1. Grundbegriffe der klassischen und Quantenstatistik

1.1 Wahrscheinlichkeit

Maxwell: „Die wahre Logik der Welt ist die Wahrscheinlichkeitsrechnung"

Interpretation:

Frequentisten (Wahrsch. = relative Häufigkeit, „objektiv")

Bayesianer (statististische Inferenz = induktives Schließen = Erweiterung d. Logik, „subjektiv")

<u>Zufallsvariable</u> beschreibt Ausgang eines Experiments, Ergebnis einer Stichprobe etc.

<u>endlich-diskrete</u> Zufallsvariable X:

spezifiziert durch Wertevorrat $\mathbb{W} = \{x_1, \dots, x_n\}$ und eine <u>Wahrscheinlichkeitsverteilung</u> $p_1, \dots, p_n$, $p_i \geq 0$, $\sum_{i=1}^{n} p_i = 1$

$\mathbb{W}$ ... Wahrscheinlichkeitsraum (sample space)

x_i ... mögliche Messwerte (allgemeiner: Zahlen, die „Elementarereignisse" repräsentieren)

p_i ... Wahrsch. f. Messung d. Wertes x_i (bzw. Eintreten des durch x_i repräsentierten Elementarereignisses)

Teilmengen $A \subseteq \mathbb{W}$ heißen Ereignisse, sie bilden den Ereignisraum $\mathbb{E} = 2^{\mathbb{W}}$ (Potenzmenge v. $\mathbb{W}$)

Die Wahrscheinlichkeit des Ereignisses A ist

$$p(A) = \sum_{x_i \in A} p_i$$

$$\Rightarrow p(\{x_i\}) = p_i, \quad p(\mathbb{W}) = 1, \quad p(\emptyset) = 0$$

$$p(A \cup B) = p(A) + p(B), \text{ falls } A \cap B = \emptyset$$

D.h. p ist ein normiertes Maß auf $\mathbb{E}$

Erwartungswert von X: $\langle X \rangle = \sum_i p_i x_i$

z.B. symmetrischer Würfel: wählen X = Augenzahl nach Wurf

$$\Rightarrow \mathbb{W} = \{1, \ldots, 6\}, \quad p_i = \frac{1}{6}$$

z.B. $A = \{1\} \cup \{3\} \Rightarrow p(A) = \frac{1}{3}$

$$\langle X \rangle = \frac{1}{6}(1 + \cdots + 6) = \frac{1}{6} \cdot \frac{6 \cdot 7}{2} = 3{,}5$$

Ist ein Spezialfall v. Gleichverteilung ($p_i = \frac{1}{n}$).

Für diese gilt allgemein $p(A) = \dfrac{\# A}{\# \mathbb{W}}$

$$= \frac{\text{Zahl d. „günstigen" Fälle}}{\text{Zahl d. möglichen Fälle}}$$

<u>Bedingte Wahrsch.</u> $p(A|B)$ = Wahrsch. v. A, wenn B gegeben

$\Rightarrow p(A \cap B) = p(A|B)\,p(B) \quad (*)$

$\Rightarrow p(B|A) = \dfrac{p(A|B)\,p(B)}{p(A)}$ Satz v. Bayes

Erweiterung auf <u>2 Zufallsvariable</u> X, Y

Elementarereignisse (x_i, y_j), $i \in I$, $j \in J$

$\Rightarrow$ Wahrsch. raum $\mathbb{W}_X \times \mathbb{W}_Y$,

<u>gemeinsame</u> Wahrscheinlichkeitsverteilung P_{ij} („bivariate Vert.")

Zshg. mit den univariaten Vert. p_i, q_j f. X bzw. Y

folgt aus den <u>Randwahrscheinlichkeiten</u> (Marginalen)

$p(A) = P(A \times \mathbb{W}_Y)$, $q(B) = P(\mathbb{W}_X \times B)$

$\Rightarrow p_i = \sum_j P_{ij}, \quad q_j = \sum_i P_{ij}$

Zwei Zufallsvariable heißen <u>statistisch unabhängig</u>,

wenn f. jedes $A \subseteq \mathbb{W}_X$ $P(A|B)$ denselben Wert $p(A)$ $\forall B \subseteq \mathbb{W}_Y$

und " $B \subseteq \mathbb{W}_Y$ $P(B|A)$ " " $q(B)$ $\forall A \subseteq \mathbb{W}_X$

hat.

$\Rightarrow$ Die Verallg. v. $(*)$, $P(A \times B) = P(A|B)\,q(B)$

impliziert dann $P(A \times B) = p(A)\,q(B)$, insbesondere

$\underline{P_{ij} = p_i q_j}$

Eine abzählbar-diskrete Zufallsvariable ist charakterisiert durch eine Folge v. Elementarwahrscheinlichkeiten $p_1, p_2, \dots$ mit

$\sum_{i=1}^{\infty} p_i = 1 \Rightarrow$ Gleichverteilung nicht möglich

Kontinuierliche Zufallsvariable

Identifizieren $\mathbb{W}$ mit (kontinuierl. Teilmenge v.) $\mathbb{R}$

Zufallsvariable X charakterisiert durch Dichtefunktion $\rho(x) \geq 0$,

$$\int_{\mathbb{R}} \rho(x)\, dx = 1$$

Für A (Borelmenge) $\subseteq \mathbb{R}$ ist

$$p(A) = \int_A \rho(x)\, dx$$

Mehrere Zufallsvariable $X_1, \dots, X_n$: $\mathbb{W} = \mathbb{R}^n$

gemeinsame Dichtefkt $\rho(\vec{x})$, $p(A) = \int_A \rho(\vec{x})\, d^n x$, $\int_{\mathbb{R}^n} \rho(\vec{x})\, d^n x = 1$

„multivariate Verteilung"

Variable statistisch unabhängig $\Leftrightarrow \rho(\vec{x}) = \rho_1(x_1)\rho_2(x_2) \cdots \rho_n(x_n)$

Erwartungswert einer Fkt. $f : \mathbb{W} \to \mathbb{R}$

$$\langle f(X) \rangle = \int d^n x\, f(\vec{x})\, \rho(\vec{x})$$

z. B. <u>Momente</u> einer Zufallsvar. X definiert durch $\langle X^n \rangle$
(müssen nicht existieren!)

$n = 1$: $\mu := \langle X \rangle$ identisch mit Mittelwert v. X → Mittelwertsatz

$n = 2$: $Var(X) := \langle (X-\mu)^2 \rangle$ <u>Varianz</u> (Schwankungsquadrat)

$\sigma(x) = \sqrt{Var(X)}$ heißt <u>Standardabweichung</u>,
wird bei Messdaten in Form eines Fehlerintervalls
(grafisch: Fehlerbalken) angegeben,
z. B. $1{,}234 \pm 0{,}056 = 1{,}234(56)$

Die Verteilung wird aus einem Histogramm durch „Binning" und „Fitting" bestimmt. Meist kommt sie einer Gauß-Verteilung nahe (wenn nicht, vermutet man einen „systematischen Fehler".
Grund: → zentraler Grenzwertsatz

<u>Gauß- oder Normalverteilung</u>

$$\rho_G(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

<u>statist. Sicherheit</u> = Wahrsch., dass eine Stichprobe im <u>Vertrauensintervall</u> $(\mu - n\sigma, \mu + n\sigma)$ liegt. Ihr Wert f. $n = 1$: 68%, $n = 2$: 95%, $n = 3$: 99,7%, $n = 4$: $1 - 6 \cdot 10^{-5}$, $n = 5$: $1 - 6 \cdot 10^{-7}$. Komplementär dazu ist die <u>statistische Signifikanz</u> einer Messung außerhalb d. Vertrauensintervalls einer <u>Nullhypothese</u>. Teilchenphysik: Messresultat gilt als Entdeckung, wenn es auf 5σ-Niveau nicht mit Nullhypothese verträglich ist

<u>Kovarianz</u> $\mathrm{Cov}(X,Y) := \langle (X - \langle X\rangle)(Y - \langle Y\rangle)\rangle = \langle XY\rangle - \langle X\rangle\langle Y\rangle$

<u>Korrelation</u> v. X_i u. X_j

$$\mathrm{Cor}(X_i, X_j) = \frac{\mathrm{Cov}(X_i, X_j)}{\sqrt{\mathrm{Var}(X_i)\,\mathrm{Var}(X_j)}}$$

$$\Rightarrow -1 \leq \mathrm{Cor}(X_i, X_j) \leq 1$$

antikorreliert perfekt korreliert

X_i, X_j statistisch u. a. $\Rightarrow \mathrm{Cov}(X_i, X_j) = \mathrm{Cor}(X_i, X_j) = 0$

Der Umkehrschluss gilt nicht!

charakteristische Funktion einer Verteilung (inverse)

$$G(k) = \langle e^{ikX}\rangle = \int dx\, \rho(x)\, e^{ikx}$$ = Fouriertr. d. Dichtefkt., existiert immer!

Wenn alle Momente existieren, ist

$$G(k) = \sum_{n=0}^{\infty} \frac{(ik)^n}{n!}\langle X^n\rangle = 1 + \sum_{1}^{\infty}\frac{(ik)^n}{n!}\langle X^n\rangle$$

$$\ln G(k) = \sum_{n=1}^{\infty}\frac{(ik)^n}{n!}\kappa_n$$ mit den <u>Kumulanten</u>

$\kappa_1 = \langle X\rangle = \mu$

$\kappa_2 = \langle X^2\rangle - \langle X\rangle^2 = \mathrm{Var}(X)$

$\kappa_3 = \langle X^3\rangle - 3\langle X^2\rangle\langle X\rangle + 2\langle X\rangle^3$

usw.

z. B. Gauß-Vert.: $G(k) = e^{i\mu k - \frac{1}{2}\sigma^2 k^2}$

$\Rightarrow \kappa_1 = \mu, \quad \kappa_2 = \sigma^2, \quad \kappa_n = 0$ f. $n \geq 3$

(vgl. QFT: Momente $\to$ n-Pt.-Funktionen, Kumulanten $\to$ zusammenhängende n-Pt.-Funktionen)

Grenzwertsätze

Der zentrale Grenzwertsatz (einfachste Version)

Seien X_i, $i = 1, \dots, N$, unabhängige und identisch verteilte Zufallsvariable. Alle Kumulanten mögen existieren u. $\langle X_i \rangle = 0$, $\mathrm{Var}(X_i) = \sigma^2$

Sei $Z_N = \frac{1}{\sqrt{N}}(X_1 + \dots + X_N)$. Dann $\exists$ $\lim_{N\to\infty} Z_N$ und ist eine Gauß'sche Zufallsvariable mit Mittel 0 u. Varianz σ^2.

Bew.: $\langle Z_N \rangle = 0$, $\mathrm{Var}(Z_N) \underset{\uparrow}{=} \frac{1}{N} \sum_{i=1}^{N} \mathrm{Var}(X_i) = \sigma^2$

$\langle X_i X_j \rangle = 0$ f. $i \neq j$

und alle höheren Kumulanten fallen f. $N \to \infty$ mindestens so rasch wie $N^{-1/2}$ ab. (κ_3: $\frac{1}{N^{3/2}} \cdot N \langle X_i^3 \rangle$) *)

$\exists$ allgemeinere Versionen dieses Satzes.

Bedeutung: Viele „kleine" zufällige Einflüsse summieren sich zu einem Gauß-verteilten Gesamteinfluss, erklärt „Natürlichkeit" d. Gauß-Verteilung

*) bei höheren Kumulanten ist wesentlich, dass Terme einander aufheben

<u>Mittelwertsatz</u> (Gesetz der großen Zahlen)

Sei X eine Zufallsvariable mit existierenden Momenten. Dann liefert der Mittelwert von N Messungen im Limes $N \to \infty$ <u>exakt</u> den Erwartungswert μ.

<u>Bew.</u>: Seien $X_1, \dots, X_N$ <u>unabhängige</u> Zufallsvariable mit <u>derselben</u> Verteilung wie X. Sei $Z_N = \frac{1}{N}(X_1 + \dots + X_N)$.

Dann ist $\langle Z_N \rangle = \frac{1}{N} \sum_{i=1}^{N} \langle X_i \rangle = \langle X \rangle$,

$$\mathrm{Var}(Z_N) = \frac{1}{N^2} \underbrace{\left[\langle (\sum X_i)^2 \rangle - (\sum \mu_i)^2 \right]}_{N\langle x^2 \rangle + \cancel{\sum_{i \neq j} \langle X_i X_j \rangle} - N\mu^2 - \cancel{\sum_{i \neq j} \mu_i \mu_j}} = \frac{1}{N} \mathrm{Var}(X)$$

und alle höheren Kumulanten fallen f. $N \to \infty$ zumindest so rasch wie N^{-2} ab.

$\Rightarrow Z_N$ wird f. $N \to \infty$ Gauß'sch mit Erwartungswert μ und Varianz $\to 0$, d. h. determiniert.

1.2 Statistische Entropie

= Maß f. Informationsdefizit einer Verteilung.

Einfachste Begründung:

Informationswert (Überraschung) eines Ereignisses mit Wahrsch. p

ist eine Fkt. $s(p)$ mit den Eigenschaften:

(i) s monoton fallend in $(0,1]$

(ii) $s(1) = 0$

(iii) Die Informationswerte statistisch u.a. Ereignisse addieren sich

$$P(A \times B) = p(A)q(B) \Rightarrow \underline{s(pq) = s(p) + s(q)}$$

$\Rightarrow s(p) = -k \ln p$ mit einer Konstante $k > 0$

Folgerung: $s(0) = \infty$ ✓

Def. statistische Entropie einer diskreten Verteilung :=

= mittlerer Informationswert eines Elementarereignisses, also

$$\underline{S(p_1, \ldots, p_n) := -k \sum_i p_i \ln p_i} \quad , \quad k > 0 \quad \text{(Shannon 1948)}^{*)}$$

$\Rightarrow S = 0$, falls ein $p_i = 1$ $\quad (0 \cdot \ln 0 = 0)$

$S = k \ln n$ f. Gleichverteilung $p_i = \frac{1}{n}$,

ist Maximalwert f. n-wertige Zufallsvariable

---

*) Die ursprüngliche Charakterisierung v. S war komplizierter

Informationstheoret. Bedeutung v. S

Wähle $k = \frac{1}{\ln 2} \Rightarrow S = -\sum_i p_i \log_2 p_i$

$\approx$ Minimalzahl v. binären (ja/nein) Fragen, die man bei geg. p_i stellen muss, um herauszufinden, welches Elementarereignis bei einer Messung eingetreten ist = Zahl d. Bits (binary digits = Binärziffern), die zur Mitteilung d. Messergebnisses benötigt

Offensichtlich im Fall der Gleichverteilung, z. B. ein Teilchen in einer von 2^n Zellen: Nummeriere die Zellen v. 1 bis 2^n, frage sukzessive bez. d. Hälfte mit den kleineren Nummern, notiere 0 bzw. 1 bei Antwort ja bzw. nein $\rightarrow$ Binärdarstellung der besetzten Zelle: n Bits

Fragestrategie bei allgemeiner Verteilung arbeitet mit dem besten Ersatz f. Halbierung im Sinn d. Wahrscheinlichkeitsmaßes

Statist. Entropie als Maß f. Wahrscheinlichkeit einer Verteilung

Benötigen Stirling-Formel $\lim_{N\to\infty} \frac{\sqrt{2\pi N}\left(\frac{N}{e}\right)^N}{N!} = 1$ (Ü)

$\Rightarrow$ f. große N gilt $N! \approx \sqrt{2\pi N}\left(\frac{N}{e}\right)^N$ in dem Sinne, dass der Quotient aus linker u. rechter Seite f. $N \to \infty$ gegen 1 geht

Im selben Sinne folgt $\ln N! \approx N(\ln N - 1)$

Ein Experiment mit n möglichen Resultaten werde N-mal ausgeführt. (bzw. N unterscheidbare Teilchen auf n Zellen aufgeteilt oder in N Kopien eines System jeweils einer von n möglichen Zuständen besetzt). Alle Resultate (Zellen, Zustände) mögen dieselbe Wahrscheinlichkeit (= a priori Wahrscheinlichkeit) $q_i = \frac{1}{n}$ haben.

Sei $N_{\{N_i\}}$ Zahl d. möglichen Realisierungen d. Serie $\{N_1, .., N_n\}$ (auf die Reihenfolge soll es nicht ankommen),

z. B. $N_{\{N, 0, .. 0\}} = 1$, allgemein ist

$$N_{\{N_i\}} = \frac{N!}{N_1! N_2! .. N_n!} \sim \left(\frac{N}{N_1}\right)^{N_1} \cdots \left(\frac{N}{N_n}\right)^{N_n}, \quad \text{da } \sum_i N_i = N$$

Haben Faktor $\sqrt{2\pi N}$ in Stirling-Formel vernachlässigt, da wir die N-te Wurzel ziehen und außerdem logarithmieren werden.

Zahl d. möglichen Realisierungen <u>aller</u> Serien

$$\sum_{N_1, .., N_n} N_{\{N_i\}} = n^N \quad (\text{Multinomialsatz f. } (\underbrace{1 + \cdots + 1}_{n})^N)$$

$\Rightarrow$ „Metawahrscheinlichkeit" f. Verteilung $\{p_i = \frac{N_i}{N}\}$

ist
$$\frac{N_{\{N_i\}}}{n^N} \equiv P_N(p_1, .., p_n) \sim \frac{1}{n^N} \cdot \frac{1}{p_1^{N_1} \cdots p_n^{N_n}}$$

$$= \frac{1}{(n p_1)^{N_1} \cdots (n p_n)^{N_n}} \xrightarrow{N \to \infty} 1 \quad \text{f. } p_i = q_i = \frac{1}{n}$$

$\Rightarrow P_N \to 0$ sonst (wurde hier nicht exakt bewiesen)

Wie zu erwarten: Die Verteilung d. gemessenen Resultate entspricht im Limes $N \to \infty$ mit Wahrscheinlichkeit 1 der a priori-Wahrscheinlichkeitsverteilung.

Interessanter ist $S(\{p_i\}) = \frac{k}{N} \ln N_{\{N_i\}}$

$$\approx -k \sum_i p_i \ln p_i = \frac{k}{N} \ln\left(n^N P_N(p_1, \ldots, p_n)\right) = \frac{k}{N} \ln P_N + k \ln n$$

ist bei festem N <u>monoton wachsende</u> Fkt. v. P_N

$\Rightarrow$ statist. Entropie ist Maß f. Wahrscheinlichkeit einer Verteilung in einer großen ($N \gg 1$) Serie v. Messungen bei gleicher a priori-Wahrsch. d. Elementarereignisse.

Insbesondere ist die Verteilung maximaler Entropie die wahrscheinlichste Verteilung.

Nur zur Charakterisierung der Gleichverteilung hat sich dieser Aufwand nicht gelohnt. Aber:

1. Nebenbedingungen können ein anderes Maximum als die Gleichverteilung erzwingen $\to$ Ü

2. Als a priori - Verteilung $\{q_i\}$ kann statt der Gleichverteilung eine andere gewählt werden. Auch dann gilt

$P_N(q_1, \dots, q_n) \xrightarrow{N\to\infty} 1$: äquivalente Formulierung d. Gesetzes d. großen Zahlen, gleichwertig mit Mittelwertsatz

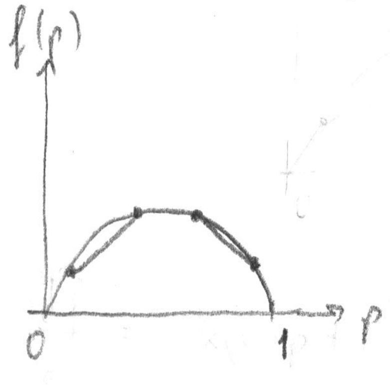
Das führt auf den Begriff d. relativen Entropie einer Verteilung $\{p_i\}$ bez. einer a-priori-Verteilung $\{q_i\}$ → Ü

Konkavität d. statist. Entropie

Sind $\{p_i\}$, $\{q_i\}$ zwei Verteilungen gleicher Mächtigkeit, dann gilt

$$S(\{\lambda p_i + (1-\lambda) q_i\}) \geq \lambda S(\{p_i\}) + (1-\lambda) S(\{q_i\})$$

Bew.: Es genügt zu zeigen, dass die Funktion $f(p) = -k\, p \ln p$ die Ungleichung $f(\lambda p + (1-\lambda) q) \geq \lambda f(p) + (1-\lambda) f(q)$ f. bel. $p, q \in (0,1)$ erfüllt. Sie bedeutet: Der Funktionsgraph v. f liegt im Intervall $(0,1)$ oberhalb jeder Sehne.*)



Das wird garantiert durch $f'' < 0$

Tatsächlich ist $f' = -k(\ln p + 1) \Rightarrow$

$f'' = -\frac{k}{p} < 0.$

*) ⟺ unterhalb jeder Tangente

Anschauliche Bedeutung d. <u>Konkavität v. S</u>:

Betr. 2 Gemische aus jeweils n Reinsubstanzen mit normierten Konzentrationen p_i bzw. q_i. Mische die beiden Gemische im Verhältnis $\lambda : 1-\lambda$ $\to$ Das Informationsdefizit (bez. d. Natur eines willkürlich entnommenen Teilchens) kann nur zunehmen: „Mischen macht gemischter"

<u>Statist. Entropie einer kontinuierlichen Zufallsvariablen</u>

Entropiefunktional $S[\rho] := -k \int dx\, \rho(x) \ln \rho(x)$

analog f. n-dim. Verteilung

$S[\rho]$ kann auch negativ sein ($\to$ Ü)

nicht invariant unter Reparametrisierung $x \mapsto y(x) \Rightarrow$

$\tilde{\rho}(y) = \frac{dx}{dy} \rho(x(y)) \leadsto$

Die natürlichere Version der statist. Entropie ist die relative Entropie $S_{rel}[\rho, \rho_0] = -k \int dx\, \rho(x) \ln \frac{\rho(x)}{\rho_0(x)}$

Betrachtet man Verteilungen auf ganz $\mathbb{R}$, divergiert $S_{rel} \to -\infty$ im Grenzfall der „Gleichverteilung" $\rho_0 \to 0$ (tatsächlich $\nexists$ auf $\mathbb{R}$ keine Gleichverteilung). Das Problem tritt nicht auf, wenn man Verteilungen auf einem <u>endlichen</u> Intervall betrachtet.

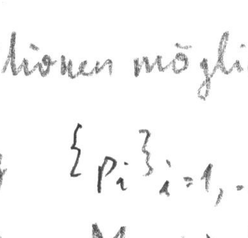
1.3 Klassische Zustände

<u>isoliertes System</u> besteht aus Teilchen, deren Dynamik festgelegt durch zeitunabh. klass. <u>Hamiltonfkt.</u> bzw. QM <u>Hamiltonoperator</u> (einschließlich evtl. <u>Randbedingungen</u>)

<u>Zustand</u> eines Systems = Darstellung d. Information über die tatsächliche <u>Geschichte</u> d. Systems (gleichwertig: <u>Anfangsdaten</u>, da die Dynamik aus diesen die Geschichte bestimmt)

<u>Klassische</u> Anfangsdaten sind i. a. Zufallsvariable, einfachster Fall: eine diskrete Zufallsvariable (z. B. nur endlich viele Teilchenpositionen möglich) $\rightarrow$ Zustand spezifiziert durch Verteilung $\{p_i\}_{i=1,\dots,n} \equiv \vec{p}$, formaler Koordinatenvektor eines <u>Punkts</u> in einer <u>Mannigfaltigkeit</u>: welche?

$p_i \geq 0$, $\sum_1^n p_i = 1$ $\Rightarrow$ die Punkte $\vec{p}$ bilden ein <u>Simplex</u> in einem affinen Raum: abgeschlossen bez. konvexer Linearkombinationen, d. h. $\vec{p}(\lambda) = \lambda \vec{p}_1 + (1-\lambda)\vec{p}_2$, $0 \leq \lambda \leq 1$ ist wieder ein Zustand

z. B. $n = 3$:  2-Simplex = Dreieck

(Simplex = affiner Begriff, benötigt keine Metrik; natürliche Metrik ist Fisher-Informationsmetrik, die Simplex zum <u>Orthanten</u> (verallg. Quadranten) einer Sphäre macht)

Die Eckpunkte d. Simplex sind Zustände maximaler Information (ein $p_i = 1 \Rightarrow S(\vec{p}) = 0$). Solche Zustände heißen extremal oder <u>rein</u>, alle anderen gemischt.

<u>Äquivalente Def.</u>: Reiner Zustand := nicht als nichttriviale konvexe Linearkombination (zweier) anderer Zustände darstellbar

Die reinen Zustände bilden als „Einheitsvektoren" eine „Basis" (bez. konvexer Linearkomb.) f. alle Zustände $\Rightarrow$ <u>jeder</u> Zustand <u>eindeutig</u> als konvexe Linearkomb. <u>reiner</u> Zustände darstellbar; gilt nur f. klassische Zustände, nicht in QT!

<u>Klassisches System mit n Freiheitsgraden</u>

Anfangsdaten = Punkte im Phasenraum mit Koordinaten

$(q_1, .., q_n, p_1, .., p_n) \equiv (\vec{q}, \vec{p})$ ($\vec{p}$ hat hier eine andere Bed. als vorhin!)

sind Zufallsvariable $\rightarrow$ <u>Zustand</u> spezifiziert durch

Dichtefkt. $\rho(\vec{q}, \vec{p}) \geq 0$, $\int d^n q\, d^n p\, \rho = 1$

Reine Zustände: $\rho(\vec{q}, \vec{p}) = \delta^{(n)}(\vec{q} - \vec{q}_0)\,\delta^{(n)}(\vec{p} - \vec{p}_0)$

bilden 2n-dim. Raum ($\cong$ Phasenraum)

gemischte Zstde. bilden offenbar ∞-dim. Funktionenraum

Machen Zustand <u>zeitabhängig</u>, um auf diese Weise Zeitentwicklung des Systems darzustellen, d.h. bilden Systemgeschichte auf die Fkt. $\varrho(\vec{q},\vec{p},t)$ ab (Alternative: zeitabhängige Observable).

Wahrsch., das System zu t im Phasenraumgebiet G anzutreffen:

$$P_t(G) = \int_G d^nq\, d^np\, \varrho(\vec{q},\vec{p},t)$$

Dynamik liefert <u>Evolutionsgleichung</u> f. ϱ (alternative Herleitung → Ü)

Erhaltung d. Wahrscheinlichkeit ⇒ (vgl. $\frac{\partial\varrho}{\partial t} + \mathrm{div}(\varrho\vec{v}) = 0$)

<u>Kontinuitätsgl.</u> $$\frac{\partial\varrho}{\partial t} + \sum_i \frac{\partial}{\partial q_i}(\varrho\dot{q}_i) + \sum_i \frac{\partial}{\partial p_i}(\varrho\dot{p}_i) = 0$$

Hamiltonsche Bew. gl.: $\dot{q}_i = \frac{\partial H}{\partial p_i}$, $\dot{p}_i = -\frac{\partial H}{\partial q_i}$ sind Fktn. v. $(\vec{q},\vec{p})$

$$\Rightarrow \frac{\partial\varrho}{\partial t} + \sum_i \frac{\partial\varrho}{\partial q_i}\cdot\dot{q}_i + \sum_i \frac{\partial\varrho}{\partial p_i}\dot{p}_i = 0$$ oder $\frac{d\varrho}{dt} = 0$ (totale od. konvektive Ableitung)

oder $\frac{\partial\varrho}{\partial t} = -\{\varrho, H\}$ <u>Liouville-Gleichung</u> *)

<u>Observable</u> = Funktion $O(\vec{Q},\vec{P})$

$\Rightarrow$ Erwartungswert $\langle O\rangle\big|_t = \int d^nq\, d^np\, \varrho(\vec{q},\vec{p},t)\, O(\vec{q},\vec{p})$

Statist. Entropie d. Zustands ϱ:

$S[\varrho] = -k\int d^nq\, d^np\, \varrho \ln\varrho$ ist <u>zeitunabhängig</u>:

---

*) ⇒ ϱ konstant entlang Teilchentrajektorien im Phasenraum. Die Trajektorien selbst verhalten sich wie die einer inkompressiblen Fl. im Ortsraum

Offenbar ist $\frac{d}{dt}(\rho \ln \rho) = (\rho \ln \rho)' \cdot \frac{d\rho}{dt} = 0 \Rightarrow$

$$\frac{\partial}{\partial t}(\rho \ln \rho) = \sum_i \left[-\frac{\partial}{\partial q_i}(\rho \ln \rho) \frac{\partial H}{\partial p_i} + \frac{\partial}{\partial p_i}(\rho \ln \rho) \frac{\partial H}{\partial q_i} \right]$$

$$\Rightarrow \int d^n q \, d^n p \, \frac{\partial}{\partial t}(\rho \ln \rho) = \int d^n q \, d^n p \, \rho \ln \rho \sum_i \left(\frac{\partial^2 H}{\partial q_i \partial p_i} - \frac{\partial^2 H}{\partial q_i \partial p_i} \right)$$

$$= 0$$

Konnten partiell integrieren, weil $\rho \ln \rho \to 0$ im ∞, da ρ integrierbar

$\Rightarrow$ Für ein isoliertes System ist die statistische Zustandsentropie konstant

Das muss so sein, weil $\rho(\vec{q}, \vec{p}, t_2)$ auf Grund d. Bew.gl. durch $\rho(\vec{q}, \vec{p}, t_1)$ festgelegt ist und daher dieselbe Information enthält

1.4 Quantenzustände

klassische Observable $\hat{=}$ reelle Funktion $O(\vec{q},\vec{p})$ auf dem Phasenraum $\mathbb{W}$

Quantenobservable $\hat{=}$ s.a. Element $\hat{O}$ einer nichtkommutativen Algebra, typischerweise erzeugt von Elementen $\hat{\vec{q}}, \hat{\vec{p}}$ mit $[\hat{q}_i, \hat{p}_j] = i\hbar\delta_{ij}$,

Quantenobservable können als Operatoren in einem Hilbertraum $\mathcal{H}$ dargestellt werden

Klass. Zustand $\hat{=}$ Funktional auf dem Raum d. Funktionen auf $\mathbb{W}$:

$$f \mapsto \langle f(\vec{Q},\vec{P})\rangle_\rho = \int d^nq\, d^np\, \rho(\vec{q},\vec{p})\, f(\vec{q},\vec{p})$$

Eigenschaften:

(i) linear: $\langle \alpha f + \beta g\rangle = \alpha\langle f\rangle + \beta\langle g\rangle$

(ii) positiv: $\langle f\rangle \geq 0$ falls $f(x) \geq 0$

(iii) normiert: $\langle 1\rangle = 1$

Analog wird ein Quantenzustand repräsentiert durch ein lineares, positives, normiertes Funktional auf der Observablen-algebra: $\hat{A} \mapsto \langle\hat{A}\rangle$ mit

(i) $\langle \alpha\hat{A} + \beta\hat{B}\rangle = \alpha\langle\hat{A}\rangle + \beta\langle\hat{B}\rangle$

(ii) $\hat{A} \geq 0$ (d.h. $\langle\psi|\hat{A}|\psi\rangle \geq 0\ \forall\psi$) $\Rightarrow \langle\hat{A}\rangle \geq 0$

(iii) $\langle \mathbb{1}\rangle = 1$

z. B. $\dim \mathcal{H} = n < \infty$

Observable $\hat{=}$ hermitische $n \times n$-Matrix A

Allgemeinster Zustand $A \mapsto \langle A \rangle$?

linear: $\langle A \rangle = \sum_{i,j} \rho_{ij} A_{ji} \equiv \mathrm{Tr}(\rho A)$ (Spur)

reell*): $\langle A \rangle^* = \sum \rho_{ij}^* A_{ji}^* = \sum \rho_{ij}^* A_{ij} = \sum \rho_{ji}^* A_{ji} \overset{!}{=} \langle A \rangle$

$\Rightarrow \sum_{i,j} (\rho_{ij} - \rho_{ji}^*) A_{ji} = 0 \quad \forall$ hermitischen A

wähle z. B. $A_{12} = A_{21} = 1$, $A_{ij} = 0$ sonst

bzw. $A_{12} = -A_{21} = i$, $A_{ij} = 0$ sonst

$\Rightarrow \rho_{12} = \rho_{21}^*$, analog $\rho_{ij} = \rho_{ji}^* \quad \forall i,j \Rightarrow$ ρ hermitisch

positiv: wähle $A_{ij} = b_i b_j^* \overset{(ii)}{\Rightarrow} \vec{b}^\dagger \rho \vec{b} \geq 0 \quad \forall \vec{b}$

$\Rightarrow$ ρ positiv

normiert: $\langle 1\!\!1 \rangle = \rho_{ij} \delta_{ij} = 1 \Rightarrow \mathrm{Tr}\,\rho = 1$

$\Rightarrow$ allgemeinstes Zustandsfunktional hat Form

$$\langle A \rangle_\rho = \mathrm{Tr}(A\rho), \quad \rho \text{ hermitisch}, \geq 0, \ \mathrm{Tr}\,\rho = 1$$

---

*) folgt aus positiv, weil jedes hermitische $A = A_+ - A_-$ mit $A_+ \geq 0$ und $A_- \geq 0$.

ϱ heißt <u>Dichtematrix</u> (od. statistischer Operator)
(weil sie den Quantenzustand analog charakterisiert wie die Dichtefunktion einen klassischen Zustand)
Falls $\dim \mathcal{H} = \infty$, muss ϱ formal dieselben Eigenschaften haben, insbesondere ein Operator der Spurklasse sein (nur für diese ist die Spur definiert).
Mit ϱ_1, ϱ_2 ist auch $\lambda \varrho_1 + (1-\lambda)\varrho_2$, $\lambda \in [0,1]$, eine Dichtematrix. Ein Zustand heißt <u>rein</u>, wenn er keine nichttriviale Zerlegung dieser Art hat. Die reinen Zustände sind genau jene mit $\mathrm{Rang}(\varrho) = 1$. Das ist in der Diagonaldarstellung von ϱ offensichtlich:

$$\text{z.B.} \begin{pmatrix} \varrho_1 & & 0 \\ & \varrho_2 & \\ 0 & & 0 \ddots \end{pmatrix} = \varrho_1 \begin{pmatrix} 1 & & 0 \\ & 0 & \\ 0 & & 0 \ddots \end{pmatrix} + \varrho_2 \begin{pmatrix} 0 & & \\ & 1 & 0 \\ 0 & & 0 \ddots \end{pmatrix}$$

$\Rightarrow$ reines ϱ ist hermitischer (orthogonaler) Projektor,
also $\underline{\varrho = |\psi\rangle\langle\psi|}$, $\langle\psi|\psi\rangle = 1$, $e^{i\alpha}|\psi\rangle$ liefert dasselbe ϱ

$\Rightarrow$ Auch f. endliches n $\exists$ ein <u>Kontinuum</u> von <u>reinen</u> Zuständen, also viel mehr als im klassischen Fall.
Dichtematrizen bilden konvexen Körper, seine extremen Punkte sind die reinen Zustände. Nicht reine Zustände heißen <u>gemischt</u>

$|\psi\rangle$ heißt der Zustandsvektor d. reinen Zustands $\rho_\psi = |\psi\rangle\langle\psi|$

Es gilt $\langle A\rangle_{\rho_\psi} = \mathrm{Tr}\,(A|\psi\rangle\langle\psi|) = \mathrm{Tr}\,(\langle\psi|A|\psi\rangle) = \langle\psi|A|\psi\rangle$,

Zyklizität d. Spur gilt auch f. Produkt v. Rechteckmatrizen

was mit der üblichen Definition des Erwartungswerts v. A „im Zustand $|\psi\rangle$" übereinstimmt.

(explizit: $\sum_{i=k,j} A_{ij}\psi_j\psi_k^* = \sum_{i,j}\psi_i^* A_{ij}\psi_j$)

<u>Beispiel</u>: allgemeinste Dichtematrix f. $n=2$ (beschreibt Spin $\frac{1}{2}$-Teilchen od "qubit")

$\rho = \frac{1}{2}(\mathbb{1} + \vec{r}\cdot\vec{\sigma})$, $0 \le |\vec{r}| \le 1$

$\Rightarrow$ Zustandsmf. $\hat{=}$ 3-dim. Einheitskugel (Bloch-Kugel)

$\vec{r} = 0$: $\rho = \frac{1}{2}\mathbb{1}$ „Spurzustand" $\Rightarrow \mathrm{Tr}(A\rho) = \frac{1}{2}\,\mathrm{Tr}\,A$, $\hat{=}$ Gleichverteilung

reine Zustände : $|\vec{r}| = 1 \Rightarrow \vec{r} = (\sin\vartheta\cos\varphi, \sin\vartheta\sin\varphi, \cos\vartheta)$

$0 \le \vartheta \le \pi$, entspricht dem Zustandsvektor

$$|\psi\rangle = \cos\frac{\vartheta}{2}|\uparrow\rangle + e^{i\varphi}\sin\frac{\vartheta}{2}|\downarrow\rangle, \qquad |\uparrow\rangle \equiv \begin{pmatrix}1\\0\end{pmatrix},\ |\downarrow\rangle \equiv \begin{pmatrix}0\\1\end{pmatrix}$$

$$\Rightarrow |\psi\rangle\langle\psi| = \begin{pmatrix} \cos^2\frac{\vartheta}{2}, & \cos\frac{\vartheta}{2}\sin\frac{\vartheta}{2}(\cos\varphi - i\sin\varphi) \\ \cos\frac{\vartheta}{2}\sin\frac{\vartheta}{2}(\cos\varphi + i\sin\varphi), & \sin^2\frac{\vartheta}{2} \end{pmatrix}$$

$$= \frac{1}{2}(\mathbb{1} + \cos\vartheta\,\sigma_3 + \sin\vartheta\cos\varphi\,\sigma_1 + \sin\vartheta\sin\varphi\,\sigma_2)$$

$\Rightarrow$ Reine Zustände liegen auf d. Einheitssphäre (Bloch-Sphäre), bilden also 2-dim. Kontinuum.

Vgl. klassischen Fall $n=2$: 1-Simplex = Intervall $\vdash\!\!\dashv$, nur 2 reine Zustde. = Endpunkte

Dichtematrix ist selbst keine Wahrscheinlichkeitsverteilung, sondern definiert eine solche f. <u>jede</u> Observable A:

$A = \sum a_i |\psi_i\rangle\langle\psi_i| \Rightarrow p_i = \mathrm{Tr}(|\psi_i\rangle\langle\psi_i|\varrho)$ ist die Wahrscheinlichkeit, im Zustand v. ϱ den Wert a_i d. Observablen zu messen (ist $\varrho = |\psi\rangle\langle\psi|$ rein, dann ist $p_i = |\langle\psi|\psi_i\rangle|^2$ die Übergangswahrsch. $|\psi\rangle \to |\psi_i\rangle$). Allerdings $\nexists$ f. nichtkommutierende Observable keine <u>gemeinsame</u> Wahrscheinlichkeitsverteilung $\Rightarrow$ QM kann nicht durch einen klassischen Wahrscheinlichkeitsraum beschrieben werden.

Im Gegensatz zu klass. gemischten Zuständen ist die Zerlegung eines gemischten Quantenzustands in reine Zustände nicht eindeutig! Bilden $|e_i\rangle$ eine ONB v. normierten Eigenzustdn. v. ϱ, dann ist $\varrho = \sum_{i=1}^{n} \varrho_i |e_i\rangle\langle e_i|$, aber $\exists$ ∞ viele andere Darstellungen $\varrho = \sum_{i=1}^{\times} \lambda_i |\psi_i\rangle\langle\psi_i|$, wobei die Zustandsvektoren $|\psi_i\rangle$ weder orthogonal noch von endlicher Anzahl sein müssen, selbst wenn $\dim \mathcal{H} = n < \infty$!

Dichtematrizen und Nichtklassizität

ϱ ordnet jeder ONB $\{e_i\}$ eine Wahrsch. verteilung $p_i = \mathrm{Tr}(|e_i\rangle\langle e_i|\varrho)$ zu. Wg. der Orthogonalität d. 1-dim. Projektoren $|e_i\rangle\langle e_i|$ kommutieren diese, sind daher unabhängig messbar und die p_i experimentell bestimmbar. Ein Zustand $|e\rangle$ kann Element von ∞ vielen ONB sein. Dennoch hat $|e\rangle\langle e|$ eine eindeutige Wahrsch. $p_e = \mathrm{Tr}(|e\rangle\langle e|\varrho)$ bei geg. ϱ. $\Rightarrow$ Die Wahrscheinlichkeitszuweisung durch ϱ ist nicht kontextuell, d.h. u.a. davon, welche Zustände außer $|e\rangle$ noch von der Messung miterfasst sind. Wie stark ist die Forderung nicht-kontextueller Wahrscheinlichkeiten alleine? Antwort gibt das Theorem v. Gleason: F. $n > 2$ ist die Annahme, dass allen Projektoren nicht-kontextuelle Wahrscheinlichkeiten zugeschrieben werden können, äquivalent zur Existenz einer Dichtematrix, so dass $p_e = \mathrm{Tr}(|e\rangle\langle e|\varrho)$, d.h. alle Möglichkeiten d. Zuweisung nicht-kontextueller Wahrsch. werden durch Dichtematrizen erfasst. $\rightarrow$ Starke Einschränkung an Theorien mit „verborgenen Variablen".

Ist es möglich, allen Projektoren <u>Wahrheitswerte</u> (d.h. nur Wahrsch. 0 od. 1) zuzuschreiben, d.h. gibt es ein ρ, so dass f. beliebige ONB $\{e_i\}$ die p_{e_i} deterministisch sind? Nein! „Nicht einmal Gott kann das Resultat jeder Messung <u>wissen</u>, einschließlich solcher, die gemacht hätten werden können, aber nicht gemacht wurden."*) Einen direkten Beweis liefert das <u>Theorem v. Kochen u. Specker</u>; das für $n > 2$ eine endliche Menge von Strahlen konstruiert, für die es keine Wahrheitswerte gibt.

<u>Zeitentwicklung</u>

<u>Schrödingerbild</u>: zeitabhängige Dichtematrix $\rho(t)$

$\rho = \sum \rho_i |e_i\rangle\langle e_i| \Rightarrow$ Zeitentwicklung wie die von $|\psi\rangle\langle\psi|$:

$$\frac{d}{dt}|\psi\rangle\langle\psi| = -\frac{i}{\hbar}H|\psi\rangle\langle\psi| + |\psi\rangle\langle\psi|\frac{i}{\hbar}H = -\frac{i}{\hbar}[H, |\psi\rangle\langle\psi|]$$

$$\Rightarrow \frac{d}{dt}\rho(t) = -\frac{i}{\hbar}[H, \rho(t)] \Rightarrow \rho(t) = e^{-\frac{i}{\hbar}Ht}\rho(0)\, e^{\frac{i}{\hbar}Ht}$$

(Bloch-Heisenberg-Gl., von Neumann-Gleichung)

vgl. <u>Heisenbergbild</u>: $\rho = \text{const}$, $\frac{d}{dt}O = \frac{i}{\hbar}[H, O]$

$$\Rightarrow O(t) = e^{\frac{i}{\hbar}Ht}\, O(0)\, e^{-\frac{i}{\hbar}Ht}$$

$\langle O\rangle|_t = \mathrm{Tr}(\rho O)$ ist in beiden Bildern dasselbe.

---

*) Im klassischen Fall ist das möglich („Laplace-Dämon")

Quantenmechanische Entropie

ϱ ordnet jeder ONB $\{e_i\}$ eine Verteilung $\{p_i\}$ zu, für die sich eine statist. Entropie $S(\{p_i\})$ definieren lässt.

Fordern $S[\varrho] = 0$ f. ϱ rein, (da reine Zustände die Zustände mit maximaler Information sind. Auch f. einen reinen Zustand können die p_i nicht deterministisch, sogar gleichverteilt sein =>

allg. Def. d. quantenmech. Entropie:

$$S[\varrho] = \min_{\text{ONB}\,\{e_i\}} S(\{p_i\})$$

Das Minimum wird angenommen, wenn die e_i Eigenvektoren v. ϱ sind =>

$$\underline{S[\varrho] = -k\sum \varrho_i \ln \varrho_i = -k \operatorname{Tr} \varrho \ln \varrho} \quad \text{von Neumann-Entropie}^{*)}$$

<u>Beispiel: Informationsverlust durch Verschränkung</u>

System Σ bestehe aus 2 Teilsystemen Σ_1, Σ_2

(Notation: $\Sigma = \Sigma_1 \times \Sigma_2$, was offenbar auf den Konfigurationsraum zutrifft) => $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$

=> # Frh.grade wird multipliziert

Seien $|e_a\rangle$, $|f_i\rangle$ ONB f. $\mathcal{H}_1$ bzw. $\mathcal{H}_2$

Betr. reinen Zustand $|\psi\rangle = \sum_{a,i} \psi_{ai} |e_a\rangle \otimes |f_i\rangle$

---

*) Die quantenmechanische Entropie ist also identisch mit der klassischen Entropie der durch die Eigenwerte v. ϱ definierten Wahrsch.verteilung

ist verschränkt (kein Produktzustand), wenn für alle Basen ≥ 2 Koeff. $\psi_{ai} \neq 0$. Die zugehörige Dichtematrix ist

$$\rho = |\psi\rangle\langle\psi| = \sum_{a,b,i,j} \psi_{ai}\psi^*_{bj} |e_a\rangle\langle e_b| \otimes |f_i\rangle\langle f_j|$$

Verzichten auf Information über Σ_2:

vgl. Randwahrsch. $p_i = \sum_j P_{ij}$ $\rightarrow$

$$(\rho_1)_{ab} = \sum_{i=j} \rho_{abij} \quad \text{oder} \quad \underline{\rho_1 = \mathrm{Tr}_{\mathcal{H}_2}\rho}$$

äquivalent: betrachten nur Observable $A \otimes 1\!\!1$

$\Rightarrow \mathrm{Tr}((A\otimes 1\!\!1)\rho) = \mathrm{Tr}_{\mathcal{H}_1}(A\rho_1)$ (weil $\mathrm{Tr} = \mathrm{Tr}_{\mathcal{H}_1} \circ \mathrm{Tr}_{\mathcal{H}_2}$, einfach nachzurechnen in Indexnot.)

$(\rho_1)_{ab} = \sum_i \psi_{ai}\psi^*_{bi} \neq \chi_a\chi^*_b$ f. irgendein $|\chi\rangle \in \mathcal{H}_1 \Rightarrow \rho_1$ ist gemischt!

Dieses Phänomen hat kein klassisches Analogon

(P_{ij} deterministisch $\Rightarrow p_i, q_j$ deterministisch)

Analog: $\rho_2 = \mathrm{Tr}_{\mathcal{H}_1}\rho$, $(\rho_2)_{ij} = \sum_a \psi_{ai}\psi^*_{aj}$

Matrixnotation: $\rho_1 = \psi\psi^\dagger$, $\rho_2 = \psi^T\psi^*$, $\psi \equiv (\psi_{ai})$

$\mathrm{Tr}\,\rho_1 = \sum_{a,i} |\psi_{ai}|^2 = \mathrm{Tr}\,\rho_2 = 1$, da $\langle\psi|\psi\rangle = 1$

Außerdem ist $\mathrm{Tr}\, \varrho_1^k = \mathrm{Tr}\, \varrho_2^k \quad \forall k \in \mathbb{N}$,

weil $\mathrm{Tr}(\psi\psi^\dagger \cdots \psi\psi^\dagger) \underset{\uparrow}{=} \mathrm{Tr}(\psi^*\psi^T \cdots \psi^*\psi^T)$

ϱ_1^k ist wie ϱ_1 selbst hermitisch
$\Rightarrow$ Spur reell

$= \mathrm{Tr}(\psi^T\psi^* \cdots \psi^T\psi^*)$

$\varrho_i \ln \varrho_i$ ist Potenzreihe in ϱ_i

$\left(\ln\varrho = \ln(\mathbb{1} + \varrho - \mathbb{1}) = \varrho - \mathbb{1} - \frac{1}{2}(\varrho - \mathbb{1})^2 + \frac{1}{3}(\varrho - \mathbb{1})^3 - \ldots\right)$

$\Rightarrow \underline{S[\varrho_1]} = -k\,\mathrm{Tr}\,\varrho_1 \ln\varrho_1 = -k\,\mathrm{Tr}\,\varrho_2\ln\varrho_2 = \underline{S[\varrho_2]}$:

„<u>Verschränkungsentropie</u>"

<u>Additivität u. Subadditivität d. Entropie</u>

Ist $\varrho = \varrho_1 \otimes \varrho_2$ ein Produktzustand, dann ist

$S[\varrho] = S[\varrho_1] + S[\varrho_2]$: Additivität d. Entropie,

daraus folgt die Extensivität d. thermodyn. Entropie; Bew. $\rightarrow$ S. 33a.

Im allgemeinen Fall ist $S[\varrho] \leq S[\varrho_1] + S[\varrho_2]$, wobei

$\varrho_1 = \mathrm{Tr}_{\mathcal{H}_2}\varrho$, $\varrho_2 = \mathrm{Tr}_{\mathcal{H}_1}\varrho$. (o. Bew.)

Die von Neumann-Entropie ist auch im Schrödingerbild <u>zeitunabhängig</u>, weil

$\varrho(t) = e^{-\frac{i}{\hbar}Ht}\varrho(0)\, e^{\frac{i}{\hbar}Ht}$ dieselben Eigenwerte wie $\varrho(0)$ hat

und $\mathrm{Tr}\,\varrho\ln\varrho$ nur von den Eigenwerten v. ϱ abhängt.

Additivität d. Entropie

a) klassisch

2 Systeme Σ_1, Σ_2 heißen unabhängig, wenn die gemeinsame Dichtefkt. $\rho(\vec{X}_1, \vec{X}_2) = \rho_1(\vec{X}_1)\rho_2(\vec{X}_2)$ erfüllt, wo $\vec{X}_i = (\vec{q}_i, \vec{p}_i)$.

$$\Rightarrow \rho \ln \rho = \rho_1 \rho_2 (\ln \rho_1 + \ln \rho_2)$$

$$\int d^{2n}X_1 \, d^{2n}X_2 \, \rho \ln \rho = \int d^{2n}X_1 \rho_1 \ln \rho_1 + \int d^{2n}X_2 \rho_2 \ln \rho_2$$

$$\Rightarrow S[\rho] = S[\rho_1] + S[\rho_2]$$

b) quantenmechanisch

$$\rho = \rho_1 \otimes \rho_2 = (\rho_1 \otimes 1\!\!1_2)(1\!\!1_1 \otimes \rho_2)$$

$$\Rightarrow \ln(\rho_1 \otimes \rho_2) = (\ln \rho_1) \otimes 1\!\!1_2 + 1\!\!1_1 \otimes \ln \rho_2$$

da $\rho_1 \otimes 1\!\!1_2 \cong \begin{pmatrix} \rho_{1,1} & & & & \\ & \ddots & & & 0 \\ & & \rho_{1,1} & & \\ & 0 & & \rho_{1,2} & \\ & & & & \rho_{1,2} \\ & & & & & \ddots \end{pmatrix}$, $1\!\!1_1 \otimes \rho_2 \cong \begin{pmatrix} \rho_{2,1} & & & & & \\ & \rho_{2,2} & & & & \\ & & \ddots & & & \\ & & & \rho_{2,n} & & \\ & & & & \rho_{2,1} & \\ & & & & & \ddots \\ & & & & & & \rho_{2,n} \\ & & & & & & & \ddots \end{pmatrix}$

$\rho_{i,k}$ ... Eigenwerte v. ρ_i

$$\Rightarrow \mathrm{Tr}\, \rho \ln \rho = \mathrm{Tr}(\rho_1 \ln \rho_1 \otimes \rho_2 + \rho_1 \otimes \rho_2 \ln \rho_2)$$

$$= \mathrm{Tr}_{\mathcal{H}_1} \rho_1 \ln \rho_1 + \mathrm{Tr}_{\mathcal{H}_2} \rho_2 \ln \rho_2$$

$$\Rightarrow S[\rho] = S[\rho_1] + S[\rho_2]$$

Beweis d. Konkavität v. $S[\varrho]$

Lemma 1: Ist e ein Einheitsvektor u. A hermitisch, dann gilt f. jede konkave Fkt. f $\quad f(\langle e|A|e\rangle) \geq \langle e|f(A)|e\rangle$

Bew.: $e = \sum \lambda_i \varphi_i$, φ_i ONB v. Eigenvektoren v. A, $\sum |\lambda_i|^2 = 1$

$$\Rightarrow \langle e|A\,e\rangle = \sum |\lambda_i^2|\, a_i\,, \qquad (A\varphi_i = a_i \varphi_i)$$

z.z. also $f(\sum |\lambda_i|^2 a_i) \geq \sum |\lambda_i|^2 f(a_i)$

Das ist aber eine allg. Eigenschaft einer konkaven Fkt. f (Jensensche Ungleichung, folgt durch Induktion aus

$$f(\alpha_1 a_1 + \alpha_2 a_2) \geq \alpha_1 f(a_1) + \alpha_2 f(a_2)\,, \quad 0 \leq \alpha_i \leq 1,\ \sum \alpha_i = 1$$

d. h. der Funktionsgraph liegt über jeder Sekante)

Lemma 2: $$\mathrm{Tr}\, f(A) = \inf_{\{e_i\}} \sum_i f(\langle e_i|Ae_i\rangle)$$

wo $\{e_i\}$ ein (nicht notwendigerweise vollständiges) ONS

Bew.: folgt aus Lemma 1, da $\sum_i f(\langle e_i|Ae_i\rangle) \geq \sum \langle e_i|f(A)e_i\rangle$ und die Gleichheit gilt, wenn die e_i Eigenvektoren v. A sind

Seien nun $|i\rangle$ die orthonormierten Eigenvektoren von $\lambda A + (1-\lambda)B$

Aus L. 2 folgt $\underline{\mathrm{Tr}\, f(\lambda A + (1-\lambda)B)} = \sum_i f(\lambda\langle i|A|i\rangle + (1-\lambda)\langle i|B|i\rangle)$

$$\geq \lambda \sum_i f(\langle i|A|i\rangle) + (1-\lambda)\sum_i f(\langle i|B|i\rangle) \geq \underline{\lambda\, \mathrm{Tr} f(A) + (1-\lambda)\,\mathrm{Tr} f(B)}$$

Setze $f(x) = -x \ln x$, $A = \varrho_1$, $B = \varrho_2$ $\Rightarrow$ $S[\varrho]$ ist konkav

Bem. I. a. gilt nicht, daß $f(\lambda A + (1-\lambda)B) \geq \lambda f(A) + (1-\lambda) f(B)$!