VI.4 Statistische Entropie

Die probabilistische Herangehensweise wird durch einen Mangel an Kenntnisse über den mikroskopischen Zustand des Systems verursacht. Man darf sich aber fragen, wie viel Information fehlt, indem nur Wahrscheinlichkeiten $\{p_k\}$ bzw. eine Wahrscheinlichkeitsdichte ρ_N bekannt sind?

Intuitiv ist die fehlende Information „klein“, wenn die Wahrscheinlichkeitsverteilung für nur wenige Zustände erhebliche Werte annimmt. Dagegen ist die fehlende Information größer, wenn sich die Verteilung über viele Zustände erstreckt. Um dieser Intuition Sinn zu geben, wird ein Maß für die fehlende Information in Wahrscheinlichkeitstheorie in § VI.4.1 eingeführt. Dann wird dieses Maß zu den Fällen der Wahrscheinlichkeitsverteilungen angewandt, die in der Beschreibung eines quantenmechanischen (§ VI.4.2) bzw. eines klassischen (§ VI.4.3) Systems eingeführt werden.

VI.4.1 Information und Wahrscheinlichkeiten

Der Einfachheit halber werden hiernach mit Ausnahme von Definition (VI.24) nur diskrete Wahrscheinlichkeitsverteilungen betrachtet. Die Verallgemeinerung auf den Fall einer stetigen Verteilung ist ziemlich einfach.

VI.4.1 a Statistische Entropie in der Informationstheorie

Es seien M Ereignisse $\{\omega_1, \ldots, \omega_M\}$ mit jeweiligen Wahrscheinlichkeiten $p_1, \ldots, p_M$. Um die *fehlende Information* (*Ignoranz*, *Informationsdefizit*, oder kurz — und ziemlich irreführend — „Information“) zu messen, die dem Einsatz dieser Wahrscheinlichkeitsverteilung entspricht, hat Claude Shannon [22] die *statistische Entropie*

$$\boxed{S_{\text{stat}}(p_1, \ldots, p_M) \equiv -k \sum_{m=1}^{M} p_m \ln p_m} \tag{VI.23}$$

mit k einer zuerst beliebigen positiven Konstante eingeführt. $S_{\text{stat}}(p_1, \ldots, p_M)$ ist somit genau der Erwartungswert von $-k \ln p_m$.

In der Informationstheorie wird gewöhnlich $k = 1/\ln 2$ angenommen, so dass

$$S_{\text{stat}}(p_1, \ldots, p_M) = \left\langle \log_2 \frac{1}{p_m} \right\rangle$$

und die statistische Entropie dimensionslos[21] ist.

> Die Größe $I(p_m) \equiv -\log_2 p_m$ ist der *Informationsgehalt* oder auch *Überraschungswert* assoziiert mit dem Ereignis ω_m (bei Shannon handelte es sich bei den Ereignissen um mögliche Nachrichten). Umso unwahrscheinlicher ω_m ist, desto aufschlussreich ist sein Auftreten.

Im Fall einer stetigen Wahrscheinlichkeitsverteilung, beschrieben durch eine Wahrscheinlichkeitsdichte $p(x)$, lautet die statistische Entropie

$$\boxed{S_{\text{stat}}\big(p(x)\big) \equiv -k \int p(x) \ln p(x) \, \mathrm{d}x.} \tag{VI.24}$$

Hier auch ist S_{stat} gleich dem Erwartungswert des Logarithmus der Verteilung.

[21] In diesem Zusammenhang ist die Einheit der statistischen Entropie das Bit.

VI.4.1 b Eigenschaften der statistischen Entropie

Die durch Gl. (VI.23) definierte statistische Entropie besitzt viele mathematische Eigenschaften, die hiernach ohne Beweis aufgezählt werden.

Symmetrie

Die Funktion $S_{\text{stat}}(p_1, \ldots, p_M)$ ist eine symmetrische Funktion ihrer M Variablen.

Die statistische Entropie ist nicht-negativ

Für alle Wahrscheinlichkeitsverteilungen $p_1, \ldots, p_M$ gilt $S_{\text{stat}}(p_1, \ldots, p_M) \geq 0$.

Minimum

Die statistische Entropie ist minimal, wenn eine der Wahrscheinlichkeiten gleich 1 ist, d.h. wenn das entsprechende Ereignis sicher ist. Dann ist $S_{\text{stat}}(0, \ldots, 1, \ldots, 0) = 0$.

Maximum

Bei fester Zahl M von möglichen Ereignissen ist die statistische Entropie $S_{\text{stat}}(p_1, \ldots, p_M)$ maximal für die diskrete Gleichverteilung $p_1 = \cdots = p_M = 1/M$, d.h. wenn die Ereignisse gleich wahrscheinlich sind. Dann ist $S_{\text{stat}}(p_1, \ldots, p_M) = k \ln M$.

Im Fall gleich wahrscheinlicher Ereignisse nimmt die statistische Entropie $S_{\text{stat}}(p_1, \ldots, p_M)$ mit der Zahl M zu.

Unmögliche Ereignisse

Wenn einige Ereignisse nicht stattfinden können, d.h. eine Wahrscheinlichkeit $p_m = 0$ haben, tragen sie zur statistischen Entropie nicht bei:

$$S_{\text{stat}}(p_1, \ldots, p_M, 0, \ldots, 0) = S_{\text{stat}}(p_1, \ldots, p_M).$$

Additivität

Seien die M Ereignisse in zwei Gruppen $A = \{\omega_1, \ldots, \omega_m\}$ und $B = \{\omega_{m+1}, \ldots, \omega_M\}$ eingeteilt. Dann ist $p'_A = p_1 + \cdots + p_m$ bzw. $p'_B = p_{m+1} + \cdots + p_M$ die Wahrscheinlichkeit, dass eines unter den Ereignissen von A bzw. B auftritt. Die statistische Entropie (VI.23) kann dann als

$$S_{\text{stat}}(p_1, \ldots, p_M) = S_{\text{stat}}(p'_A, p'_B) + p'_A S_{\text{stat}}\left(\frac{p_1}{p'_A}, \ldots, \frac{p_m}{p'_A}\right) + p'_B S_{\text{stat}}\left(\frac{p_{m+1}}{p'_B}, \ldots, \frac{p_M}{p'_B}\right) \qquad \text{(VI.25)}$$

umgeschrieben werden. Der erste Beitrag beschreibt die Auswahl zwischen A und B, die zwei nächsten die Auswahl innerhalb A bzw. B (z.B. ist p_1/p'_A die Wahrscheinlichkeit für das Auftreten von ω_1 unter der Bedingung, dass A auftritt). Somit beschreibt Gl. (VI.25) den Gewinn von Information in zwei Schritten.

Unteradditivität

Seien zwei Ereignismengen $A = \{\omega_1, \ldots, \omega_M\}$ und $B = \{\omega'_1, \ldots, \omega'_N\}$, und kombinierte Ereignisse ω''_{mn} bestehend aus der Vereinigung von einem Ereignis aus A und einem aus B, mit Wahrscheinlichkeiten p''_{mn}. Dann gilt

$$S_{\text{stat}}(p''_{11}, \ldots, p''_{mn}, \ldots, p''_{MN}) \leq S_{\text{stat}}(p_1, \ldots, p_M) + S_{\text{stat}}(p'_1, \ldots, p'_N), \qquad \text{(VI.26)}$$

mit $p_m = \displaystyle\sum_{n=1}^{N} p''_{mn}$ bzw. $p'_n = \displaystyle\sum_{m=1}^{M} p''_{mn}$ der Auftretenswahrscheinlichkeit für ω_m bzw. ω'_n.

Die Gleichheit findet nur dann statt, wenn A und B stochastisch unabhängig sind, so dass $p''_{mn} = p_m p'_n$ für alle m, n: die fehlende Information über die kombinierten Ereignisse ist kleiner, wenn die Ereignisse von A und solche von B nicht unabhängig sind.

Konkavität

Seien zwei Wahrscheinlichkeitsverteilungen $p_1, \ldots, p_M$ und $p'_1, \ldots, p'_M$ für dieselben Ereignisse. Wenn λ eine reelle Zahl mit $0 < \lambda < 1$ ist, dann bilden die $\lambda p_m + (1-\lambda)p'_m$ für $m = 1, \ldots, M$ eine neue Wahrscheinlichkeitsverteilung, und es gilt

$$S_{\text{stat}}\big(\lambda p_1 + (1-\lambda)p'_1, \ldots, \lambda p_M + (1-\lambda)p'_M\big) \geq \lambda S_{\text{stat}}(p_1, \ldots, p_M) + (1-\lambda)S_{\text{stat}}(p'_1, \ldots, p'_M), \quad \text{(VI.27)}$$

wobei die Gleichung nur dann gilt, wenn $p_m = p'_m$ für jedes m.

Die Ungleichung bedeutet, dass die Mischung zweier „identischen" Systeme mit unterschiedlichen Wahrscheinlichkeitsverteilungen zu einem System führt, das weniger bekannt als die ursprünglichen Systeme: der Informationsdefizit über das dritte System ist größer als der Mittelwert der Defizite für die zwei Anfangssysteme.

VI.4.2 Statistische Entropie eines quantenmechanischen Systems

Der Begriff der statistischen Entropie kann jetzt zu den Wahrscheinlichkeitsverteilungen angewandt werden, mit denen makroskopische quantenmechanische Systeme beschrieben werden.

VI.4.2 a Statistische Entropie und Dichteoperator

Sei $\hat{\rho}$ der Dichteoperator zugeordnet einem quantenmechanischen Zustandsgemisch, mit den Eigenwerten $p_1, \ldots, p_M$. Die entsprechende statistische Entropie wird durch

$$\boxed{S\big(\hat{\rho}\big) \equiv -k_B \sum_m p_m \ln p_m = -k_B \operatorname{Tr}\big(\hat{\rho} \ln \hat{\rho}\big)} \quad \text{(VI.28)}$$

definiert, wobei sich die zweite Gleichung einfach in einer Basis prüfen lässt, in der $\hat{\rho}$ diagonal ist. Dabei bezeichnet k_B die *Boltzmann*[(an)]*-Konstante*, $k_B = 1,38 \cdot 10^{-23}$ J K^{-1}.

Bemerkung: Die erste Gleichung entspricht der Definition der Entropie nach Gibbs (1878) und die zweite, der von Neumann Entropie (1927).

VI.4.2 b Eigenschaften

Die allgemeinen mathematischen Eigenschaften der statistischen Entropie lassen sich auf $S\big(\hat{\rho}\big)$ übertragen.

Minimum

Die statistische Entropie (VI.28) ist minimal und gleich 0, wenn sich das System in einem reinem Zustand befindet.
Dies bedeutet insbesondere, dass $S\big(\hat{\rho}\big)$ immer nicht-negativ ist.

Maximum

Wenn die möglichen Mikrozustände einen endlich-dimensionalen Hilbert-Raum $\mathscr{H}$ der Dimension W aufspannen, dann ist die statistische Entropie (VI.28) maximal im Fall gleich wahrscheinlicher Zustände, mit

$$\boxed{S\big(\hat{\rho}\big) = k_B \ln W,} \quad \text{(VI.29)}$$

entsprechend der Definition der Entropie nach Boltzmann.

Diese Eigenschaft, die der Charakterisierung des mikrokanonischen Ensembles zugrunde liegt, wird im § VII.3.1 hiernach bewiesen.

(Unter-)Additivität

Sei ein physikalisches System bestehend aus zwei unabhängigen Teilsystemen A und B, mit den jeweiligen Dichteoperatoren $\hat{\rho}_A$ auf dem Hilbert-Raum $\mathscr{H}_A$ und $\hat{\rho}_B$ auf $\mathscr{H}_B$. Die statistische

[(an)] L. Boltzmann, 1844–1906

Entropie des Gesamtsystems ist die Summe der statistischen Entropien der Teilsysteme

$$S(\hat{\rho}) = S(\hat{\rho}_A) + S(\hat{\rho}_B). \tag{VI.30}$$

Wenn die Teilsysteme korreliert sind, $S(\hat{\rho}) < S(\hat{\rho}_A) + S(\hat{\rho}_B)$: der Dichteoperator $\hat{\rho}$ beschreibt die Korrelationen zwischen A und B, und somit enthält mehr Information als $\hat{\rho}_A$ und $\hat{\rho}_B$ zusammen.

Konkavität

Seien auf einem Hilbert-Raum $\mathscr{H}$ zwei Dichteoperatoren $\hat{\rho}_A$, $\hat{\rho}_B$ sowie eine reelle Zahl $0 < \lambda < 1$. Dann ist

$$S\big(\lambda\hat{\rho}_A + (1-\lambda)\hat{\rho}_B\big) \geq \lambda S(\hat{\rho}_A) + (1-\lambda)S(\hat{\rho}_B), \tag{VI.31}$$

mit Gleichung nur für $\hat{\rho}_A = \hat{\rho}_B$. Die Vereinigung zweier makroskopischen gemischten Zustände in eine einzige Mischung führt zum Wachstum der „Unordnung“ gemessen durch die statistische Entropie.

VI.4.3 Statistische Entropie eines klassischen Systems

Die fehlende Information entsprechend der Wahrscheinlichkeitsdichte auf dem Phasenraum, mit der ein makroskopisches klassisches System beschrieben wird, kann ebenfalls mit der statistischen Entropie gemessen werden. Im Fall eines klassischen System mit unbekannter Teilchenzahl N lautet die statistische Entropie

$$S(\rho) = -k_B \sum_N \int \tilde{\rho}_N(\{q_i\},\{p_i\}) \ln \tilde{\rho}_N(\{q_i\},\{p_i\})\, \mathrm{d}^{6N}\mathcal{V}. \tag{VI.32}$$

Diese statistische Entropie genügt denselben Eigenschaften wie sein quantenmechanisches Pendant (VI.28), mit einer wichtigen Ausnahme. Da eine Wahrscheinlichkeitsdichte nicht unbedingt nach oben durch 1 beschränkt ist, kann $S(\rho)$ negativ sein. Zudem hat $S(\rho)$ kein Minimum: für bestimmte Phasenraumwahrscheinlichkeitsdichten kann $S(\rho)$ gegen $-\infty$ gehen. Solche Dichten erfordern aber eine gleichzeitige Kenntnis von den Teilchenpositionen und -impulsen, die durch die Heisenberg[ao]-Unschärferelation der Quantenmechanik verboten ist.

Literatur zum Kapitel VI

- Landau & Lifschitz, *Lehrbuch der Theoretischen Physik. Band V: Statistische Physik* [7], Kap. I.
- Reif, *Fundamentals of statistical and thermal physics* [10], Kap. 2.1–2.5.
- Schwabl, *Statistische Mechanik* [23], Kap. 1.

[ao] W. Heisenberg, 1901–1976