

Entropie – Statistische Zugänge

Physikdidaktik, Albert-Ludwigs-Universität Freiburg

22. Mai 2024

Professor Dr. Thomas Filk



Weitere Kurztexte hier: <https://physikdidaktik.uni-freiburg.de/kurztexte/>

universität freiburg



Inhaltsverzeichnis

1	Entropie – Statistische Zugänge	3
1.1	Die Shannon-Information	3
1.1.1	Die Shannon-Information als fehlende Information	3
1.1.2	Die Maximierung der Shannon-Information	5
1.2	Die Boltzmann-Entropie	7
1.3	Die Entropie des idealen Gases	9
1.3.1	Die thermodynamische Entropie eines idealem Gases	9
1.3.2	Die Entropie eines idealen Gases – ein Skalenargument	10
1.3.3	Die quantenmechanische Herleitung der Boltzmann-Entropie eines idealen Gases	10
1.3.4	Die Boltzmann-Entropie eines idealen Gases in der kanonischen Gesamtheit	12

Kapitel 1

Entropie – Statistische Zugänge

Wir unterscheiden mindestens drei verschiedene Entropiebegriffe: (1) die thermodynamische Entropie, die mit einem Fluss an Wärmeenergie verbunden ist, (2) die Boltzmann-Entropie der statistischen Mechanik und (3) die Shannon-Entropie der Informationstheorie. Der thermodynamische Entropiebegriff wird in einen anderen Kurzttext behandelt (siehe „Der zweite Hauptsatz der Thermodynamik“). In diesem Kurzttext geht es im Wesentlichen um die Shannon-Entropie, die man als „Maß für Unkenntnis“ interpretieren kann, sowie der Boltzmann-Entropie, die als „Maß für Unordnung“ gedeutet werden kann.

Schwieriger ist der Beweis, dass diese beiden Definitionen von Entropie äquivalent zum thermodynamischen Entropiebegriff sind. Auch Boltzmann hat dies zunächst nur für das ideale Gas gezeigt. Diese Äquivalenz soll auch hier bewiesen werden.

1.1 Die Shannon-Information

Im Jahre 1948 veröffentlichte Claude Elwood Shannon (1916–2001) sein Hauptwerk *A Mathematical Theory of Communication* [1] und formulierte dort ein Maß für Information (heute als *Shannon-Information* bekannt). Bald darauf wurde der Bezug dieser Größe zum Entropiebegriff in der statistischen Mechanik deutlich. Shannon ging es hauptsächlich um ein Maß für den Verlust an Information, der durch fehlerhafte Signalübertragung bei Kommunikationssystemen auftritt.

1.1.1 Die Shannon-Information als fehlende Information

Wir stellen uns ein System vor, das viele verschiedene Zustände annehmen kann. Mit Ω bezeichnen wir die Menge dieser Zustände und $\{p_i\}$ sei eine Wahrscheinlichkeitsverteilung, die angibt, mit welcher Wahrscheinlichkeit p_i ein bestimmter Zustand $i \in \Omega$ vorliegt. Statt Zustand schreibe ich manchmal auch Mikrozustand oder Ereignis, wobei hier eigentlich ein Elementarereignis gemeint ist. Wir fragen nun nach einem Maß für die *fehlende Information*, um aus der Kenntnis der Wahrscheinlichkeiten $\{p_i\}$ auf einen ganz bestimmten realisierten Zustand schließen zu können.

Als Einheit der Information gilt das Bit, dessen beide Werte 1 oder 0 bzw. TRUE oder FALSE sein können. Als Maß für die fehlende Information wählen wir die minimale Anzahl von Bits (gemessen durch die Antworten auf Ja/Nein-Fragen), mit denen wir die gewünschte Information erlangen. Konkret: Wie viele Ja/Nein-Fragen müssen wir im Mittel mindestens stellen und beantwortet bekommen, um von der allgemeinen Information über die Wahrscheinlichkeitsverteilung zu der Information der speziellen Realisierung zu gelangen?

Wir betrachten zunächst als Beispiel ein Schachbrett mit 64 Feldern. Die allgemeine Information sei: Auf einem Feld des Schachbretts befindet sich eine Figur. Diese Information umfasst die Menge der Möglichkeiten Ω (64 Felder) sowie die Wahrscheinlichkeitsverteilung, in diesem Fall die Gleichverteilung: $p_i = 1/64$. Gesucht ist das konkrete Feld, auf dem sich die Figur befindet. Wir können natürlich alle Felder einzeln erfragen: Steht die Figur auf Feld 1? Steht sie auf Feld 2? etc. Im Durchschnitt würden wir in diesem Fall 32 Fragen benötigen, bis das besetzte Feld bekannt ist.

Bei geschickter Wahl der Fragen benötigen wir jedoch genau 6 Ja/Nein-Fragen, um das Feld zu finden. Dazu unterteilen wir die Möglichkeiten immer in zwei möglichst gleichwahrscheinliche Bereiche (siehe Abb. 1.1). Die erste Frage könne lauten: Befindet sich die Figur auf der linken Spielfeldhälfte? Die zweite Frage wäre beispielsweise: Ist die Figur in der oberen Hälfte? usw. Da sich mit jeder beantworteten Frage die Anzahl der Möglichkeiten halbiert, ist das gesuchte Feld nach sechs Fragen gefunden. Wir können also sagen: Die fehlende Information zwischen der Situation, wo man nur weiß, dass sich auf einem Feld des Spielfelds eine Figur befindet, bis zu der Kenntnis, um welches Feld es sich dabei genau handelt, beträgt 6 Bits.

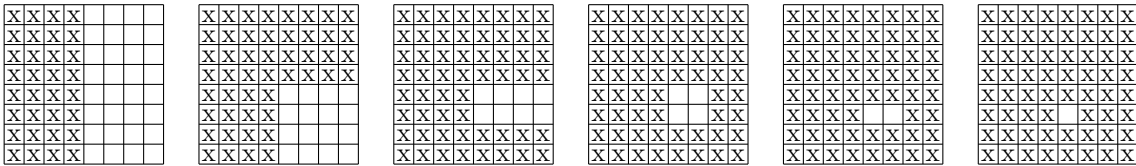


Abbildung 1.1: Nach 6 geschickt gestellten Ja/Nein-Fragen kann die Anzahl der Möglichkeiten von 64 auf 1 reduziert werden. Die erste Frage könnte lauten: Befindet sich die Figur in der linken Hälfte? Die Antwort „Nein“ lässt nur noch 32 Möglichkeiten zu (linkes Bild). Die zweite Frage: Befindet sich die Figur auf der oberen Hälfte? reduziert die Anzahl der verbliebenen Möglichkeiten auf 16 (zweites Bild), etc.

Mit jeder Antwort auf eine geschickt gestellte Frage verringert sich die Anzahl der Möglichkeiten um die Hälfte und damit verdoppeln sich die Wahrscheinlichkeiten für die verbliebenen Möglichkeiten. Nachdem bei dem Schachbrett zu Beginn die Wahrscheinlichkeit für jedes Feld $1/64$ betrug, fallen nach Beantwortung der ersten Frage 32 Felder weg und für die 32 verbliebenen Felder wird die Wahrscheinlichkeit, dort die Figur zu finden, zu $1/32$. Die Fragen enden, wenn nur noch ein Ereignis, das nun die Wahrscheinlichkeit 1 hat, übrig bleibt. Nun ist der Mikrozustand bekannt. Im vorliegenden Fall muss man $6 = -\log_2(1/64) = \log_2 64$ Fragen stellen, bis das Feld auf dem Schachbrett bekannt ist.

Allgemein ist die optimale Strategie die folgende: Sei Ω die Menge der möglichen Zustände und $\{p_i\}$ die Wahrscheinlichkeitsverteilung für die einzelnen Zustände i . Wir teilen nun Ω in zwei disjunkte Teilmengen A_1 und A_2 auf – also $\Omega = A_1 \cup A_2$ und $A_1 \cap A_2 = \emptyset$ –, sodass $p(A_1) = \sum_{i \in A_1} p_i = p(A_2) = \sum_{i \in A_2} p_i = 1/2$. Die erste Frage lautet dann beispielsweise: Befindet sich der gesuchte Mikrozustand in A_1 ? Lautet die Antwort „Ja“, fahren wir mit A_1 fort. Lautet die Antwort „Nein“, fahren wir mit A_2 fort. Wir unterteilen nun die verbliebene Menge an Mikrozuständen wieder in zwei Teilmengen mit jeweils einer Wahrscheinlichkeit von $1/2$ und stellen die nächste Frage.

War p_i die anfängliche Wahrscheinlichkeit für ein Ereignis i und ist dieses nach n Schritten der Ja/Nein-Auswahl noch übrig, so ist dessen Wahrscheinlichkeit nach n Schritten somit $2^n p_i$. Für irgendein n_i an Fragen wird diese Wahrscheinlichkeit gleich 1, der Zustand i liegt also mit Sicherheit vor. Somit ist

$$2^{n_i} p_i = 1 \quad \text{oder} \quad n_i = -\log_2 p_i . \quad (1.1)$$

Das bedeutet, die Anzahl n_i der notwendigen Ja/Nein-Fragen, bis ein Mikrozustand, der zu Beginn mit der Wahrscheinlichkeit p_i vorliegt, als realisiertes Ereignis bekannt ist, ist $n_i = -\log_2 p_i$.

Einem Ereignis, von dem bekannt ist, dass es mit der Wahrscheinlichkeit p_i vorliegen kann, ordnen wir somit die (fehlende) Information

$$I_i = -\log_2 p_i \quad (1.2)$$

zu. Die mittlere (fehlende) Information zu einer Wahrscheinlichkeitsverteilung $\{p_i\}$ ist dann

$$\bar{I}[\{p_i\}] = \sum_i I_i p_i = -\sum_i p_i \log_2 p_i. \quad (1.3)$$

Diese Größe bezeichnet man als die *Shannon-Information* zu einer Verteilung $\{p_i\}$.

Im Allgemeinen handelt es sich hier nur um eine idealisierte untere Grenze für die mittlere Anzahl an Ja/Nein-Fragen, die man stellen muss, bis ein Mikrozustand bekannt ist. Der Grund ist, dass $1/p_i$ im Allgemeinen keine Potenz von 2 ist und somit die Aufteilung einer Teilmenge $A \subset \Omega$ in zwei Unterteilmengen, die jeweils exakt die Hälfte der verbliebenen Wahrscheinlichkeiten haben, meist nicht möglich sein wird. Doch jedes Abweichen von dieser Optimalstrategie erhöht die mittlere Anzahl der notwendigen Fragen, bis der Mikrozustand bekannt ist.

1.1.2 Die Maximierung der Shannon-Information

Wir stellen nun die Frage, für welche Wahrscheinlichkeitsverteilung auf einer Menge Ω von diskreten Elementarereignissen die Shannon-Information maximal wird. Dabei setzen wir bestimmte Informationen über diese Wahrscheinlichkeiten voraus. In diesem Zusammenhang ist es oft einfacher (und näher an den in der Physik verwendeten Größen), wenn man nicht den Logarithmus zur Basis 2 sondern den natürlichen Logarithmus verwendet. Wegen

$$\log_2 x = (\log_2 e) \ln x \quad \text{bzw.} \quad \ln x = (\ln 2) \log_2 x \quad (1.4)$$

sind die beiden Funktionen proportional zueinander und die Extrempunkte sind dieselben.

Die Gleichverteilung

Wir beginnen mit der Suche nach einer Wahrscheinlichkeitsverteilung, welche die Shannon-Information maximiert und von der nur bekannt ist, dass es sich um eine normierte Wahrscheinlichkeit handelt, d.h., dass $\sum_i p_i = 1$. Dazu müssen wir \bar{I} nach p_i variieren, allerdings unter der Zwangsbedingung, dass die Summe der Wahrscheinlichkeiten eins bleibt. Dies berücksichtigen wir durch einen Lagrange-Multiplikator λ , d.h. wir schreiben zunächst:

$$I_\lambda = -\sum_i p_i \ln p_i + \lambda \left(\sum_i p_i - 1 \right) \quad (1.5)$$

und maximieren I_λ sowohl bezüglich p_k (k beliebig) als auch bezüglich λ . Dies führt auf die Gleichungen:

$$\frac{\partial I_\lambda}{\partial p_k} = -\ln p_k - p_k \frac{1}{p_k} + \lambda = -\ln p_k - 1 + \lambda = 0 \quad \text{und} \quad \frac{\partial I_\lambda}{\partial \lambda} = \left(\sum_i p_i - 1 \right) = 0. \quad (1.6)$$

Die erste Gleichung bedeutet, dass p_i unabhängig von i ist, also für alle Zustände gleich, und die zweite Gleichung normiert diese Bedingung. Insgesamt findet wir somit, dass die Gleichverteilung $p_i = 1/|\Omega|$ die maximale Shannon-Information hat. Bei einer Gleichverteilung ist daher die fehlende Information am größten. Dies entspricht in der Thermodynamik der sogenannten mikrokanonischen Verteilung.

Die Boltzmann-Verteilung

Ganz ähnlich können wir nun auch die Frage nach der Verteilung mit der größten Shannon-Information stellen, bei der wir nur den Mittelwert einer Zufallsvariablen kennen. Sei $E : \Omega \rightarrow \mathbb{R}$ eine Zufallsvariable, deren Wert für ein Elementarereignis i mit E_i bezeichnet wird. Der Mittelwert der Zufallsvariablen ist dann

$$\bar{E} = \sum_i E_i p_i. \quad (1.7)$$

Falls es sich bei E_i um die Energie eines Zustands i handelt, bezeichnet man den Erwartungswert in der Thermodynamik als innere Energie.

Wir bilden nun die Varianz der Shannon-Information, allerdings mit zwei Zwangsbedingungen an die Wahrscheinlichkeiten $\{p_i\}$: Die Summe soll eins sein (Normierung) und der Erwartungswert von E soll \bar{E} sein. Für die zweite Bedingung führen wir den Lagrange-Multiplikator β ein (das Vorzeichen ist Konvention):

$$I_{\lambda,\beta} = - \sum_i p_i \ln p_i + \lambda \left(\sum_i p_i - 1 \right) - \beta \left(\sum_i E_i p_i - \bar{E} \right) \quad (1.8)$$

Die Bedingungen für eine stationäre Wahrscheinlichkeitsverteilung (dass es sich hierbei um ein Maximum handelt, muss getrennt gezeigt werden) lauten nun:

$$\frac{\partial I_{\lambda,\beta}}{\partial p_k} = - \ln p_k - 1 + \lambda - \beta E_k = 0 \quad (1.9)$$

sowie die beiden Zwangsbedingungen:

$$\frac{\partial I_{\lambda,\beta}}{\partial \lambda} = \left(\sum_i p_i - 1 \right) = 0 \quad \text{und} \quad \frac{\partial I_{\lambda,\beta}}{\partial \beta} = - \left(\sum_i E_i p_i - \bar{E} \right) = 0 \quad (1.10)$$

Die Lösung zu der ersten Gleichung lautet:

$$p_k = \frac{1}{Z} e^{-\beta E_k} \quad (1.11)$$

mit $Z = e^{\lambda+1}$ als Normierungsfaktor. Die beiden Zwangsbedingungen legen Z und β fest. Formal lässt sich Z in der Form

$$Z = \sum_i e^{-\beta E_i} \quad (1.12)$$

schreiben; dies bezeichnet man in der statistischen Mechanik als die kanonische Zustandssumme. Die größte Shannon-Information einer Wahrscheinlichkeitsverteilung, von der nur der Mittelwert einer Zufallsvariablen bekannt ist, hat somit die Boltzmann-Verteilung bezüglich dieser Variablen. Der Lagrange-Multiplikator β wird durch die Zwangsbedingung an den Mittelwert von E festgelegt.

Die Gauss-Verteilung

Als letztes Beispiel betrachten wir die Verteilung $p(x)$ zu einer kontinuierlichen reellen Variablen x , für die der Mittelwert sowie die Varianz bekannt sind. Mittelwert und Varianz von $p(x)$ sind gegeben durch

$$\bar{x} = \int_{-\infty}^{\infty} x p(x) dx \quad \text{und} \quad \sigma^2 = \int_{-\infty}^{\infty} x^2 p(x) dx. \quad (1.13)$$

Außerdem soll die Verteilung normiert sein:

$$\int_{-\infty}^{\infty} p(x) dx = 1. \quad (1.14)$$

Für die Shannon-Information nehmen wir im kontinuierlichen Fall das Funktional

$$I = - \int_{-\infty}^{\infty} p(x) \ln p(x) dx. \quad (1.15)$$

Diesmal gibt es drei Lagrange-Multiplikatoren:

$$I_{\lambda,\beta,\alpha}[p(x)] = - \int_{-\infty}^{\infty} p(x) \ln p(x) dx + \lambda \left(\int_{-\infty}^{\infty} p(x) dx - 1 \right) \quad (1.16)$$

$$- \beta \left(\int_{-\infty}^{\infty} x p(x) dx - \bar{x} \right) - \alpha \left(\int_{-\infty}^{\infty} x^2 p(x) dx - \sigma^2 \right) \quad (1.17)$$

Die Ableitungen nach den Lagrange-Multiplikatoren ergeben wiederum die Zwangsbedingungen. Die Ableitung nach der Funktion $p(x)$ führen wir als Variationsableitung, d.h., wir schreiben $\delta I = I[p(x) + \delta p(x)] - I[p(x)]$ und entwickeln $I[p(x) + \delta p(x)]$ bis zur linearen Ordnung in $\delta p(x)$:

$$\begin{aligned} I[p(x) + \delta p(x)] &= - \int_{-\infty}^{\infty} (p(x) + \delta p(x)) \ln(p(x) + \delta p(x)) dx + \lambda \left(\int_{-\infty}^{\infty} (p(x) + \delta p(x)) dx - 1 \right) \\ &\quad - \beta \left(\int_{-\infty}^{\infty} x(p(x) + \delta p(x)) dx - \bar{x} \right) - \alpha \left(\int_{-\infty}^{\infty} x^2(p(x) + \delta p(x)) dx - \sigma^2 \right) \end{aligned} \quad (1.18)$$

$$= I[p(x)] + \int_{-\infty}^{\infty} \delta p(x) [-\ln p(x) - 1 + \lambda - \beta x + \alpha x^2] dx + o(|\delta p(x)|). \quad (1.19)$$

Damit die 1. Variation von $I[p(x)]$ verschwindet, muss der lineare Term in $\delta p(x)$ für alle $\delta p(x)$ verschwinden. Das führt auf die Gleichung

$$-\ln p(x) - 1 + \lambda - \beta x - \alpha x^2 = 0 \quad (1.20)$$

oder

$$p(x) = \exp(-1 + \lambda - \beta x - \alpha x^2). \quad (1.21)$$

Es handelt sich bei $p(x)$ somit um eine Gauß-Verteilung. Wir können die drei Lagrange-Multiplikatoren durch die vorgegebenen Werte \bar{x} , σ sowie eine Normierung ersetzen und erhalten:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right). \quad (1.22)$$

Die Gauß-Verteilung hat daher die größte Shannon-Information unter der Bedingung, dass von einer Wahrscheinlichkeitsverteilung ihr Mittelwert und ihre Varianz bekannt sind.

1.2 Die Boltzmann-Entropie

Ludwig Boltzmann (1844–1906) entwickelte seine Vorstellung einer statistischen Entropie, die durch den Logarithmus der Anzahl der Mikrozustände zu einem gegebenen Satz von Makrozuständen gegeben ist, in den 70er Jahren des 19. Jahrhunderts. Diese Beziehung lautet:

$$S = k_B \ln |\Omega|, \quad (1.23)$$

wobei $k_B = 1,380\,649 \cdot 10^{-23} \text{ J/K}$ die Boltzmann-Konstante ist und $|\Omega|$ die Anzahl der Mikrozustände zu einem gegebenen Makrozustand bezeichnet. Diese Formel (mit Boltzmanns Bezeichnung W für die Anzahl der Zustände) findet sich heute auf seinem Grabstein auf dem Wiener Zentralfriedhof (siehe Abb. 1.2).

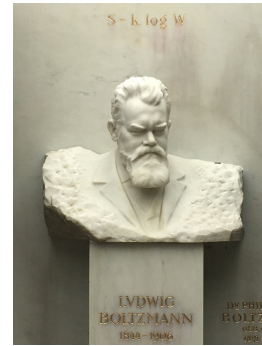


Abbildung 1.2: Der Grabstein von Ludwig Boltzmann auf dem Zentralfriedhof von Wien. Oben ist die Formel für die statistische Entropie erkennbar. (eigenes Bild)

In strenger Form gilt diese Formel zunächst nur für eine mikrokanonische Gesamtheit, also ein thermodynamisches System, das vollkommen isoliert von seiner Umgebung ist, bei dem insbesondere auch kein Austausch von Wärmeenergie mit der Umgebung möglich ist (d.h., dieses System hat adiabatische Wände). Ein solches thermodynamischer System ist durch einen Satz von makroskopischen Größen charakterisiert: die Energie E , das Volumen V , die Stoffmenge $n = N/N_A$, wobei N die Anzahl der Teilchen in dem System und N_A die Avogadro-Konstante sind, sowie eventuell weitere makroskopisch beobachtbare und kontrollierbare Größen. Die Menge $\Omega(E, V, n)$ der Mikrozustände, die mit diesen makroskopischen Größen verträglich sind, hängt natürlich von den vorgegebenen Parametern ab.

In der klassischen Mechanik ist „Anzahl der Mikrozustände“ eigentlich nicht gut definiert. Zunächst einmal handelt es sich bei festem E, V, n um einen Unterraum im $6N$ -dimensionalen Phasenraum der klassischen Teilchen. Das Volumen dieses Unterraums, das als Maß für die Anzahl der Zustände dienen könnte, verschwindet jedoch für eine feste Energie E (es handelt sich um eine Hyperfläche im Phasenraum, deren Dimension um 1 kleiner ist als die des Phasenraums selbst). Daher betrachtet man für die erlaubten Energien ϵ meist ein Energieintervall, $\epsilon \in [E - \Delta E, E]$, bzw. die Vollkugel $\epsilon \leq E$. Es zeigt sich, dass bei sehr vielen Teilchen (N in der Größenordnung der Avogadro-Zahl) hier kein messbarer Unterschied besteht. Außerdem müssen zwei „Überreste“ der Quantentheorie berücksichtigt werden: die Ununterscheidbarkeit der Teilchen – dies führt auf einen Faktor $1/N!$ für die Abzählung verschiedener Mikrozustände – und die Tatsache, dass Zustände im Phasenraum innerhalb eines Volumens von h^{3N} (h die Planck’sche Konstante) nicht auflösbar sind, d.h., das Volumen im Phasenraum ist in Einheiten von h^{3N} anzugeben. In der Quantentheorie sind die Zustände diskret und sollten der Ununterscheidbarkeit Rechnung tragen, sodass die Anzahl der Zustände mit einer Energie kleiner als ein vorgegebener Wert E wohl definiert ist.

In der mikrokanonischen Gesamtheit haben alle erlaubten Zustände dieselbe Wahrscheinlichkeit, sodass $p_i = 1/|\Omega|$ und somit die Shannon-Information

$$S = - \sum_i p_i \ln p_i = \sum_i p_i \ln |\Omega| = \ln |\Omega| \quad (1.24)$$

ist. Bis auf die Boltzmann-Konstante $k_B = 1,380\,649 \cdot 10^{-23} \text{ J/K}$, die eher historische Gründe hat, stimmt dies mit der Boltzmann-Entropie überein. Der Faktor k_B hat zwei Gründe: Zum einen hat die Entropie in der Thermodynamik die Dimension „Energie/Temperatur“ also J/K, und zum anderen ist die Anzahl der Mikrozustände so immens groß, dass man riesige Zahlen für $\ln |\Omega|$ erhält. Damit diese Zahlen mit den in der Thermodynamik vertrauten Größen übereinstimmen, berücksichtigt man den Faktor k_B .

Zu Beginn dieses Kurztextes wurde behauptet, dass die Boltzmann-Entropie als „Maß für Unordnung“ verstanden werden kann. Wenn man unter Ordnung „jedes Ding befindet sich an seinem Platz“ versteht, gibt es nur einen (oder wenige) geordnete Zustände. Je mehr Dinge sich nicht an

ihrem Platz befinden, umso größer ist die Unordnung, umso größer wird aber auch die Anzahl der Möglichkeiten, die für einen ungeordneten Zustand bestehen. Die Unordnung ist umso größer, je länger man suchen muss, um einen bestimmten Gegenstand zu finden, bzw. je mehr Möglichkeiten es gibt, wo sich ein Gegenstand befinden kann. Die Entropie wächst typischerweise mit der Temperatur und bei einer absoluten Temperatur $T = 0$ gibt es in der Quantentheorie nur einen Grundzustand. Daher ist die Entropie dort null. Je höher die Temperatur wird, umso ungeordneter wird die Bewegung der Teilchen in einem System. In diesem Sinne ist die obige Aussage zu verstehen.

1.3 Die Entropie des idealen Gases

Ein Teil der großen Leistung von Ludwig Boltzmann bestand darin zu zeigen, dass die von ihm angegebene Formel für die Entropie $S = k_B \ln |\Omega|$ mit der thermodynamischen Entropie übereinstimmt. Konkret bewiesen hat er das für ein ideales Gas. Im Folgenden werden die Boltzmann-Entropie sowie die thermodynamische Entropie eines idealen Gases auf verschiedenen Wegen hergeleitet.

1.3.1 Die thermodynamische Entropie eines idealen Gases

In der Thermodynamik berechnet man die Entropie nach der Vorschrift von Sommerfeld (siehe Der zweite Hauptsatz der Thermodynamik), indem man die Größe $\delta Q/T$ (Wärmemenge dividiert durch die absolute Temperatur) entlang eines reversiblen Weges integriert, also

$$S - S_0 = \int_0^1 \frac{\delta Q}{T}. \quad (1.25)$$

Die Wärmemenge ist allgemein gegeben durch $\delta Q = dE + pdV$; dies ist der Energiesatz – der erste Hauptsatz der Thermodynamik – für eine rein mechanische Arbeit $\delta W = pdV$. Speziell für ein ideales Gas gelten die kalorische Zustandsgleichung $E = \frac{3}{2}nRT$ sowie die ideale Gasgleichung $pV = nRT$ (n Stoffmenge in Mol, $R = N_A k_B = 8,314\,462\,618\,153\,24\text{ J}/(\text{K} \cdot \text{mol})$ ist die Gaskonstante mit der Avogadro-Zahl N_A), sodass

$$\delta Q = \frac{3}{2}nR dT + \frac{nRT}{V} dV. \quad (1.26)$$

Offenbar handelt es sich hierbei nicht um ein totales Differential, denn dann müsste $\frac{\partial}{\partial V} \frac{3}{2}nR = 0$ gleich $\frac{\partial}{\partial T} \frac{nRT}{V} = \frac{nR}{V}$ sein, was offensichtlich nicht gilt, allerdings ist

$$\frac{\delta Q}{T} = \frac{3}{2} \frac{nR}{T} dT + \frac{nR}{V} dV \quad (1.27)$$

ein totales Differential (wie es der zweite Hauptsatz auch fordert) und es gilt:

$$S(T, V) = \int \frac{\delta Q}{T} = \int \frac{3}{2} \frac{nR}{T} dT + \int \frac{nR}{V} dV = \frac{3}{2} nR \ln \frac{T}{T_0} + nR \ln \frac{V}{V_0}, \quad (1.28)$$

wobei T_0 und V_0 die Temperatur und das Volumen am Referenzpunkt S_0 sind. Dies ist die gesuchte Funktion für die Entropie des idealen Gases. Allerdings war schon Gibbs bekannt, dass diese Größe nicht additiv ist (d.h., wenn man das System verdoppelt, sodass sich V und n verdoppeln, sollte sich auch S verdoppeln). Wegen der Ununterscheidbarkeit der Teilchen muss man einen Term $Nk_b \ln N$ (wobei N die Teilchenzahl ist) subtrahieren. Dies spielt an dieser Stelle aber keine Rolle.

Um den Zusammenhang zur statistischen Herleitung der Entropie zu verdeutlichen, ersetzen wir noch die Stoffmenge n (in Mol) durch die Teilchenzahl $N = nN_A$ und die Gaskonstante R durch die Boltzmann-Konstante $R = N_A k_B$. Damit folgt für die Entropie:

$$S(T, V) = \frac{3}{2} N k_B \ln \frac{T}{T_0} + N k_B \ln \frac{V}{V_0}. \quad (1.29)$$

1.3.2 Die Entropie eines idealen Gases – ein Skalenargument

Die Entropie des idealen Gases Gl. 1.29 lässt sich auch durch ein einfaches Skalenargument nachvollziehen. Dazu betrachten wir zunächst die zu der Entropie gehörende „Menge an Mikrozuständen“ im Sinne von Boltzmann:

$$|\Omega| = \exp\left(\frac{S - S_0}{k_B}\right) = \left(\frac{T}{T_0}\right)^{\frac{3}{2}N} \left(\frac{V}{V_0}\right)^N = \left(\frac{E}{E_0}\right)^{\frac{3}{2}N} \left(\frac{V}{V_0}\right)^N. \quad (1.30)$$

Bei der zweiten Gleichung wurde ausgenutzt, dass die Energie eines idealen Gases nicht vom Volumen abhängt und proportional zur Temperatur T ist. Wir können diese Abhängigkeiten durch einfache Skalenargumente verstehen:

1. Wenn wir das Volumen V verdoppeln, verdoppelt sich für jedes Teilchen das Phasenraumvolumen, d.h. wir erhalten einen Faktor 2^N . Allgemeiner gilt: Wenn wir das Volumen um einen Faktor λ skalieren, skalieren wir das Phasenraumvolumen pro Teilchen um denselben Faktor λ und somit das Phasenraumvolumen von N Teilchen um den Faktor λ^N . Daher ist es naheliegend, dass das Phasenraumvolumen von N freien Teilchen proportional zu V^N ist.
2. Wenn wir die Energie eines Teilchens verdoppeln, erhöhen wir jede seiner Impulskomponenten um $\sqrt{2}$, da die kinetische Energie – nur diese gibt es bei einem idealen Gas – proportional zu p^2 ist. Insgesamt erhöht sich das Phasenraumvolumen pro Teilchen um $2^{3/2}$, da ein Teilchen drei Impulskomponenten hat. Für N Teilchen bedeutet dies einen Faktor $2^{3N/2}$. Allgemeiner: Wenn wir die Gesamtenergie um einen Faktor λ skalieren, skaliert das Phasenraumvolumen von N Teilchen um den Faktor $\lambda^{3N/2}$. Damit ist das Phasenraumvolumen von N Teilchen proportional zu $E^{3N/2}$.

Insgesamt folgt daher, dass sich für ein ideales Gas die Anzahl der Zustände zu einer mittleren Energie E und einem Volumen V wie $E^{3N/2}V^N$ verhält. Da $|\Omega|$ eine dimensionslose Zahl sein muss und sich diese Überlegungen immer auf eine Referenzenergie E_0 und ein Referenzvolumen V_0 beziehen, folgt Gl. 1.30.

1.3.3 Die quantenmechanische Herleitung der Boltzmann-Entropie eines idealen Gases

Wir betrachten ein N -Teilchensystem in einem kubischen Kasten der Kantenlänge L , also vom Volumen $V = L^3$. Die quantenmechanischen Energiezustände für ein freies Teilchen in einem solchen Kasten sind

$$\epsilon_{n_1, n_2, n_3} = \frac{\hbar^2 \pi^2}{2mL^2} (n_1^2 + n_2^2 + n_3^2) \quad n_i = 1, 2, \dots \in \mathbb{N}. \quad (1.31)$$

Für N freie Teilchen sind die erlaubten Energiezustände durch die Summe der Einteilchenenergien gegeben:

$$\epsilon = \frac{\hbar^2 \pi^2}{2mL^2} \sum_{i=1}^{3N} n_i^2. \quad (1.32)$$

Wir bestimmen zunächst die kumulative Anzahl $n(E, N)$ der Zustände von N freien Teilchen, so dass deren Energie kleiner oder gleich E ist. Das führt auf die Frage: Wie viele Kombinationen von natürlichen Zahlen n_i gibt es, sodass $\sum_i^{3N} n_i^2 \leq R^2$ für eine vorgegebene (große) Zahl R ? Wir lösen dieses Problem in zwei Schritten:

1. Wir bestimmen zunächst das Volumen einer d -dimensionalen Kugel.
2. Wir lösen damit das genannte kombinatorische Problem.

Schließlich betrachten wir für die Entropie nur die führenden Terme.

Das Volumen der d -dimensionalen Kugel

Das Volumen einer d -dimensionalen Kugel vom Radius R ist gegeben durch

$$\text{Vol}(R,d) = \int d\Omega \int_0^R r^{d-1} dr = \frac{1}{d} O(d) R^d, \quad (1.33)$$

wobei $d\Omega$ die Winkelintegration in d Dimensionen bezeichnet und $O(d)$ ist die Oberfläche der Einheitskugel (Kugel vom Radius $R = 1$). Diese Oberfläche bestimmen wir folgendermaßen: Wir berechnen das Integral

$$\left(\int_{-\infty}^{\infty} e^{-x^2} dx \right)^d = \pi^{d/2} \quad (1.34)$$

in Kugelkoordinaten. Dort gilt

$$\left(\int_{-\infty}^{\infty} e^{-x^2} dx \right)^d = \int d\Omega \int_0^{\infty} r^{d-1} e^{-r^2} dr = \frac{1}{2} O(d) \Gamma(d/2), \quad (1.35)$$

wobei die Definition der Γ -Funktion

$$\int_0^{\infty} r^{d-1} e^{-r^2} dr = \frac{1}{2} \int_0^{\infty} y^{d/2-1} e^{-y} dy = \frac{1}{2} \Gamma(d/2) \quad (1.36)$$

verwendet wurde. Aus diesen beiden Ergebnissen erhalten wir:

$$O(d) = \frac{2\pi^{d/2}}{\Gamma(d/2)} \quad \Rightarrow \quad \text{Vol}(R,d) = \frac{2}{d} \frac{\pi^{d/2}}{\Gamma(d/2)} R^d = \frac{\pi^{d/2}}{\Gamma(\frac{d}{2} + 1)} R^d. \quad (1.37)$$

Die Kombinatorik

Die Bedingung $\sum_{i=1}^{3N} n_i^2 \leq R^2$ lässt sich folgendermaßen interpretieren: Gesucht ist die Anzahl der Punkte mit ganzzahligen, positiven Koordinaten n_i innerhalb einer $3N$ -dimensionalen Kugel vom Radius R . Da diese Punkte ein $3N$ -dimensionales Gitter bilden und jede Gitterzelle das Volumen 1 hat, können wir die Frage in guter Näherung dadurch beantworten, dass wir das Volumen einer $3N$ -dimensionalen Kugel vom Radius R betrachten.

Dieses Volumen ist durch Gl. 1.37 gegeben. Allerdings sind wir nicht an beliebigen ganzzahligen $3N$ -Tupeln interessiert, sondern nur an den positiven. Da in dem Kugelvolumen jede Koordinate positive und negative Werte annehmen kann, erhalten wir das Volumen im rein positiven „Quadranten“, indem wir noch durch 2^{3N} dividieren. Außerdem müssen wir noch die Ununterscheidbarkeit der Teilchen berücksichtigen, sodass wir noch durch $N!$ dividieren müssen. Die gesuchte Anzahl an Kombinationen $n(E,V,N)$ ist somit näherungsweise:

$$n(E,V,N) = \frac{1}{2^{3N} N!} \text{Vol}(R,3N) \quad \text{mit} \quad R^2 = \frac{2mL^2}{\hbar^2 \pi^2} E. \quad (1.38)$$

Wir könnten für das kombinatorische Problem nun noch Korrekturterme bestimmen, die sich daraus ergeben, dass die Zahlen n_i nicht 0 sein dürfen. Für große Werte von R , sodass Terme der Ordnung N/R vernachlässigt werden können, spielen diese Korrekturen jedoch keine Rolle. Als Ergebnis erhalten wir:

$$\begin{aligned} n(E,V,N) &= \frac{1}{2^{3N} N!} \left(\frac{2}{3N} \frac{\pi^{3N/2}}{\Gamma(3N/2)} \right) \left(\frac{2mL^2}{\hbar^2 \pi^2} E \right)^{3N/2} \\ &= \frac{1}{N! \hbar^{3N}} \left(\frac{2}{3N} \frac{1}{\Gamma(3N/2)} \right) \left(\frac{m}{2\pi} E \right)^{3N/2} V^N. \end{aligned} \quad (1.39)$$

Man erkennt an diesem Ergebnis schon die wesentlichen Elemente: (1) der Faktor $1/(N! \hbar^{3N})$, der sich aus der Quantentheorie ergibt (dass man den Phasenraum mit h und nicht mit \hbar normieren sollte, wird in Abschnitt 1.3.4 gezeigt), sowie (2) die Abhängigkeit $E^{3N/2} V^N$.

Näherungen

In Abschnitt 1.3.1 wurde erwähnt, dass man zur thermodynamischen Entropie noch einen Korrekturterm anbringen muss, um das Gibbs'sche Paradoxon zu vermeiden: Die Entropie sollte eine additive Größe sein, d.h., wenn sich die extensiven Größen E und V verdoppeln, soll sich auch die Entropie verdoppeln. Dazu betrachten wir die führende N -Abhängigkeit vom Logarithmus von $n(E,N)$.

Für die Γ -Funktion sowie den Term $N!$ verwenden wir die sogenannte Stirling'sche Näherung:

$$N! \sim \sqrt{2\pi N} \left(\frac{N}{e}\right)^N \quad \text{und} \quad \Gamma(x) \sim \sqrt{\frac{2\pi}{x}} \left(\frac{x}{e}\right)^x. \quad (1.40)$$

Da wir letztendlich am Logarithmus von $n(E,N)$ interessiert sind und nur an der Abhängigkeit von N , E und V , betrachten wir nur die Terme, die im Logarithmus proportional zu N sind (alle anderen Terme verschwinden für große Werte von N im Vergleich):

$$\frac{S}{k_B} = \ln n(E,V,N) \approx N \ln \left(\frac{E^{\frac{3}{2}} V C}{N^{\frac{5}{2}}} \right) \quad \text{mit} \quad C = \frac{e^{\frac{5}{2}} m^{\frac{3}{2}}}{(3\pi)^{\frac{3}{2}} \hbar^3}. \quad (1.41)$$

Die Konstante C enthält die Masse m der Teilchen und \hbar in einer Form, sodass die Dimensionen stimmen, das Argument des Logarithmus also dimensionslos ist. Die anderen Konstanten wie π und e sind für die Entropie nicht relevant. Wichtig ist das Skalenverhalten von S unter der Transformation $(E,V,N) \rightarrow (\lambda E, \lambda V, \lambda N)$. Man erkennt, dass sich im Argument des Logarithmus alle Potenzen von λ wegheben, sodass insgesamt von dem führenden Faktor N nur ein Faktor λ auftritt, wie man es von der Entropie als extensiver Größe auch erwartet. Ohne den Faktor $N!$ wäre das nicht der Fall.

Man beachte, dass die Energie proportional zu N ist, d.h., wenn wir die Energie über die Beziehung $E = \frac{3}{2} N k_B T$ durch die absolute Temperatur T ersetzen, erhalten wir im Logarithmus die Kombination $T^{3/2} V/N$, die ebenfalls das richtige Skalenverhalten besitzt. V/N ist das Volumen pro Teilchen

1.3.4 Die Boltzmann-Entropie eines idealen Gases in der kanonischen Gesamtheit

Abschließend bestimmen wir die freie Energie in der kanonischen Gesamtheit. Aus dieser Größe erhalten wir nicht nur unmittelbar die thermische Zustandsgleichung (also die ideale Gasgleichung $p = p(V,T)$) sowie die kalorische Zustandsgleichung (die Beziehung zwischen der Energie und der Temperatur und dem Volumen, $E = E(V,T)$, wobei beim idealen Gas die Energie nicht vom Volumen abhängt) sondern auch die Entropie des idealen Gases.

Die kanonische Zustandssumme

Die kanonische Zustandssumme (einschließlich des Faktors $1/(N!h^{3N})$ aus der Quantentheorie – man beachte, dass hier die Planck'sche Konstante h und nicht \hbar steht) lautet in der klassischen Näherung:

$$Z(T,V,N) = \frac{1}{N!h^{3N}} \int_{-\infty}^{\infty} \prod_i^{3N} dp_i \int_{-\infty}^{\infty} \prod_i^{3N} dx_i \exp \left(-\beta \sum_{i=1}^{3N} \frac{p_i^2}{2m} \right). \quad (1.42)$$

Die Integrale über die Ortsfreiheitsgrade x_i liefern insgesamt einen Beitrag V^N . Die Integrale über die Impulsfreiheitsgrade sind Gauß'sche Integrale und liefern für jeden Freiheitsgrad den Beitrag

$$\int_{-\infty}^{\infty} dp \exp \left(-\frac{p^2}{2mk_B T} \right) = \sqrt{2\pi mk_B T}. \quad (1.43)$$

Damit folgt insgesamt für die Zustandssumme Z bzw. die freie Energie F :

$$Z = \frac{1}{N!h^{3N}} V^N (2\pi m k_B T)^{3N/2} = \exp\left(-\frac{F}{k_B T}\right). \quad (1.44)$$

Die kalorische Zustandsgleichung

Aus der allgemeinen Darstellung der kanonischen Zustandssumme $Z = \sum_i e^{-\beta E_i}$ erhalten wir für den Erwartungswert der Energie:

$$\bar{E} = -\frac{\partial}{\partial \beta} \ln Z = \frac{1}{Z} \sum_i E_i e^{-\beta E_i} \quad (1.45)$$

und damit für das ideale Gas die kalorische Zustandsgleichung:

$$\bar{E} = -\frac{\partial}{\partial \beta} \ln \beta^{-3N/2} = \frac{3N}{2} \beta^{-1} = \frac{3N}{2} k_B T. \quad (1.46)$$

Die thermische Zustandsgleichung – Die ideale Gasgleichung

Aus der allgemeinen Beziehung $\left(\frac{\partial F}{\partial V}\right)_{T,N} = -p$ erhalten wir die ideale Gasgleichung:

$$p = \frac{\partial}{\partial V} (k_B T \ln V^N) = N k_B T \frac{1}{V} \quad \text{bzw.} \quad pV = N k_B T. \quad (1.47)$$

Die Entropie

In der kanonischen Gesamtheit erhält man die Entropie aus der Beziehung

$$\left(\frac{\partial F}{\partial T}\right)_{V,N} = -S. \quad (1.48)$$

Mit $F = -k_B T \ln Z$ folgt:

$$S = k_B \ln Z + k_B T \frac{\partial \ln Z}{\partial T} = k_B \ln Z + k_B \frac{3N}{2} \quad (1.49)$$

oder

$$\frac{S}{k_B} = \ln Z + \frac{3N}{2}. \quad (1.50)$$

In diesem Ausdruck betrachten wir wiederum nur die führenden Terme der Stirling-Formel, wobei der Term $\frac{3N}{2} = N \ln e^{3/2}$ berücksichtigt wird:

$$\ln Z + \frac{3}{2}N \approx N \ln \left(\frac{e^{5/2} (2\pi m k_B T)^{3/2} V}{N h^3} \right) \quad (1.51)$$

Ersetzen wir in dieser Formel die Temperatur T durch die Energie E , also $k_B T = \frac{2}{3} \frac{E}{N}$, erhalten wir

$$\frac{S}{k_B} \approx N \ln \left(\frac{e^{5/2} (4\pi m E)^{3/2} V}{N^{5/2} 3^{3/2} h^3} \right) = N \ln \left(\frac{E^{3/2} V}{N^{5/2}} C \right) \quad (1.52)$$

mit

$$C = \frac{e^{5/2} (4\pi)^{3/2} m^{3/2}}{3^{3/2} h^3} = \frac{e^{5/2} m^{3/2}}{(3\pi)^{3/2} h^3}. \quad (1.53)$$

Wir erhalten somit exakt dasselbe Ergebnis wie aus der quantenmechanischen Berechnung, einschließlich desselben Faktors C aus Gl. 1.41. Man beachte, dass die beiden Konstanten C nur deshalb identisch sind, weil wir (1) die klassische Zustandssumme mit $1/h^{3N}$ normiert haben (und nicht mit $1/h^{3N}$) und weil wir bei der Berechnung der Entropie aus der freien Energie auch den konstanten Term $\frac{3}{2}N$ berücksichtigt haben.

Literaturverzeichnis

- [1] Shannon, C. E.; *A Mathematical Theory of Communication*; Bell System Technical Journal. Short Hills N.J. (1948) 379–423 & 623–656.