedrigste Wert, den die molekulare Zustandssumme annehmen kann? Was im allgemeinen Fall ihr größter Wert? Wie sieht das in einem Zweiniveau-System aus?
- ❓ Wie verhalten sich die Besetzungszahlen für beliebig hohe Temperaturen? Was bedeutet das für die Besetzungsinversion, wie sie z.B. für die Funktion eines Lasers konstituierend ist, in einem System im thermischen Gleichgewicht?

Sehr einfach ausgedrückt erlaubt uns Shannons Definition, die Entropie als Maß für das Unwissen über ein System (bzw. seine Konfiguration) zu interpretieren. Entropie ist also ein Maß für die Menge an Information, die wir benötigen, um aus der Kenntnis des vorliegenden Makrozustandes des Systems den tatsächlich vorliegenden Mikrozustand zu bestimmen. Je mehr Realisierungsmöglichkeiten das System hat (je mehr Zustände thermisch zugänglich sind, d.h. je größer die molekulare Zustandssumme ist), desto mehr Information wird benötigt. Entsprechend lässt sich Entropie als Maß für die Unkenntnis über die Zustände aller einzelnen Teilchen beschreiben.

Es sei hier noch betont, dass Shannons Informationsbegriff ein quantitativer und kein qualitativer Informationsbegriff ist. Die Bedeutung einer Zeichenfolge (Semantik) wird von seiner Theorie *nicht* erfasst. Das ist aber für den Kontext, in dem er die Theorie entwickelte (Nachrichtenübertragung) nicht von Belang. Shannons Informationsbegriff findet dafür in vielen Disziplinen, u.a. der Bioinformatik, rege Anwendung, da er ein Maß für die Nichtzufälligkeit einer Zeichenkette liefert.

Weiterführende Literatur

Die Darstellung in diesem Artikel folgt in weiten Teilen jener in [4, Kap. 15]. Boltzmanns Originalarbeit ist [2], die Gleichung wurde in ihrer uns heute bekannten Form von Max Planck aufgestellt [3].

Für den Vergleich zwischen Boltzmann-Entropie und Shannon-Information vgl. u.a. [6, Kap. 2.3]. Dort findet sich ebenfalls die erwähnte mathe-

matische Begründung für die Verwendung des Logarithmus. Die Originalarbeit von Shannon, in der er seine Formulierung der Information bzw. Entropie einführt, ist [5].



Weitere Informationen und ggf. aktuellere Fassungen dieses Artikels finden sich auf der zugehörigen Webseite des Autors:

<https://www.till-biskup.de/de/lehre/entropie/>

Referenzen

- [1] Rudolf Clausius. *Über den zweiten Hauptsatz der mechanischen Wärmetheorie. Ein Vortrag, gehalten in einer allgemeinen Sitzung der 31. Versammlung deutscher Naturforscher und Aerzte zu Frankfurt a. M. am 23. September 1867.* Braunschweig: Friedrich Vieweg und Sohn, 1867.
- [2] Ludwig Boltzmann. Über die Beziehung zwischen dem zweiten Hauptsatz der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respektive den Sätzen über das Wärmegleichgewicht. *Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Classe der Kaiserlichen Akademie der Wissenschaften. II. Abtheilung* 76 (1877), S. 373–435.
- [3] Max Planck. Über das Gesetz der Energieverteilung im Normalspectrum. *Annalen der Physik* 4 (1901), S. 553–563.
- [4] Peter W. Atkins und Julio de Paula. *Physikalische Chemie.* 5. Auflage. Weinheim: Wiley-VCH, 2013.
- [5] Claude Shannon. A mathematical theory of communication. *The Bell System Technical Journal* 27 (1945), S. 379–423.
- [6] Roland Glaser. *Biophysik.* 4., völlig überarbeitete Auflage. Jena: Gustav Fischer, 1996.