

Vorlesungsskript

Statistische Mechanik

Jun.-Prof. Dr. Martin Kliesch

Tex-Dokument kompiliert am 12. März 2020

Ergänzende Literatur und Grundlage des Skripts

- F. Schlögl, *Probability and Heat – Fundamentals of Thermostatistics* [1]: Referenz zur verallgemeinerten kanonischen Verteilung und den damit direkt verbundenen Konzepten
- F. Schwabl, *Statistische Physik* [2]
- W. Nolting, *Grundkurs Theoretische Physik 6 – Statistische Physik* [3] (mit Übungsaufgaben mit Lösungen)
- Kluge und Neugebauer – Einführung in die Thermodynamik
- F. Scheck, *Theoretische Physik 5 – Statistische Theorie der Wärme* [4] (mit Behandlung einiger geometrischer Aspekte)
- T. Fließbach, *Statistische Physik* [5]
- R. Egger, *Statistische Mechanik – Vorlesungsskript* (dicht am Schwabl [2]), zum Download [6]
- Referenzen zu den mathematischen Teilen des Skripts:
J. Elstrodt, *Maß- und Integrationstheorie* [7];
Rockafellar, *Convex Analysis* [8];
T. M. Cover and J. A. Thomas, *Elements of information theory* [9]

Inhaltsverzeichnis

1. VL	4
1 Einleitung	6
1.1 Motivation	6
1.2 Notation	7
1.3 Wahrscheinlichkeitstheorie	8
1.3.1 Ereignisse und Wahrscheinlichkeitsmaße	8
2. VL	10
1.3.2 Zufallsvariablen	11
3. VL	15
1.4 Informationstheorie	17
1.4.1 Shannon-Entropie	18
4. VL	19
1.4.2 Kullback-Leibler-Divergenz (Zusatzinfo)	20
1.4.3 Differentielle Entropien	20
1.5 Konvexität und Legendre-Transformation (5. VL)	22
1.5.1 Legendre-Transformation (5. VL)	22
1.5.2 Eine Konstruktion von konvexen Funktionen	23
1.6 1-Formen	24
1.6.1 Integrierende Faktoren	25
1.6.2 Homogene Funktionen	25
1.7 Spektralkalkül	25
2 Klassische statistische Mechanik	26
2.1 Die verallgemeinerte kanonische Verteilung	27
5. VL	29
6. VL	30
2.1.1 Beispiel: Kanonische Verteilung	30
2.1.2 Beispiel: Großkanonische Verteilung	32
2.1.3 Ein Dissipations-Fluktuations-Satz	33
2.2 Begründung der Gleichgewichtsthermodynamik	34
7. VL	34
2.2.1 Der Satz von Liouville	35
2.2.2 Gleichgewichtsverteilungen und thermodynamische Variablen	35
2.2.3 Zeit und Irreversibilität	37
8. VL	38
2.2.4 Extensive und intensive Größen	38
2.2.5 Der thermodynamische Limes	39
2.3 Die klassischen Ensemble	40
9. VL	40
2.3.1 Das kanonische Ensemble	41
2.3.2 Gibbs-Ensemble	42
10. VL	43
2.3.3 Großkanonische Ensemble	44
2.3.4 Das Magnetfeld-Ensemble	46
2.3.5 Das mikrokanonische Ensemble	46
11. VL	47
2.3.6 Wiederholung und Ensembleübersicht	47

3	Quantenstatistik	47
3.1	Quantenmechanikgrundlagen	48
3.1.1	Dichteoperatoren	49
	12. VL	49
3.1.2	Zeitentwicklung	51
3.1.3	Zusammengesetzte Systeme	52
	13. VL	53
3.1.4	Reduzierte Zustände/Marginale	53
3.2	Informationsmaße für Quantenzustände	54
	14. VL	57
3.3	Analogietabelle-Tabelle: Klassische Mechanik vs. Quantenmechanik	57
3.4	Der verallgemeinerte kanonisch Dichteoperator	58
3.5	Dynamische Äquilibration und Thermalisierung (Zusatzinfo)	60
3.6	Ununterscheidbare Teilchen	62
	15. VL	63
3.7	Ideale Quantengase I	64
3.7.1	Großkanonisches Ensemble	64
4	Thermodynamik	67
	16. VL	67
4.1	Euler-Gleichung und Gibbs-Duhem-Relation	68
	17. VL	69
4.1.1	Beispiel: Ideales Gas	69
4.1.2	Erster Hauptsatz der Thermodynamik	70
4.2	Thermodynamische Potentiale	70
	18. VL	73
4.2.1	Maxwell-Relationen	73
4.2.2	Suszeptibilitäten und thermodynamische Stabilität	75
	19. VL	77
4.2.3	Gleichgewichtsbedingungen	79
4.3	Thermodynamische Prozesse	83
4.3.1	Grundbegriffe	83
4.3.2	Irreversible Expansion eines Gases: Gay-Lussac-Versuch (1807)	83
	21. VL	85
4.3.3	Mischung von Gasen	85
4.3.4	Isotherme Expansion eines idealen Gases	85
4.3.5	Adiabatische Expansion eines idealen Gases	86
	22. VL	86
4.3.6	Der Carnot-Prozess	86
4.3.7	Der zweite Hauptsatz	88
4.3.8	Absolute Temperatur (Zusatzinfo)	89
	23. VL	89
4.3.9	Phänomenologische Entropie	89
4.3.10	Der 3. Hauptsatz der Thermodynamik	91
5	Modellsysteme und Phasenübergänge	92
5.1	Ideales Molekül-Gas	92
	24. VL	93
5.2	Reale Gase	93
5.2.1	Thermische Zustandsgleichung	94
5.2.2	Kalorische Zustandsgleichung	96
5.2.3	Der Joule-Thomson-Prozess	97
	25. VL	98
5.2.4	Phasenübergänge	98
	26. VL	100

5.3	Ideale Quantengase II	101
5.3.1	Zustandsgleichung	102
	27. VL	104
5.3.2	Klassischer Grenzfall	104
5.3.3	Ideale Bosegase	105
	28. VL	106
5.3.4	Ideales Fermigas	108
	29. VL	110
5.4	Das Ising-Modell	110
5.4.1	Zustandssumme	112
5.4.2	Das 1D-Ising-Modell	112
5.4.3	2D Isingmodell	114
	Bibliographie	115

1. VL

Vorbereitung

- Fragen, Wünsche, Feedback, usw. an teaching@mkliesch.eu
- Literatur: gerne eine **Mail senden, wenn etwas in der Bibliothek fehlt**

Umfrage zu Vorkenntnissen

- Wahrscheinlichkeitstheorie
 - σ -Algebren
 - Zufallsvariablen, z.B. Varianz
 - Momente und Kumulanten
 - Zentraler Grenzwertsatz
- Differentialformen
- Mechanik: Hamiltonfunktion
- Quantenmechanik
 - Wie wird ein Spin- $\frac{1}{2}$ -Teilchen mathematisch beschrieben?
 - Und n Spin- $\frac{1}{2}$ -Teilchen? Dimension?
 - Dichtematrizen
- Anwesende studieren Physik/Medizin-Physik. Noch etwas anderes?

Formalien

- Sciebo-Verzeichnis [6] (Material und Übungsblätter)
- Prüfungsvoraussetzung: mindestens 50% der Punkte auf den Übungsblättern müssen erreicht werden und es muss ein Mal in einer Übung vorgerechnet werden.
- Äquivalente Prüfungszulassungen aus vorangegangenen Veranstaltungen werden anerkannt, sofern eine Bestätigung vorliegt.
- Die 1. Übung findet in der 1. Vorlesungswoche statt, das 1. Tutorium in der zweiten.
- Es gibt eine zusätzliche Übung Donnerstags 16:30h - 18:30h, falls sich zeitnahe ausreichend viele Studierende anmelden.
- Zusammenarbeit ist erlaubt, muss aber auf den Abgaben verwerkt werden.
- Die Abgaben sollen nach Möglichkeit in Papierform erfolgen. Eine elektronische Abgabe ist trotzdem möglich [10].

- Die 1. Klausur findet am 3.2.2020 um 11:30h–14:30h in den Hörsälen 2531.HS 5 K, L & J statt.
Die 2. Klausur wird voraussichtlich am 3.4.2020 stattfinden. Es sind werden keine Hilfsmittel zugelassen sein.
- Die Klausur wird auf den Übungszetteln und kurzen Verständnisfragen zur Vorlesung basieren. In der Klausur wird es eine kurze Formelsammlung geben.
- Das Skript bietet eine Zusammenfassung der Vorlesung.

1 Einleitung

1.1 Motivation

Ziele der statistische Mechanik/Physik:

- Begründung der Thermodynamik
- Genauere thermodynamische Voraussagen (z.B. Dissipations-Fluktuations-Aussagen, Nichtgleichgewicht, ...)
- Bereitstellung allgemein anwendbarer Methoden (z.B. Kosmologie, Ökonomie)
- Untersuchungsobjekte: Vielteilchensysteme

Grundidee:

- Beschreibung makroskopischer Systeme, die sich aus mikroskopischen Teilchen zusammensetzen.
 - Der *Mikrozustand* eines Vielteilchensystem ist klassisch durch Phasenraumkoordinaten $\xi := (q_1, p_1, \dots, q_N, p_N)^\top$ gegeben. Typischerweise ist $N \sim 6 \cdot 10^{23}$.
- Aus den Anfangsbedingungen (Orte und Impulse) aller Teilchen ergibt sich sämtliche Information über das entsprechende makroskopische System.
- Kenntnisse aller Anfangsbedingungen: unrealistisch und für relevante Fragestellungen unnötig
- Idee: Erfasse Nichtwissen/Ignoranz gegenüber mikroskopischen Details mithilfe von statistischen Methoden (wie Mittelung)
 - Der *Makrozustand* wird durch wenige anwendungsrelevante makroskopische Variablen bestimmt.
- Wichtiges Beispiel: Ideales Gas im (Quasi-)Gleichgewicht. Physikalische makroskopische Größen sind: Energie E , Temperatur T , Volumen V , Druck p , Teilchenzahl N , chemisches Potential μ

Kurz: Die klassische statistische Mechanik liefert probabilistische Beschreibungen beobachtbarer physikalischer Größen von Vielteilchensystemen.

Daher beginnt die Vorlesung mit *Wahrscheinlichkeits-* und *Informationstheorie*. Die Informationstheorie führt zum Jaynesschen Prinzip der maximalen Entropie. Dieses Prinzip erlaubt das makroskopische Wissen über ein System effektiv zu nutzen und liefert einen roten Faden für weite Teile der statistischen Mechanik und Thermodynamik.

Auf diesen mathematischen Teil folgt der physikalische Teil der Vorlesung, s. Inhaltsverzeichnis.

Begriffsklärung: Die *Statistik* ist die Wissenschaft des Umgangs mit Daten, was deren Sammlung, Aufbereitung, Analyse, Interpretation und Präsentation mit einschließt. Die *mathematische Statistik* ist die Wissenschaft der Anwendung von Wahrscheinlichkeitstheorie auf die Statistik. Die *Wahrscheinlichkeitstheorie* ist die Wissenschaft der mathematischen Modellierung von Zufallsgeschehen, deren zentralen Objekte zufällige Ereignisse, Zufallsvariablen und stochastische Prozesse sind. Die *Stochastik* umfasst die *Wahrscheinlichkeitstheorie* und die *mathematische Statistik*.

Die *statistische Physik* ist das Fachgebiet, welches die Stochastik auf physikalische (oder physikalisch motivierte) Fragestellungen anwendet (*stochastische Physik* wäre ein zutreffenderer Begriff). Häufig wird *statische Mechanik* als synonym für statistische Physik verwendet. Man meint damit doch meist konkreter die Physik vieler mikroskopischer Teilchen, die mithilfe der Stochastik formalisiert wird.

1.2 Notation

- (i) $\{x_1, x_2, \dots\}$ bezeichnet die *ungeordnete Menge* und (x_1, x_2, \dots) die *geordnete Menge* von Elementen x_1, x_2, \dots .
- (ii) $\emptyset := \{ \}$ bezeichnet die *leere Menge*.
- (iii) Die *Potenzmenge* einer Menge Ω ist die Menge aller Teilmengen von Ω und wird mit $\mathcal{P}(\Omega)$ bezeichnet.
- (iv) Für $A, B \subset \Omega$ bezeichnen $A \cap B$, $A \cup B$, $A \setminus B$ jeweils den *Schnitt*, die *Vereinigung* und die mengentheoretische *Differenz* von A und B .
- (v) Für eine natürliche Zahl n setzen wir

$$[n] := \{1, 2, \dots, n\}. \quad (1.1)$$

- (vi) Die Ableitung einer Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ wird mit df bezeichnet und wir fassen df als Abbildung $df : \mathbb{R}^n \rightarrow L(\mathbb{R}^n, \mathbb{R})$ auf, wobei $L(\mathbb{R}^n, \mathbb{R})$ den Vektorraum der linearen Abbildungen von \mathbb{R}^n nach \mathbb{R} bezeichnet; häufig wird $L(\mathbb{R}^n, \mathbb{R})$ mit Zeilenvektoren identifiziert. Der Gradient $\nabla f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ist der duale Vektor zu df ; häufig als Spaltenvektor geschrieben.
- (vii) Häufig bezeichnet man die Koordinatenfunktion $\mathbb{R}^n \rightarrow \mathbb{R}$, $x \mapsto x_i$ ebenfalls mit x_i . Da die Koordinatenfunktion x_i linear ist, ist $dx_i : \mathbb{R}^n \rightarrow L(\mathbb{R}^n, \mathbb{R})$ eine konstante Funktion. Genauer: Wenn wir dx_i bei $p \in \mathbb{R}^n$ auswerten erhalten wir

$$d_p x_i = x_i. \quad (1.2)$$

Damit kann man df als

$$df = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i \quad (1.3)$$

schreiben.

- (viii) Eine Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ heißt *konvex* falls

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \quad (1.4)$$

für alle $x, y \in \mathbb{R}^n$ und $\lambda \in [0, 1]$.

Informationen, die nicht in der Vorlesung behandelt wurden sind in Grau geschrieben, bzw. haben eine graue Überschrift.

1.3 Wahrscheinlichkeitstheorie

1.3.1 Ereignisse und Wahrscheinlichkeitsmaße

1.3.1.1 Ereignisräume

- Ein Mengensystem $\Sigma \subset \mathcal{P}(\Omega)$ (Potenzmenge) über einer Menge Ω heißt σ -Algebra wenn
 - (i) $\Omega \in \Sigma$,
 - (ii) $\Omega \setminus A \in \Sigma$ für alle $A \in \Sigma$ (abgeschlossen unter Komplementen), und
 - (iii) $\bigcup_{j=1}^{\infty} A_j \in \Sigma$ für alle $A_1, A_2, \dots \in \Sigma$ (abgeschlossen unter abzählbaren Vereinigungen).
- Ein *Ereignisraum* ist ein Tupel (Ω, Σ) wobei $\Sigma \subseteq \mathcal{P}(\Omega)$ eine σ -Algebra ist.
Beispiele:
 - Würfel: $\Omega = \{1, \dots, 6\}$ und $\Sigma = \mathcal{P}(\Omega)$
 - Punktteilchen in 1D (reelle Zahlen): $\Omega = \mathbb{R}$ und $\Sigma \subsetneq \mathcal{P}(\Omega)$.
Die Menge Σ ist häufig die Borel σ -Algebra. Die Details sind nicht wichtig für die Vorlesung, doch es sei gesagt dass $\Sigma = \mathcal{P}(\Omega)$ zu keinem konsistenten Wahrscheinlichkeitsbegriff führen kann (Banach-Tarski-Paradox).
- Ereignisse* werden durch Elemente von Σ modelliert, z.B. $\{1, 3, 5\}$ wird mit dem Ereignis „der Würfel zeigt eine ungerade Augenzahl“ identifiziert.
- Logische Operationen für $A, B \in \Sigma$ sind

$$\begin{aligned}
 \text{NOT } A &= \Omega \setminus A \\
 A \text{ AND } B &= A \cap B \\
 A \text{ OR } B &= A \cup B \\
 A \text{ XOR } B &= A \Delta B
 \end{aligned} \tag{1.5}$$

wobei $A \Delta B := (A \cup B) \setminus (A \cap B)$ die symmetrische Differenz bezeichnet.

- Eine σ -Algebra ist insbesondere abgeschlossen unter diesen Operationen.
- Freiwillige Übung: Zeigen Sie folgendes:
 - $(\mathcal{P}(\Omega), \Delta, \cap)$ ist ein Ring
 - Dieser Ring ist auch eine Algebra über dem Körper $\{0, 1\}$ mit den Operationen $0 \cdot A = \emptyset$ und $1 \cdot A = A$.
 - Damit ist eine σ -Algebra $\Sigma \subset \mathcal{P}(\Omega)$ eine Unteralgebra.
 - Umgekehrt ist eine Unteralgebra $\Sigma' \subset \mathcal{P}(\Omega)$ auch eine σ -Algebra, falls (i) $\Omega \in \Sigma'$ und (ii) Σ' abgeschlossen unter abzählbaren Vereinigungen ist.

1.3.1.2 Wahrscheinlichkeitsmaße

Zwei Mengen $A, B \subset \Omega$ heißen *disjunkt*, wenn $A \cap B = \emptyset$.

Definition 1.1 ((Wahrscheinlichkeits)maße [A. Kolmogorov 1930er]):

Ein *Wahrscheinlichkeitsmaß* auf einem Ereignisraum (Ω, Σ) ist eine Abbildung

$$\mathbb{P} : \Sigma \rightarrow \mathbb{R}, \tag{1.6}$$

die folgendes erfüllt:

- (i) $0 \leq \mathbb{P}[A] \leq 1$ für alle $A \in \Sigma$,
- (ii) $\mathbb{P}[\Omega] = 1$, und

(iii) \mathbb{P} ist σ -additiv, d.h.

$$\mathbb{P}\left[\bigcup_{j=1}^{\infty} A_j\right] = \sum_{j=1}^{\infty} \mathbb{P}[A_j] \quad (1.7)$$

für alle paarweise disjunkte Ereignisse $A_1, A_2, \dots \in \Sigma$.

Ein Maß $\mu : \Sigma \rightarrow \mathbb{R}$ ist ähnlich definiert, wie ein Wahrscheinlichkeitsmaß, nur dass $\mu \geq 0$ anstelle von (i) und (ii) gelten muss.

Das Tripel $(\Omega, \Sigma, \mathbb{P})$ wird *Wahrscheinlichkeitsraum* und (Ω, Σ, μ) *Maßraum* genannt. Für $\Omega = \mathbb{R}^n$ nennt man die durch

$$F(t) := \mathbb{P}[(-\infty, t_1] \times \dots \times (-\infty, t_n)] \quad (1.8)$$

definierte Funktion (*kumulative Verteilungsfunktion*). Falls \mathbb{P} eine Wahrscheinlichkeitsdichtefunktion hat (s.u.), so ist diese als Ableitung von ρ gegeben.

Bemerkung: Praktisch wird der σ -Algebra meist wenig Aufmerksamkeit geschenkt. Es genügt in den für uns relevanten Fälle sämtliche Eigenschaften von \mathbb{P} auf gutartigen Mengen wie z.B. offenen oder abgeschlossenen Mengen oder beliebigen Intervallen zu überprüfen.

Es folgt aus der Definition eines Maßes z.B. dass $\mathbb{P}[A] \leq \mathbb{P}[B]$ durch $A \subset B$ impliziert wird (Übungsaufgabe).

1.3.1.3 Diskrete und kontinuierliche Wahrscheinlichkeitsmaße

Es gibt zwei für uns wichtige Klassen von Wahrscheinlichkeitsmaßen:

- (i) Ein *Kontinuierliches* oder (oder *absolutstetiges*) Wahrscheinlichkeitsmaß auf \mathbb{R}^n ist durch eine *Wahrscheinlichkeitsdichtefunktion* (WDF) $\rho : \mathbb{R}^n \rightarrow [0, 1]$ gegeben, d.h., ρ ist integrierbar ist und

$$\mathbb{P}[A] = \int_A \rho(x) \, dx \quad (1.9)$$

für alle $A \in \Sigma$.

- (ii) Wir nennen ein Wahrscheinlichkeitsmaß auf \mathbb{R}^n *diskret* (oder *rein atomar*) falls eine abzählbare Menge $\Gamma = \{x_1, x_2, \dots\} \subset \mathbb{R}^n$ existiert, so dass

$$\mathbb{P}[A] = \sum_{x \in A \cap \Gamma} \mathbb{P}[\{x\}] \quad (1.10)$$

für alle Ereignisse $A \in \Sigma$.

Dann ist $\mathbb{P}[\mathbb{R}^n \setminus \Gamma] = 0$.

Die Funktion $p : \mathbb{R}^n \mapsto [0, 1]$, $x \mapsto p_x := \mathbb{P}[\{x\}]$ nennt man auch Wahrscheinlichkeitsfunktion oder Zähldichte.

Das Maß

$$\delta_x(A) = \begin{cases} 1 & \text{falls } x \in A \\ 0 & \text{sonst} \end{cases} \quad (1.11)$$

nennt man auch *Diracmaß* (es mithilfe der Dirac-Delta-„funktion“ geschrieben werden). Damit lässt sich das diskrete Maß (1.10) auch als Konvexkombination

$$\mathbb{P} = \sum_{x \in \mathbb{R}^n} p_x \delta_x \quad (1.12)$$

schreiben.

Falls $p_x \neq 0$ für nur endlich viele x nennt man es auch *Wahrscheinlichkeitsvektor*. In diesem Fall ist

$$p_x = \lim_{N \rightarrow \infty} \frac{N_x}{N}, \quad (1.13)$$

wobei der das Ereignis x bei N gezogenen Ereignissen N_x Mal auftritt.

2. VL

1.3.1.4 Bedingte Wahrscheinlichkeiten und unabhängige Ereignisse

Für $\mathbb{P}[B] > 0$ definiert

$$\mathbb{P}[A|B] := \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]} \quad (1.14)$$

die durch B bedingte Wahrscheinlichkeit von A ; hier ist ein XKCD¹ dazu. Zwei Ereignisse A und B sind (stochastisch) unabhängig wenn $\mathbb{P}[A \cap B] = \mathbb{P}[A] \mathbb{P}[B]$. Also, für $\mathbb{P}[B] > 0$ wenn $\mathbb{P}[A|B] = \mathbb{P}[A]$.

Aus der Definition (1.14) folgt die Ketten- oder Produktregel,

$$\mathbb{P}[A \cap B] = \mathbb{P}[A|B] \mathbb{P}[B]. \quad (1.15)$$

Außerdem folgt direkt Bayes Theorem

$$\mathbb{P}[B|A] = \frac{\mathbb{P}[A|B] \mathbb{P}[B]}{\mathbb{P}[A]} \quad (1.16)$$

für $\mathbb{P}[A] \neq 0 \neq \mathbb{P}[B]$.

1.3.1.5 Zusammensetzen von Wahrscheinlichkeitsräumen

Für zwei Wahrscheinlichkeitsräume $(\Omega_1, \Sigma_1, \mathbb{P}^{(1)})$ und $(\Omega_2, \Sigma_2, \mathbb{P}^{(2)})$ können wir den zusammengesetzten Wahrscheinlichkeitsraum $(\Omega, \Sigma, \mathbb{P})$ bilden, wobei $\Omega := \Omega_1 \times \Omega_2$ (kartesisches Produkt) und das Produktmaß über

$$\mathbb{P}[A_1 \times A_2] := \mathbb{P}^{(1)}[A_1] \mathbb{P}^{(2)}[A_2], \quad (1.17)$$

wobei $A_1 \in \Sigma_1$ und $A_2 \in \Sigma_2$, definiert wird. Falls $\mathbb{P}^{(1)}$ und $\mathbb{P}^{(2)}$ Zähldichten $p^{(1)}$ und $p^{(2)}$ haben, ist die Zähldichte p von dem Produktmaß \mathbb{P} über

$$p_{(i,j)} = p_i^{(1)} p_j^{(2)} \quad (1.18)$$

gegeben. Falls $\mathbb{P}^{(1)}$ und $\mathbb{P}^{(2)}$ WDFen $\rho^{(1)} : \Omega_1 \rightarrow [0, 1]$ und $\rho^{(2)} : \Omega_2 \rightarrow [0, 1]$ haben, ist die WDF ρ von dem Produktmaß \mathbb{P} über

$$\rho(x, y) = \rho^{(1)}(x) \rho^{(2)}(y) \quad (1.19)$$

gegeben.

1.3.1.6 Marginalverteilung

Gegeben einen Wahrscheinlichkeitsraum $(\Omega_1 \times \Omega_2, \Sigma, \mathbb{P})$ können wir die *Marginalverteilung* auf Ω_1 bilden. Sie ist durch ein Wahrscheinlichkeitsmaß $\mathbb{P}^{(1)}$ auf Ω_1 gegeben mit

$$\mathbb{P}^{(1)}[A] := \mathbb{P}[A \times \Omega_2] \quad (1.20)$$

für Ereignisse $A \subset \Omega_1$. Falls \mathbb{P} eine Zähldichte p hat ist sie durch Wahrscheinlichkeiten der Form $p_{(i,j)}$ mit $i \in \Omega_1$ und $j \in \Omega_2$ gegeben und die Marginalverteilung hat die

¹<https://xkcd.com/795/>

Zähldichte $p^{(1)}$ mit

$$p_i^{(1)} := \sum_{j \in \Omega_2} p_{(i,j)} . \quad (1.21)$$

Falls \mathbb{P} eine WDF ρ hat, hat die Marginalverteilung die WDF $\rho : \Omega_1 \rightarrow [0, 1]$ mit

$$\rho^{(1)}(x) := \int_{\Omega_2} \rho(x, y) dy . \quad (1.22)$$

Ähnlich ist die Marginalverteilung auf Ω_2 definiert.

1.3.1.7 Zusammenfassung

- Ereignisse werden durch Teilmengen modelliert, Mengenoperationen entsprechen logischen Operationen.
- Wahrscheinlichkeitsmaße weisen Ereignissen Wahrscheinlichkeiten zu.
- Diskrete und kontinuierliche Wahrscheinlichkeitsmaße sind wichtige Beispiele.
- Kettenregel und stochastisch unabhängige Ereignisse

1.3.2 Zufallsvariablen

Motivation: Wir wollen z.B. die Augenzahl eines Würfels als eine Variable auffassen, die mit einer Wahrscheinlichkeitsverteilung assoziiert ist.

Sei $(\Omega, \Sigma, \mathbb{P})$ ein Maßraum. Eine reelle Zufallsvariable ist eine Abbildung

$$X : \Omega \rightarrow \mathbb{R} \quad (1.23)$$

die bestimmte Konsistenzbedingungen erfüllt (sie muss messbar sein). Diese Konsistenzbedingungen sind automatisch für alle in der Vorlesung auftretenden und anschaulich vorstellbaren Funktionen erfüllt. Häufig werden alle Details zur σ -Algebra verschwiegen, was wir hier auch tun. Per Definition wird \mathbb{P} als die Wahrscheinlichkeitsverteilung von X aufgefasst.

Kurz: Eine Zufallsvariable ist eine Funktion mit einer assoziierten Wahrscheinlichkeitsverteilung.

Häufig wird das zu X zugehörige Maß auch als \mathbb{P}_X bezeichnet. Ähnlich wird die zugehörige Verteilungsfunktion mit F_X und die WDF, falls sie existiert, mit ρ_X bezeichnet.

Die Definition von Zufallsvariablen kann direkt auf vektorwertige Zufallsvariablen verallgemeinert werden.

Definition 1.2 (Zufallsvektor):

Sei $(\Omega, \Sigma, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Wir fassen den \mathbb{R}^N ebenfalls als ein Ereignisraum auf (ausgestattet mit der Borelschen σ -Algebra). Dann ist ein *Zufallsvektor* (im \mathbb{R}^N) eine Abbildung

$$X : \Omega \rightarrow \mathbb{R}^N , \quad (1.24)$$

die sehr milde Konsistenzbedingungen erfüllt (sie muss messbar sein). Im Fall $N = 1$ nennt man X *Zufallsvariable*.

Für die für uns wichtigen diskreten und kontinuierlichen Maße (Abschnitt 1.3.1.3) ist es äußerst schwierig, Abbildungen zu konstruieren, welche die erwähnten Konsistenzbedingungen verletzen, also nicht messbar sind. Daher wird dieser Aspekt nicht weiter betrachtet.

Häufig betrachten wir Ereignisse der Form

$$\{X \in A\} := X^{-1}(A) := \{\omega \in \Omega : X(\omega) \in A\}, \quad (1.25)$$

d.h. das Ereignis dass X einen Wert annimmt, der zu einem Ereignis $A \subset \Omega$ gehört. Ähnlich wird

$$\{X \leq x\} := \{\omega \in \Omega : X_i(\omega) \leq x_i \ \forall i = 1, \dots, N\}, \quad (1.26)$$

$\{X = x\}$, $\{X \notin A\}$ usw. definiert. Man schreibt meistens $\mathbb{P}[X \in A] := \mathbb{P}[\{X \in A\}]$. Jeder Zufallsvektor (1.24) hat auch eine Verteilungsfunktion, die durch

$$\begin{aligned} F_X : \mathbb{R}^N &\rightarrow \mathbb{R} \\ x &\mapsto \mathbb{P}[X \leq x] \end{aligned} \quad (1.27)$$

gegeben ist. Sie stimmt mit der Verteilungsfunktion von dem zu X zugehörigen Maß überein (schnelle Übungsaufgabe).

Wir sagen, dass ein Zufallsvektor X im \mathbb{R}^N eine WDF hat, falls das zugehörigen Wahrscheinlichkeitsmaß eine WDF ρ_X hat. In diesem Fall ist

$$F_X(x) = \int_{-\infty}^{x_1} dy_1 \cdots \int_{-\infty}^{x_N} dy_N \rho_X(y). \quad (1.28)$$

Es ist hilfreich zu bemerken, dass die Verteilungsfunktion (1.27) nicht explizit vom Ereignisraum Ω abhängt.

Beispiele 1.3:

- Die Augenzahl X eines Würfels ist eine Zufallsvariable. Wenn der Würfel fair ist, spricht man auch von einer *gleichverteilten Zufallsvariable*.
- Jede Funktion der Augenzahl eines Würfels ist eine Zufallsvariable. Zum Beispiel kann man eine Zufallsvariable Y definieren, die 0 annimmt, falls der Würfel eine gerade Augenzahl hat und 1 sonst.
- Die Wahrscheinlichkeit, dass eine ungerade Zahl gewürfelt wurde ist $\mathbb{P}[X \in \{1, 3, 5\}]$, wobei X die Augenzahl ist.
- In der statistischen Mechanik spielen zufällig Phasenraumkoordinaten der Form $\xi := (q_1, p_1, \dots, q_N, p_N)^\top$ eine zentrale Rolle. Diese sind Zufallsvektoren im \mathbb{R}^{6N} .

1.3.2.1 Unabhängige Zufallsvariablen

Der Begriff der statistischen Unabhängigkeit kann von Ereignissen auf Zufallsvariablen und -vektoren erweitert werden.

Definition 1.4 (Unabhängige Zufallsvektoren):

Man nennt zwei Zufallsvektoren $X : \Omega \rightarrow \mathbb{R}^N$ und $Y : \Omega \rightarrow \mathbb{R}^{N'}$ *unabhängig* wenn

$$\mathbb{P}[X \in A \text{ AND } Y \in B] = \mathbb{P}[X \in A] \mathbb{P}[Y \in B] \quad (1.29)$$

für alle Ereignisse $A \subset \mathbb{R}^N$ und $B \subset \mathbb{R}^{N'}$.

Für Zufallsvariablen genügt es, die Unabhängigkeit auf Intervallen der Form $A = (\infty, a]$ und $B = (\infty, b]$ zu überprüfen, d.h. für abhängige Zufallsvariablen gibt es a und b , so dass die entsprechenden Intervalle die obige Gleichung verletzen. Eine ähnliche Aussage mit mehrdimensionalen Intervallen kann für Zufallsvektoren gemacht

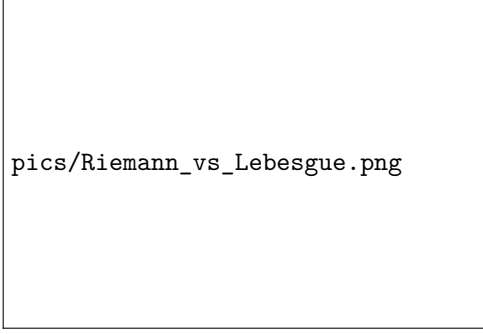


Abbildung 1.1: Für eine gegebene Funktion können approximierende Treppenfunktionen auf zwei Weisen konstruiert werden. (i) Eine equidistante Unterteilung der x-Achse (oben) für zur Definition des Riemann-Integrals. (ii) Eine equidistante Unterteilung der y-Achse (unten) für zur Definition des Lebesgue-Integrals. Zur Integration wird jeweils ein Maß auf dem Definitionsbe- reich benötigt. Für nicht-negative Funktionen gilt: Wenn sowohl Riemann-Integral als auch das Lebesgue-Integral existieren sind sie auch gleich.

werden. Zufallsvektoren X und Y sind also unabhängig, wenn ihre Verteilungsfunk- tionen faktorisieren, d.h. dass

$$F_{(X,Y)}(x,y) = F_X(x)F_Y(y) \quad (1.30)$$

für alle x, y . Falls (X, Y) eine WDF ρ hat und X und Y unabhängig sind, so faktori- siert diese ebenfalls,

$$\rho(x, y) = \rho_X(x)\rho_Y(y), \quad (1.31)$$

wobei ρ_X und ρ_Y dann die jeweiligen WDFen von X und Y sind. Eine analoge Aussage gilt, wenn (X, Y) eine Zähldichte hat.

1.3.2.2 Erwartungswerte

Auf einem Wahrscheinlichkeitsraum (Ω, Σ, μ) ist immer ein Integral

$$f \mapsto \int f(\omega) d\mu(\omega) \quad (1.32)$$

für hinreichend „gutartige,, (d.h. messbare) Funktionen f definiert. Eine konstante Funktion $C : \Omega \rightarrow \mathbb{R}$ mit $C(\omega) = c$ kann ganz einfach über eine Menge $A \in \Sigma$ integriert werden:

$$\int_A C(\omega) d\mu(\omega) := c\mu(A). \quad (1.33)$$

Eine zentrale Eigenschaft des Integrals (1.32) ist, dass es linear ist. Damit ist klar, wie Treppenfunktionen integriert werden. Wenn f eine ausreichend gutartige Funk- tion ist, kann sie durch Treppenfunktionen approximiert werden, s. Abbildung 1.1. Durch immer genauer werdende Approximationen kann dann ihr Integral berechnet werden. Das führt zum Integral für die kontinuierlichen und diskreten Wahrschein- lichkeitsmaße aus Abschnitt 1.3.1.3.

- (i) Wenn ein Zufallsvektor X im \mathbb{R}^N kontinuierlich verteilt ist, d.h. ein kontinuier- liches Wahrscheinlichkeitsmaß hat, dann hat er auch eine WDF ρ .
- (ii) Wenn ein Zufallsvektor Y im \mathbb{R}^N diskret verteilt ist hat er eine Zähldichte p .

Definition 1.5 (Erwartungswert):

Seien $X, Y : \mathbb{R}^n \rightarrow \mathbb{R}^N$ Zufallsvektoren mit jeweiligen Warscheinlichkeitsmaßen μ und ν . X habe eine WDF ρ und $Y : \mathbb{R}^n \rightarrow \mathbb{R}^N$ eine Zähldichte p . Dann sind

die *Erwartungswerte* dieser Zufallsvektoren jeweils als

$$\begin{aligned}\mathbb{E}[X] &:= \int X d\mu \equiv \int_{\mathbb{R}^n} X(\omega) d\mu(\omega) = \int_{\mathbb{R}^n} X(\omega) \rho(\omega) d\omega \\ \mathbb{E}[Y] &:= \int Y d\nu \equiv \int_{\mathbb{R}^n} Y(\omega) d\nu(\omega) = \sum_{\omega \in \mathbb{R}^n} Y(\omega) p_\omega\end{aligned}\quad (1.34)$$

gegeben, falls sie existieren. Anderenfalls nennt man den entsprechenden Zufallsvektor *nicht integrierbar*.

Häufig schreibt man auch $\langle X \rangle \equiv \mathbb{E}[X]$, $\mathbb{E}_{X \sim \mu}[X]$, $\mathbb{E}_{X \sim \rho}[X]$, $\mathbb{E}_{Y \sim p}[X]$ für die entsprechenden Erwartungswerte.

Sei nun $f : \mathbb{R}^N \rightarrow \mathbb{R}^M$ eine (messbare) Funktion und X ein Zufallsvektor im \mathbb{R}^N . Dann ist $f(X)$ ein Zufallsvektor im \mathbb{R}^M und kann entsprechend auch wieder einen Erwartungswert $\mathbb{E}[f(X)]$ haben.

Freiwillige Übungsaufgabe: Betrachten Sie die beiden Fälle (i) dass X eine WDF und (ii) eine Zähldichte hat. Berechnen Sie die WDF/Zähldichte von $f(X)$ für den Spezialfall, dass f injektiv ist.

Hier sind wichtige Beispiele für $\mathbb{E}[f(X)]$ (wir nehmen an, dass alle auftretenden Erwartungswerte existieren):

- Die *Varianz* einer Zufallsvariablen X ist $\text{Var}[X] := \mathbb{E}[(X - \mathbb{E}[X])^2]$. Daraus ergibt sich die *Standardabweichung* $\sqrt{\text{Var}[X]}$.
- Die *Kovarianz* von zwei Zufallsvariablen X und Y mit gleichem Ereignisraum Ω ist

$$\begin{aligned}\text{Cov}(X, Y) &:= \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] \\ &= \mathbb{E}[XY] - \mathbb{E}[X] \mathbb{E}[Y].\end{aligned}\quad (1.35)$$

- Die *charakteristische Funktion* eines Zufallsvektors X ist über

$$\varphi_X(q) := \mathbb{E}[\exp(i \langle q, X \rangle)] \quad (1.36)$$

definiert, wobei $\langle \cdot, \cdot \rangle$ das kanonische Skalarprodukt auf \mathbb{R}^N bezeichnet. Häufig schreibt man auch nur φ anstelle von φ_X wenn klar ist, welcher Zufallsvektor gemeint ist.

Die Varianz $\text{Var}[X]$ misst, wie stark X streut, d.h. wie groß die erwartete quadratische Abweichung vom Mittelwert $\mathbb{E}[X]$ ist.

1.3.2.3 Korrelationen

Zwei Zufallsvariablen X, Y heißen *unkorreliert*, falls

$$\mathbb{E}[XY] = \mathbb{E}[X] \mathbb{E}[Y] \quad (1.37)$$

ist. Zum Beispiel gilt:

$$X, Y \text{ unabhängig} \Rightarrow \mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)] \mathbb{E}[g(Y)] \quad (1.38)$$

für alle (messbaren) Funktionen f und g (schnelle Übungsaufgabe). Doch es gibt auch Zufallsvariablen, die nicht unabhängig aber unkorreliert sind (Übungsblatt). Die Kovarianz $\text{Cov}(X, Y)$ ist ein Maß für die Korrelationen von X und Y .

Anschaung: Bei medizinischen Patienten ist ein hoher Alkoholkonsum korreliert mit Lebererkrankungen. Man beachte, dass weder ein hoher Alkoholkonsum das Auftreten einer Lebererkrankung impliziert noch umgekehrt, [hier ist ein XKCD²](https://xkcd.com/552/) zu dieser Tatsache. Zufallsvariablen können auch antikorreliert sein (negative Kovarianz), z.B. das

²<https://xkcd.com/552/>

regelmäßige Betreiben von Ausdauersport ist antikorreliert mit dem Auftreten von Herz-Kreislauf-Erkrankungen.

3. VL

1.3.2.4 Momente, Kumulanten

In diesem Abschnitt betrachten wir Zufallsvariablen. Das k -te Moment einer Zufallsvariablen X ist $\mathbb{E}[X^k]$. Beispielsweise ist die Varianz $\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$ (schnelle Übungsaufgabe), also ein Polynom von Momenten.

Aus der charakteristischen Funktion φ von X lassen sich alle Momente m_k berechnen, falls sie existieren:

$$m_k := \mathbb{E}[X^k] = (-i)^k \varphi^{(k)}(0) \quad (k\text{-te Ableitung}) \quad (1.39)$$

(schnelle Übungsaufgabe), d.h. die charakteristische Funktion dient als *Momentenerzeugende*. Wir nehmen nun an, dass alle Momente existieren ist. Dann ist

$$\varphi(q) = \sum_{k=0}^{\infty} \frac{(iq)^k}{k!} m_k. \quad (1.40)$$

Es gilt $\varphi(0) = 1$. Damit kann $\ln(\varphi(q))$ als Potenzreihe um $q = 0$ entwickelt werden, was wir als

$$\ln \varphi(q) = \sum_{k=1}^{\infty} \frac{(iq)^k}{k!} \kappa_k \quad (1.41)$$

schreiben. Diese Entwicklung heißt *Kumulantenentwicklung* und die Entwicklungskoeffizienten κ_k heißen *Kumulanten* von X . $\ln \varphi$ wird *Kumulantenerzeugende* von X genannt und es gilt

$$\kappa_k = (-i)^k \left. \frac{d^k}{dq^k} \right|_{q=0} \ln \varphi(q). \quad (1.42)$$

Durch direktes Differenzieren oder indem man die Reihenentwicklung $\ln(1+x) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{x^k}{k}$ auf $\ln(\varphi)$ anwendet und (1.40) einsetzt findet man die ersten Kumulanten (freiwillige Rechenaufgabe)

$$\kappa_1 = m_1 \quad (1.43)$$

$$\kappa_2 = m_2 - m_1^2 = \mu_2 \quad (1.44)$$

$$\kappa_3 = m_3 - 3m_2m_1 + 2m_1^3 = \mu_3 \quad (1.45)$$

$$\kappa_4 = \mu_4 - 3\mu_2^2, \quad (1.46)$$

wobei $\mu_k := \mathbb{E}[(X - \mathbb{E}[X])^k]$ das k -te *zentrale Moment* von X bezeichnet (per Konvention setzte man $\mu := \mu_1 := m_1$).

1.3.2.5 Normalverteilung und zentraler Grenzwertsatz

Die Normalverteilung nimmt eine gewisse Sonderstellung unter den Wahrscheinlichkeitsverteilungen ein, was durch den zentralen Grenzwertsatz gerechtfertigt werden kann. Man schreibt $X \sim \mathcal{N}(\mu, \sigma^2)$ falls X die WDF

$$\rho(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (1.47)$$

hat und nennt X dann *normalverteilt mit Erwartungswert μ und Varianz σ^2* . Wir bezeichnen die Korrespondierende Verteilungsfunktion mit

$$\Phi_{\mu,\sigma^2}(z) = \int_{-\infty}^z \rho(x) dx. \quad (1.48)$$

Freiwillige Übungsaufgabe: Überprüfen der Normiertheit und dass die charakteristische Funktion durch

$$\varphi_{\mu,\sigma^2}(q) = \exp\left(i\mu q - \frac{1}{2}\sigma^2 q^2\right) \quad (1.49)$$

gegeben ist.

Damit ist die Kumulantenerzeugende durch

$$\ln \varphi_{\mu,\sigma^2}(q) = i\mu q - \frac{1}{2}\sigma^2 q^2 \quad (1.50)$$

gegeben und die Kumulanten (1.42) ergeben sich als

$$\kappa_1 = \mu \quad (1.51)$$

$$\kappa_2 = \sigma^2 \quad (1.52)$$

$$\kappa_k = 0 \quad \forall k \geq 3. \quad (1.53)$$

Satz 1.6 (Zentraler Grenzwertsatz):

Seien X_1, X_2, \dots, X_N unabhängige Zufallsvariablen mit existierenden ersten und zweiten Momenten und setze

$$Z_N = \frac{1}{\sqrt{N}} \sum_{i=1}^N X_i. \quad (1.54)$$

Dann gilt für jedes $z \in \mathbb{R}$

$$\lim_{N \rightarrow \infty} \mathbb{P}[Z_N \leq z] = \Phi_{\mu,\sigma^2}(z) \quad (1.55)$$

wobei

$$\begin{aligned} \mu &= \lim_{N \rightarrow \infty} \frac{1}{\sqrt{N}} \sum_{i=1}^N \mathbb{E}[X_i], \\ \sigma^2 &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \text{Var}[X_i]. \end{aligned} \quad (1.56)$$

Bemerkungen:

- Insbesondere konvergiert das *empirische Mittel*

$$\bar{X}^N := \frac{1}{N} \sum_{i=1}^N X_i \quad (1.57)$$

gegen eine Gaussverteilung die beliebig scharf um den mittleren Mittelwert der X_i konzentriert ist.

- Die Zufallsvariablen (X_i) können diskret verteilt sein. Der Grenzwert (1.55) ergibt eine kontinuierliche Verteilung.

Für den folgenden Beweis ist folgende Beobachtung zur charakteristischen Funktion

(1.36) hilfreich. Für unabhängige Zufallsvektoren X und Y gilt

$$\begin{aligned}\varphi_{X+Y}(q) &= \mathbb{E}[\exp(i\langle q, X \rangle) \exp(i\langle q, Y \rangle)] \\ &= \mathbb{E}[\exp(i\langle q, X \rangle)] \mathbb{E}[\exp(i\langle q, Y \rangle)] \\ &= \varphi_X(q) \varphi_Y(q).\end{aligned}\tag{1.58}$$

Aus der Definition (1.36) folgen auch direkt dass

$$\varphi_{\lambda X}(q) = \varphi_X(\lambda q)\tag{1.59}$$

für alle $\lambda \in \mathbb{R}$ und q .

Sei X nun eine Zufallsvariable. Dann gilt weiterhin, dass

$$\varphi_{\lambda+X}(q) = e^{i\lambda q} \varphi_X(q)\tag{1.60}$$

für alle $\lambda \in \mathbb{R}$ und q . Falls X eine WDF ρ hat finden wir ihre charakteristische Funktion φ

$$\rho(x) = \frac{1}{2\pi} \int dq e^{-iqx} \varphi(q)\tag{1.61}$$

unter der Benutzung der Fouriertransformation und ihrer Inversen.

Beweis von Satz 1.6. Für unseren Beweis nehmen wir an, dass alle Kumulanten $\kappa_k^{(X_i)}$ existieren und uniform in i beschränkt sind.

Wir verwenden die Notation $\varphi[X](q) := \varphi_X(q)$ und $\varphi_i := \varphi_{X_i}$. Dann ist

$$\begin{aligned}\ln \varphi_Z(q) &= \ln \varphi \left[\sum_{i=1}^N X_i \right] (q/\sqrt{N}) \\ &= \sum_{i=1}^N \ln \varphi_i(q/\sqrt{N})\end{aligned}\tag{1.62}$$

für jedes q . Nun fixieren wir q und benutzen die Kummulantenentwicklung (1.41),

$$\begin{aligned}\ln \varphi_Z(q) &= \sum_{i=1}^N \sum_{k=1}^{\infty} \frac{(iq/\sqrt{N})^k}{k!} \kappa_k^{(X_i)} \\ &= \sum_{i=1}^N \left(iq \mathbb{E}[X_i]/\sqrt{N} - \frac{1}{2} q^2 \text{Var}[X_i]/N \right) + O(1/\sqrt{N}) \\ &= i\mu q - \frac{1}{2} q^2 \sigma^2 + O(1/\sqrt{N}) \\ &= \ln \varphi_{\mu, \sigma^2}(q) + O(1/\sqrt{N}).\end{aligned}\tag{1.63}$$

Daraus folgt

$$\lim_{N \rightarrow \infty} \varphi_Z(q) = \varphi_{\mu, \sigma^2}(q)\tag{1.64}$$

für alle $q \in \mathbb{R}$. Diese punktweise Konvergenz von φ_Z impliziert die punktweise Konvergenz der Verteilungsfunktion von Z (Stetigkeitssatz von Lévy). \square

1.4 Informationstheorie

Die Informationstheorie wurde von Claude Shannon (1948) begründet und beschäftigt sich mit Quantifizierung, Speicherung und Übertragung von Informationen. Das umfasst insbesondere, Datenkodierung und -kompression.

Die Quantifizierung von Information lässt sich besonders gut mithilfe von Kompression veranschaulichen. Wir nehmen an, dass wir unabhängige Stichproben einer Zu-

fallsvariable X bekommen, welche m unterschiedliche Werte $x \in \{a_1, a_2, \dots, a_m\}$ annehmen kann. Wir bezeichnen die Wahrscheinlichkeiten dieser Werte mit

$$p_i := \mathbb{P}[X = a_i], \quad i = 1, 2, \dots, m \quad (1.65)$$

und nehmen an, dass diese Werte (und auch die a_i) explizit bekannt sind.

Wir betrachten nun N unabhängige Stichproben von X , d.h. wir betrachten N unabhängige Kopien $X_1, \dots, X_N \sim p$ von X . Man spricht hier auch von N *unabhängig und identisch verteilte Zufallsvariablen* (kurz: iid. für independent and identically distributed).

- Wie viele Bits (binäre Variablen) werden benötigt, um eine Stichprobe von (X_1, \dots, X_N) zu speichern?

Eine Möglichkeit für eine Speicherung wäre a_i jeweils mit der Binärdarstellung von i zu kodieren. Dieses Kodierungsschema benötigt $\lceil \log_2(m) \rceil$ N viele Bits, wobei $\lceil x \rceil := \min\{k \in \mathbb{Z} : x \leq k\}$ das Aufrunden bezeichnet.

Übung: Finden Sie Beispiele, wo weniger Bits zur Kodierung benötigt werden.

1.4.1 Shannon-Entropie

Die *Shannon-Entropie* oder auch nur *Entropie* stellt sich als ein sinnvolles Informationsmaß heraus. Sie kann als ein Maß für die „Zufälligkeit“ einer Zufallsvariable aufgefasst werden.

Definition 1.7 (Shannon-Entropie):

Sei X ein Zufallsvektor mit Zähldichte p . Dann ist die Entropie (zur Basis $b > 0$) von X als

$$H(X) := H(p) := - \sum_i p_i \log_b(p_i) \quad (1.66)$$

definiert.

Die Basis b legt eine Informationseinheit fest, die meistens aus dem Kontext heraus klar ist. In der Informatik, wo Informationen von Bits getragen werden, wird typischerweise $b = 2$ gewählt. In der statistischen Physik wählen wir die Einheiten so dass die Entropie als

$$S(X) := -k_B \sum_i p_i \ln(p_i) \quad (1.67)$$

gegeben ist, wobei k_B die Boltzmannkonstante bezeichnet.

Eigenschaften:

- Für diskret verteilte Zufallsvariable X gilt $H(X) \geq 0$. Falls X eine „konstante“ Zufallsvariable ist, gilt $H(X) = 0$, d.h., falls $\mathbb{P}[X = x] = 1$ für irgendein x und $\mathbb{P}[X = y] = 0$ für alle $y \neq x$.
- Sei X eine gleichverteilte und Y eine beliebige Zufallsvariable auf $\{a_1, a_2, \dots, a_m\}$. Dann ist $H(Y) \leq H(X) = \log_b(m)$.
- Die Entropie ist konkav:

$$H(\lambda p + (1 - \lambda)q) \geq \lambda H(p) + (1 - \lambda)H(q) \quad (1.68)$$

für alle $\lambda \in [0, 1]$ und Wahrscheinlichkeitsvektoren $p, q \in [0, 1]^m$. Diese Eigenschaft folgt aus der Konkavität des Logarithmus (kurze Übungsaufgabe).

(iv) Die Entropie additiv für unabhängige Zufallsvektoren X und Y :

$$H((X, Y)) = H(X) + H(Y) \quad (1.69)$$

(einfache Übungsaufgabe). Allgemein kann man mithilfe der Nichtnegativität der Kullback-Leibler-Divergenz (1.76) zeigen, dass für eine Zähldichte p mit Marginalverteilungen $p^{(1)}$ und $p^{(2)}$ immer

$$H(p) \leq H(p^{(1)}) + H(p^{(2)}) \quad (1.70)$$

gilt. Wir werden diese Eigenschaft später beweisen (Lemma 1), wenn wir Verallgemeinerungen der Shannon-Entropie und Kullback-Leibler-Divergenz auf Quantenzustände kennengelernt haben.

Satz 1.8 (Shannons Kodierungs-Satz (informell) [1, Kap. 5]):

N iid. diskrete Zufallsvariablen, die jeweils die Entropie $H(X)$ haben, können in mindestens $N H(X)$ viele Bits komprimiert werden, wobei die Wahrscheinlichkeit von Datenverlust im Limes $N \rightarrow \infty$ verschwindet.

Falls die Zufallsvariablen in weniger als $N H(X)$ Bits komprimiert werden, konvergiert die Wahrscheinlichkeit Daten zu verlieren gegen 1.

4. VL

Wir können die Entropie als Erwartungswert schreiben. Dafür machen wir aus der dem Wahrscheinlichkeitsvektor (p_i) selbst eine Zufallsvariable, indem wir die Zufallsvariable X mit Bild $\{i\}$ für i einsetzen, was $p_{i=X} = p_X$ ergibt (also eine Zufallsvariable mit dem gleichen Ereignisraum wie X und endlich vielen Werten in $[0, 1]$). Wir definieren den *Informationsgehalt* von i bzgl. p als

$$\iota(i) := -\log_b(p_i) \quad (1.71)$$

und erhalten damit

$$H(X) = \mathbb{E}[\iota(X)]. \quad (1.72)$$

Der Informationsgehalt ι gibt an, wie überraschend es ist, einen bestimmten Wert für X zu beobachten: je unwahrscheinlicher ein Wert ist, desto überraschender ist er. In diesem Sinne gibt die Entropie den mittleren Informationsgehalt einer Zufallsvariablen X an, der durch die Kodierung gespeichert werden muss.

Der Informationsgehalt kann auch axiomatisch hergeleitet werden. Dazu verlangen wir, dass

1. ι stetig ist,
2. $\iota(j) = f(p_j)$ für eine Funktion f , und
3. ι additiv ist.

Um den letzten Punkt genauer zu formulieren, seien $X, Y \sim p$ zwei iid. Zufallsvariablen mit Wahrscheinlichkeitsvektor $p \in [0, 1]^n$. Dann ist der Wahrscheinlichkeitsvektor von (X, Y) durch $P_{(i,j)} = p_i p_j$ mit $i, j \in \{1, \dots, n\}$. Additivität des Informationsgehalts bedeutet

$$\iota((X, Y)) = \iota(X) + \iota(Y). \quad (1.73)$$

Dass kann i.A. nur gelten, wenn

$$f(p_i p_j) = f(p_i) + f(p_j) \quad (1.74)$$

für alle i und j und alle Wahrscheinlichkeitsvektoren p . D.h. wir suchen eine nichttri-

viale Funktion $f : [0, 1] \rightarrow \mathbb{R}$, welche die Funktionalgleichung

$$f(xy) = f(x) + f(y) \quad (1.75)$$

erfüllt. Außerdem ist es sinnvoll zu fordern, dass der Informationsgehalt ι stetig ist. Daraus folgt $\iota(x) = \log_b(x)$ für alle $x \in [0, 1]$ und eine freie Konstante b (ohne Beweis).

1.4.2 Kullback-Leibler-Divergenz (Zusatzinfo)

Was passiert nun, wenn man in der Situation von Shannon-Kodierung einen Code für eine Wahrscheinlichkeitsverteilung p benutzt, der für eine andere Wahrscheinlichkeitsverteilung q optimiert wurde? In diesem Fall werde zusätzliche Bits zur Speicherung der von p gezogenen Daten notwendig (Satz 1.8). Man kann zeigen [9, Kapitel 2.3], dass im Limes vieler Daten die Zahl der zusätzlich benötigten Bits genau $ND(p\|q)$ beträgt, wobei N die Zahl der iid. gezogenen Buchstaben aus $\{a_1, \dots, a_m\}$ und $D(p\|q)$ wie folgt definiert ist.

Definition 1.9 (Relative Entropie):

Seien p und q zwei Zähldichten auf dem gleichen Ereignisraum. Die *relative Entropie* oder auch *Kullback-Leibler-Divergenz* zwischen p und q ist

$$D(p\|q) \equiv D_{\text{KL}}(p\|q) := - \sum_x p_x \ln \left(\frac{q_x}{p_x} \right). \quad (1.76)$$

Eigenschaften:

- (i) Für alle Wahrscheinlichkeitsverteilungen p und q ist

$$D(p\|q) \geq 0 \quad (1.77)$$

(der Beweis folgt später) und

$$D(p\|p) = 0. \quad (1.78)$$

Die Nichtnegativität werden wir später ausnutzen.

- (ii) Falls $p_x > 0$ für ein x mit $q_x = 0$ ist $D(p\|q) = \infty$.
 (iii) Im Allgemeinen ist $D(p\|q) \neq D(q\|p)$.

1.4.3 Differentielle Entropien

Wir benötigen auch einen Entropiebegriff für kontinuierlich verteilte Zufallsvariablen. Dazu fixieren wir eine Zufallsvariable X mit stetiger WDF $\rho : \mathbb{R} \rightarrow [0, 1]$. Dann führen wir eine Diskretisierung von \mathbb{R} ein,

$$\mathbb{R} = \bigcup_{i \in \mathbb{Z}} (i\Delta, (i+1)\Delta], \quad (1.79)$$

wobei $1 \gg \Delta > 0$. Dem Intervall $[\Delta i, \Delta(i+1)]$ können wir die Wahrscheinlichkeit

$$p_{\Delta,i} := \mathbb{P}[X \in (i\Delta, (i+1)\Delta]] = \int_{i\Delta}^{(i+1)\Delta} \rho(x) dx \approx \Delta \cdot \rho(i\Delta) \quad (1.80)$$

zuordnen. Dies führt zu der Entropie der diskretisierten Verteilung

$$\begin{aligned}
H_{\Delta}(\rho) &:= H(p_{\Delta}) = - \sum_{i \in \mathbb{Z}} \Delta \rho(i\Delta) \ln(\Delta \rho(i\Delta)) \\
&= - \sum_{i \in \mathbb{Z}} \Delta \rho(i\Delta) (\ln(\rho(i\Delta)) + \ln(\Delta)) \\
&\approx - \sum_{i \in \mathbb{Z}} \Delta \rho(i\Delta) \ln(\rho(i\Delta)) - \ln(\Delta).
\end{aligned} \tag{1.81}$$

Der Term $\ln(\Delta)$ hängt nicht von der Verteilung ρ ab, weshalb wir ihn einfach weglassen können. Daher definieren wir die (differentielle) Entropie von ρ als

$$H(\rho) := - \lim_{\Delta \searrow 0} \sum_{i \in \mathbb{Z}} \Delta \rho(i\Delta) \ln(\rho(i\Delta)) = - \int_{\mathbb{R}} \rho(x) \ln(\rho(x)) dx. \tag{1.82}$$

Die Verallgemeinerung dieser Idee auf mehrdimensionale Ereignisräume führt zu der folgenden Definition der (*differentiellen*) Entropie.

Definition 1.10 (Entropien für kontinuierliche Verteilungen):

Die *Entropie* eines kontinuierlich verteilten Zufallsvektors X mit WDF $\rho : \mathbb{R}^N \rightarrow [0, 1]$ ist als

$$H(X) := H(\rho) := - \int_{\mathbb{R}^N} \rho(x) \log(\rho(x)) dx \tag{1.83}$$

definiert.

Die *relative Entropie* zwischen WDFen $\rho : \mathbb{R}^N \rightarrow [0, 1]$ und $\sigma : \mathbb{R}^N \rightarrow [0, 1]$ ist als

$$D(\rho \parallel \sigma) := - \int_{\mathbb{R}^N} \rho(x) \log\left(\frac{\sigma(x)}{\rho(x)}\right) dx \tag{1.84}$$

definiert.

Die differentielle Entropie H erbt alle Eigenschaften der Shannon-Entropie, bis auf (i) und (ii), d.h. die Entropie kontinuierlicher Verteilungen kann beliebige Werte in \mathbb{R} annehmen und insbesondere negativ werden. Das wird uns aber nicht stören, da wir in der statistischen Physik an Entropiedifferenzen und -änderungen interessiert sind.

Die differentielle relative Entropie D erbt alle Eigenschaften der relativen Entropie. Insbesondere gilt

$$D(\rho \parallel \sigma) \geq 0 \tag{1.85}$$

für alle WDFen ρ und σ . Das liegt daran, dass die relative Entropie durch eine die relativen Entropien der diskretisierten Verteilungen approximiert werden kann: Eine ähnliche Diskretisierung wie (1.81) ergibt

$$\begin{aligned}
D_{\Delta}(\rho \parallel \sigma) &:= S(p_{\Delta} \parallel q_{\Delta}) = - \sum_{i \in \mathbb{Z}} \Delta \rho(i\Delta) \ln\left(\frac{\Delta \sigma(i\Delta)}{\Delta \rho(i\Delta)}\right) \\
&= - \sum_{i \in \mathbb{Z}} \Delta \rho(i\Delta) \ln\left(\frac{\sigma(i\Delta)}{\rho(i\Delta)}\right),
\end{aligned} \tag{1.86}$$

so dass

$$\lim_{\Delta \searrow 0} D_{\Delta}(\rho \parallel \sigma) = D(\rho \parallel \sigma). \tag{1.87}$$

1.5 Konvexität und Legendre-Transformation (5. VL)

Die konvexe Analysis beschäftigt sich mit konvexen Mengen und konvexen Funktionen. In der statistischen Mechanik begegnen wir häufig konkaven Funktionen, also Funktionen deren Negatives konvex ist. Daher sind einige wenige Aussagen aus der konvexen Analysis sehr hilfreich.

1.5.1 Legendre-Transformation (5. VL)

Eine umfassende Erläuterung der Legendre-Transformation in Bezug auf Anwendungen in der Physik kann in Ref. [11] gefunden werden und der vollständige mathematische Hintergrund in dem Buch [8].

Wir erinnern uns, dass die *Legendre-Transformierte* (oder das *Fenchel Duale* oder *konvex Konjugierte*) f^* einer Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$ über

$$f^*(y) := \sup_x \{\langle y, x \rangle - f(x)\} \quad (1.88)$$

definiert ist [8, Section 12]. Man kann zeigen, dass f^* immer eine konvexe unterhalbstetige Funktion auf \mathbb{R}^n ist [8, Theorem 12.1]. Wenn f ebenfalls konvex und unterhalbstetig ist gilt $(f^*)^* = f$ [8, Theorem 12.2].

Für die statistische Physik ist es hilfreich, die Legendre-Transformation für konkave Funktionen über

$$\begin{aligned} \tilde{f}(y) &:= \inf_x \{\langle y, x \rangle - f(x)\} \\ &= -(-f)^*(-y) \end{aligned} \quad (1.89)$$

zu definieren. Dann ist \tilde{f} immer eine konkave oberhalbstetige Funktion und falls f eine konkave und oberhalbstetige ist gilt

$$\tilde{\tilde{f}} = f \quad (1.90)$$

Nun nehmen wir an, dass f eine konkave oberhalbstetige Funktion ist und dass das Infimum in (1.89) für ein gegebenes y angenommen wird. Falls $\bar{x} = \bar{x}(y)$ ein kritischer Punkt der ebenfalls konkaven Funktion

$$x \mapsto f(x) - \langle y, x \rangle \quad (1.91)$$

ist, ergibt sich dass

$$\tilde{f}(y) = \langle y, \bar{x} \rangle - f(\bar{x}) \quad (1.92)$$

(kritische Punkte konkaver Funktionen sind globale Maxima). Die gleiche Gleichung gilt für die Legendre-Transformation im konvexen Fall.

Nun nehmen wir zusätzlich an, dass f stetig differenzierbar und ∇f invertierbar ist; d.h. es existiert eine Abbildung $h : \mathbb{R}^n \rightarrow \mathbb{R}^n$ so dass $y = h \circ \nabla f(y)$. Dann ist $\bar{x} = h(y)$, d.h.

$$\begin{aligned} \tilde{f}(y) &= \langle y, h(y) \rangle - f(h(y)) \\ &= \langle y, (\nabla f)^{-1}(y) \rangle - f((\nabla f)^{-1}(y)) . \end{aligned} \quad (1.93)$$

1.5.1.1 Homogene Funktionen 1. Grades (Zusatzinfo)

Hier betrachten wir ein Beispiel, wo die Legendretransformierte nicht mithilfe von Ableitungen berechnet wird. Dazu definieren wir die *Indikatorfunktion* $\delta_C : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$ einer konvexen Menge $C \subset \mathbb{R}^n$ über

$$\delta_C(y) := \begin{cases} 0 & \text{falls } y \in C, \\ \infty & \text{sonst} \end{cases} \quad (1.94)$$

und bemerken, dass diese Funktionen konvex sind. Wir nennen eine Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ *homogen p -ten Grades*, falls

$$f(\lambda x) = \lambda^p f(x) \quad \forall \lambda > 0. \quad (1.95)$$

Die Legendretransformierte einer konkaven Funktion 1. Grades $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ist

$$\begin{aligned} \tilde{f}(y) &= \inf_x \{ \langle y, x \rangle - \lambda f(x/\lambda) \} \\ &= \inf_x \{ \langle y, \lambda x \rangle - \lambda f(x) \} \\ &= \lambda \tilde{f}(y) \end{aligned} \quad (1.96)$$

für alle $\lambda > 0$. Daher ist $\tilde{f}(y) \in \{-\infty, 0\}$ für alle y . Genauer ist

$$\tilde{f}(y) = -\delta_C(y) \quad (1.97)$$

wobei

$$C := \{y : \langle y, x \rangle \geq f(x) \ \forall x\} \quad (1.98)$$

eine abgeschlossene konvexe Menge ist. Da f konkav ist, ist C konvex und \tilde{f} damit auch tatsächlich wieder konkav. Wenn f unterhalbstetig ist, so ist $f = \tilde{\tilde{f}} = \tilde{\delta}_C$. Das ergibt

$$f(x) = \tilde{\delta}_C(x) = \inf_{y \in C} \langle y, x \rangle. \quad (1.99)$$

Die Funktion $-\tilde{\delta}_C = \delta_C^*$ nennt man *Trägerfunktion* von C . Weitere Details zu Trägerfunktionen können in [8, Abschnitt 13] gefunden werden.

Freiwillige Übungsaufgabe: Zeigen Sie, dass für eine oberhalbstetige konkave homogene Funktion 1. Grades f tatsächlich $\tilde{\tilde{f}} = f$ ist.

1.5.2 Eine Konstruktion von konvexen Funktionen

Es gibt eine weitere mathematische Tatsache, die uns helfen wird, wenn wir die verallgemeinerte kanonische Verteilung herleiten.

Satz 1.11 ([8, Theorem 5.7]):

Sei $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ linear und $f : \mathbb{R}^n \rightarrow \mathbb{R}$ konvex. Dann ist die durch

$$f_A(y) := \inf \{ f(x) : Ax = y \} \quad (1.100)$$

definierte Funktion $f_A : \mathbb{R}^m \rightarrow \mathbb{R}$ ebenfalls konvex.

Beweis. Wir überprüfen direkt die Definition von konvexen Funktionen (1.4). Sei also $\lambda \in]0, 1[$ und $y_1, y_2 \in \mathbb{R}^m$. Dann ist

$$\begin{aligned} \lambda f_A(y_1) + (1 - \lambda) f_A(y_2) &= \inf \{ \lambda f(x_1) : Ax_1 = y_1 \} + \inf \{ (1 - \lambda) f(x_2) : Ax_2 = y_2 \} \\ &= \inf \{ \lambda f(x_1) + (1 - \lambda) f(x_2) : Ax_1 = y_1, Ax_2 = y_2 \} \\ &\geq \inf \{ f(\lambda x_1 + (1 - \lambda)x_2) : Ax_1 = y_1, Ax_2 = y_2 \} \\ &= \inf \{ f(x'_1 + x'_2) : Ax'_1 = \lambda y_1, Ax'_2 = (1 - \lambda)y_2 \}, \end{aligned} \quad (1.101)$$

wobei wir $x'_1 := \lambda x_1$ und $x'_2 := (1 - \lambda)x_2$ gesetzt haben. Dann ist

$$\begin{aligned} \lambda f_A(y_1) + (1 - \lambda) f_A(y_2) &\geq \inf \{ f(x'_1 + x'_2) : A(x'_1 + x'_2) = \lambda y_1 + (1 - \lambda)y_2 \} \\ &= f_A(\lambda y_1 + (1 - \lambda)y_2). \end{aligned} \quad (1.102)$$

□

1.6 1-Formen

(In der 9. VL behandelt.)

Wir erinnern uns an die Ableitung $df : \mathbb{R}^n \rightarrow L(\mathbb{R}^n, \mathbb{R})$ einer differenzierbaren Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ aus Punkt (vii), Abschnitt 1.2. Die Ableitung ist ein Beispiel für sogenannte 1-Formen.

Eine 1-Form (oder auch *Pfaffsche Form*) auf \mathbb{R}^n ist eine Abbildung

$$\omega : \mathbb{R}^n \rightarrow L(\mathbb{R}^n, \mathbb{R}). \quad (1.103)$$

Wir schreiben $\omega_p \in L(\mathbb{R}^n, \mathbb{R})$ für die entsprechende Linearform bei einem Punkt $p \in \mathbb{R}^n$. Ähnlich schreiben wir $d_p f \in L(\mathbb{R}^n, \mathbb{R})$ für df ausgewertet bei $p \in \mathbb{R}^n$.

Man kann ω immer als

$$\omega = \sum_{i=1}^n \phi_i dx_i \quad (\text{Basisentwicklung}) \quad (1.104)$$

schreiben, wobei $\phi_i : \mathbb{R}^n \rightarrow \mathbb{R}$ durch

$$\phi_i(p) = \omega_p(e_i) \quad (1.105)$$

gegeben sind mit den i -ten kanonischen Basisvektoren $e_i \in \mathbb{R}^n$.

1-Formen können entlang einer Kurve $\gamma : [a, b] \rightarrow \mathbb{R}^n$ integriert werden:

$$\int_{\gamma} \omega := \int_a^b \omega_{\gamma(t)}(\dot{\gamma}(t)) dt \quad (\text{Kurvenintegral}). \quad (1.106)$$

Wichtiges Beispiel: Ist $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ein Vektorfeld, so definiert

$$\omega^F := \langle F, \cdot \rangle \quad (1.107)$$

eine 1-Form mit $\omega_p^F = \langle F(p), \cdot \rangle \in L(\mathbb{R}^n, \mathbb{R})$. Das Kurvenintegral von ω^F entlang $\gamma : [a, b] \rightarrow \mathbb{R}^n$ ist dann

$$\int_{\gamma} \omega^F = \int_a^b \langle F(\gamma(t)), \dot{\gamma}(t) \rangle dt, \quad (1.108)$$

also das ganz gewöhnliche Kurvenintegral eines Vektorfeldes. Wenn man F als Kraft interpretiert nennt man das Integral (1.108) auch *Arbeitsintegral*.

Wir nennen eine 1-Form ω auf \mathbb{R}^n *geschlossen* falls die folgenden *Integrabilitätsbedingungen* für die Darstellung (1.104) erfüllt sind:

$$\frac{\partial \phi_i}{\partial x_j} = \frac{\partial \phi_j}{\partial x_i} \quad (1.109)$$

für alle $i, j \in [n]$. Wir nennen ω *exakt* falls eine differenzierbare Funktion f existiert, so dass

$$\omega = df. \quad (1.110)$$

Man nennt f dann *Potential* von ω und für eine beliebige Kurve $\gamma : [a, b] \rightarrow \mathbb{R}^n$ gilt dann

$$\int_{\gamma} \omega = f(\gamma(b)) - f(\gamma(a)), \quad (1.111)$$

d.h. das Arbeitsintegral ist wegunabhängig. Daraus folgt auch für jeden geschlossenen Weg γ dass

$$\oint_{\gamma} df = 0. \quad (1.112)$$

Verständnisaufgabe: Zeigen Sie, dass exakte Formen geschlossen sind.

Die Analysis lehrt uns, dass die Umkehrung auch gilt:

Satz 1.12:

Geschlossene 1-Formen auf dem \mathbb{R}^n sind exakt. Allgemeine 1-Formen mit wegunabhängigem Arbeitsintegral sind ebenfalls exakt.

Die Umkehrung der ersten Aussage des Satzes gilt nicht:

$$\omega = \frac{x dy - y dx}{x^2 + y^2} \quad (1.113)$$

ist eine geschlossene aber nicht exakte 1-Form (Übungsaufgabe).

1.6.1 Integrierende Faktoren

Gegeben eine 1-Form ω auf dem \mathbb{R}^n nennen wir eine Funktion $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ einen *integrierenden Faktor* für ω falls $\varphi \omega$ exakt ist.

In der Physik bezeichnen wir 1-Formen, die nicht exakt sind, häufig auch als $\omega = df$ oder $\omega = \delta f$.

Freiwillige Übungsaufgabe: Finden Sie eine nicht exakte 1-Form mit einem integrierenden Faktor.

1.6.2 Homogene Funktionen

(In der 16. VL behandelt.)

Wir nennen eine Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ *homogen k -ten Grades* falls

$$f(tx) = t^k f(x) \quad \forall t > 0. \quad (1.114)$$

Satz 1.13 (Eulersche Identität):

Seien $f : \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}$ eine differenzierbare Funktion k -ten Grades mit $k \geq 1$. Dann ist

$$kf = \sum_{j=1}^n \frac{\partial f}{\partial x_j} dx_j. \quad (1.115)$$

Beweis. Differenzieren von (1.114) liefert

$$\begin{aligned} kt^{k-1}f(x) &= \sum_{j=1}^n \frac{\partial f}{\partial x_j}(tx) \frac{\partial(tx_j)}{\partial t} \\ &= \sum_{j=1}^n \frac{\partial f}{\partial x_j}(tx) x_j. \end{aligned} \quad (1.116)$$

Auswerten bei $t = 1$ und die Identifikation (1.2) beendet den Beweis. \square

1.7 Spektralkalkül

(In der 13. VL behandelt.)

Hier wollen wir den Spektralkalkül für endlichdimensionale Hilberträume kurz erklären. Wir nennen einen Operator $A \in L(\mathcal{H})$ *normal*, falls

$$[A, A^\dagger] = 0. \quad (1.117)$$

Wichtige Beispiele für normale Operatoren sind selbstadjungierte ($A^\dagger = A$) und unitäre ($A^\dagger = A^{-1}$) Operatoren. Ist eine Basis von \mathcal{H} vorgegeben, so nennen $A \in L(\mathcal{H})$ *unitär diagonalisierbar*, falls ein unitärer Operator $U \in L(\mathcal{H})$ existiert, so dass

$$A = UDU^\dagger, \quad (1.118)$$

wobei $D = \text{diag}(a_1, \dots, a_{\dim(\mathcal{H})})$ *diagonal* ist. Dann lässt sich A auch bzgl. einer Orthonormalbasis (ONB) aus Eigenvektoren $\{|a_j\rangle\} \subset \mathcal{H}$ als

$$A = \sum_j a_j |a_j\rangle\langle a_j| \quad (1.119)$$

schreiben.

Aus der linearen Algebra ist der folgende Zusammenhang bekannt.

Satz 1.14 (Unitäre Diagonalisierung):

Ein Operator A auf einem Hilbertraum ist genau dann unitär diagonalisierbar wenn A normal ist.

Der *Spektralkalkül* hilft uns, Funktionen auf das Spektrum von normalen Operatoren anzuwenden und dadurch neue Operatoren zu erhalten.

Definition 1.15 (Spektralkalkül):

Sei $f : \mathbb{C} \rightarrow \mathbb{C}$ eine Funktion und $A \in L(\mathcal{H})$ ein normaler Operator von der Form (1.118). Dann definieren wir eine $f(A) \in L(\mathcal{H})$ als

$$f(A) = Uf(D)U^\dagger, \quad (1.120)$$

wobei man im Matrix-Fall $f(D)$ aus D erhält, indem man f auf die Diagonale der Diagonalmatrix D anwendet, $(f(D))_{i,j} := f(D_{i,j})\delta_{i,j}$.

Bemerkung: Der Spektralkalkül lässt sich auch auf den Fall $\dim(\mathcal{H})$ verallgemeinern: Wenn A ein rein diskretes Spektrum (Punktspektrum) wird $f(D)$ unverändert über die Eigenwerte von A berechnet. Falls A (zusätzlich) ein kontinuierliches Spektrum hat ist D ein Multiplikationsoperator mit einer Funktion $x \mapsto a(x)$ (wie z.B. für den Ortsoperator ist $a(x) = x$) und $f(D)$ ist durch $f \circ a$ gegeben.

Alternativ zu (1.120) kann man $f(A)$ über eine Potenzreihe $f(x) = \sum_j a_j x^j$ als

$$f(A) = \sum_{j=0}^{\infty} a_j A^j \quad (1.121)$$

definieren, wobei $A^0 := \mathbb{1}$ und $A^{j+1} := AA^j$ die Potenzen von A definiert.

Aufgabe: Zeigen Sie, dass die Definition (1.120) mit der Definition über Potenzreihen (1.121) übereinstimmt.

2 Klassische statistische Mechanik

In diesem Kapitel lernen wir die verschiedenen Ensemble der statistischen Mechanik (kanonisch, großkanonisch, mikrokanonisch, ...) und die damit verbundenen physikalischen Größen (Energie, Temperatur, Druck, ...) kennen. Dabei handelt es sich

um stochastische Beschreibungen von Vielteilchensystemen mittels ebendieser makroskopischen Größen.

Allgemein kann man Wahrscheinlichkeitsverteilungen (i) durch beobachtete Frequenzen (1.13) oder (ii) über Erwartungswerte schätzen/charakterisieren, wobei in der statistischen Physik typischerweise die letztere Strategie verfolgt wird. Ein makroskopischer thermischer Zustand wird hier durch makroskopische Größen beschrieben, die wir als Mittelwerte über mikrokanonische Zustände beschreiben. Allerdings gibt es hier typischerweise viel weniger makroskopische Größen als mikroskopische Freiheitsgrade, so dass diese bei weitem nicht ausreichen, um *eine* Wahrscheinlichkeitsverteilung festzulegen. Daraus ergibt sich die Frage, ob und wie man systematisch physikalisch sinnvolle Wahrscheinlichkeitsverteilung findet, die Vielteilchensysteme beschreiben. Ein besonders erfolgreicher Ansatz, der von E.T. Jaynes formalisiert wurde, ist eine Wahrscheinlichkeitsverteilung zu verwenden, welche unter den gegebenen makroskopischen Größen die Entropie maximiert. Dies führt zur *verallgemeinerten kanonischen Verteilung*, wie wir im folgenden Abschnitt sehen werden. Diese Verteilung verallgemeinert die klassischen Ensemble der statistischen Mechanik.

Von nun an werden wir die Entropie von einer Verteilung p mit

$$S(p) := -k \sum_x p_x \ln(p_x) \quad (2.1)$$

bezeichnen, wobei k eine Konstante ist (H bezeichnet traditionell auch Hamiltonfunktionen und die thermodynamische Enthalpie).

2.1 Die verallgemeinerte kanonische Verteilung

Dieser Abschnitt folgt vom Entwurf her [1, Kapitel 1.3] und wird mit wenig Mehraufwand mathematisch weitgehend rigoros präsentiert.

Wir betrachten ein Vielteilchensystem, das durch einen Ereignisraum Ω beschrieben ist. Weiterhin nehmen wir an, dass die relevanten makroskopische Größen als Erwartungswerte bzgl. der zu bestimmenden Wahrscheinlichkeitsverteilung schreibbar und als

$$\mathbb{E}[M^{(\nu)}] \equiv \langle M^{(\nu)} \rangle = M^{(\nu)} \quad \text{mit} \quad \nu = 1, 2, \dots, m \quad (2.2)$$

vorgegeben sind. Hier beschreibt jedes $M^{(\nu)}(x)$ eine physikalische Größe, die durch die Koordinate $x \in \Omega$ eindeutig bestimmt ist. Wir schreiben $M_i^{(\nu)}$ statt $M^{(\nu)}(x)$ für den diskreten Fall. Zur kürzeren Notation kann es auch hilfreich sein, $M = (M^{(\nu)})_{\nu \in [m]}$ als Zufallsvektor aufzufassen.

Beispiele:

- $\Omega = \mathbb{R}^{6N}$ ist der Phasenraum von N Teilchen und $M^{(1)}(x)$ ist die Energie eines Teilchens mit Phasenraumkoordinate x .
- $\Omega = \bigcup_{N=0}^{N_{\max}} \mathbb{R}^{6N}$ ist der Phasenraum von einem System das maximal N_{\max} vielen Teilchen enthält und $M^{(2)}(x)$ ist die Teilchenzahl, die zu x gehört.

Meistens interessiert man sich in der statistischen Physik nicht explizit dafür, wie der Ereignisraum genau gegeben ist.

Nun fehlt noch eine Wahrscheinlichkeitsverteilung für gegebene

$$M := (M^{(\nu)})_{\nu \in [m]}. \quad (2.3)$$

Hier möchten wir eine Verteilung wählen, die möglichst wenige Annahmen macht. Man spricht auch von einer *vorurteilsfreien Schätzung*. Dies führt zum Jaynessches Prinzip der maximalen Entropie, was in der informationstheoretischen Interpretation (Shannon-Kodierung, Satz 1.8) der Entropie begründet ist. Jaynes formulierte dies

wie folgt [12]: „[The] maximum entropy distribution may be asserted for the positive reason that it is uniquely determined as the one which is maximally noncommittal¹ with respect to missing information.“

Das Jaynessche Prinzip [12]:

Unter allen Zuständen eines physikalischen Systems, die kompatibel mit dem vorhandenen Wissen über das System sind, ist jener zu wählen, welcher die Entropie maximiert.

Genauer ist hier der Zustand des Systems der Zustand unseres Wissens über das System, welcher durch eine Wahrscheinlichkeitsverteilung (und später durch eine Dichtematrix) gegeben ist. Diese Auffassung ist kompatibel mit Bayesscher Inferenz [13, 14]

Für den Fall, dass die Zufallsvariablen $M^{(\nu)}$ jeweils maximal n verschiedene Werte annehmen können haben wir damit das folgende Minimierungsproblem unter Nebenbedingungen (NB) zu lösen (die Nebenbedingungen $p_i \geq 0$ lassen wir weg, da sie automatisch erfüllt sein werden):

$$\begin{aligned} & \text{minimiere} && -S(p) \\ & \text{unter NB} && \sum_{i=1}^n p_i = 1 \\ & && \sum_{i=1}^n M_i^{(\nu)} p_i = M^{(\nu)} \quad \forall \nu \in [m] \end{aligned} \quad (2.4)$$

Dies ist ein Minimierungsproblem mit stetiger konvexer Funktion $-S(p)$ unter affinen Nebenbedingungen. Die konvexe Analysis lehrt uns [8], dass hier jeder kritische Punkt, den man mittels Lagrange-Multiplikatoren finden kann, auch ein globales Minimum ist.

Die Lagrange-Funktion zu (2.4) ist als

$$\Lambda(p, \tilde{\lambda}) = k \sum_{i=1}^n p_i \ln(p_i) - \tilde{\lambda}_0 \left(1 - \sum_{i=1}^n p_i \right) - \sum_{\nu=1}^m \tilde{\lambda}_\nu \left(M^{(\nu)} - \sum_{i=1}^n M_i^{(\nu)} p_i \right) \quad (2.5)$$

gegeben, wobei $\tilde{\lambda}_0$ der Lagrange-Multiplikator zu $\sum_i p_i = 1$ und $\tilde{\lambda}_\nu$ jeweils der Lagrange-Multiplikator zu der NB bzgl. $M^{(\nu)}$ ist. Ab jetzt werden wir die Einsteinsche Summenkonvention (ESK) für die Summation über $\nu \in [m]$ verwenden. Dann erfüllt ein kritischer Punkt zu (2.4)

$$k \ln(p_i) + k + \tilde{\lambda}_0 + \tilde{\lambda}_\nu M_i^{(\nu)} = 0 \quad \forall i \in [n]. \quad (2.6)$$

Das ergibt

$$p_i = \exp\left(\Psi - \lambda_\nu M_i^{(\nu)}\right) \quad (2.7)$$

mit $\Psi := -\frac{(k+\lambda_0)}{k}$ und

$$\lambda_\nu := \tilde{\lambda}_\nu / k. \quad (2.8)$$

Eine analoge Argumentation kann man für den kontinuierlichen Fall unter Verwendung von Variationsrechnung machen (Übungsblatt). Zusammengefasst ergibt das folgendes.

¹sich nicht festlegen wollen, unverbindlich

Verallgemeinerte kanonische Verteilung

Bei gegebenen Erwartungswerten (2.2) ist die *verallgemeinerte kanonische Verteilung* die Verteilung, welche die Entropie (s. Abschnitt 1.4) maximiert und ist als

$$p_i = \exp(\Psi - \lambda_\nu M_i^{(\nu)}) \quad (\text{diskreter Fall}), \quad (2.9)$$

$$\rho(x) = \exp(\Psi - \lambda_\nu M^{(\nu)}(x)) \quad (\text{kontinuierlicher Fall}) \quad (2.10)$$

(mit ESK) gegeben, wobei Ψ die eine Normierungskonstante ist und $(\lambda_\nu)_{\nu \in [m]}$ noch zu bestimmende reelle Konstanten sind, die durch (2.2) festgelegt sind.

Die Entropie der verallgemeinerten kanonischen Verteilung wird mit $S(\mathbf{M})$ bezeichnet.

5. VL

Die Legendre-Transformation und ein Satz aus der konvexen Analysis (Abschnitt 1.5.1) wurde ebenfalls in der 5. VL behandelt.

Bemerkungen:

- Da Ψ dem Lagrange-Multiplikator zur Normiertheit entspricht betrachten wir Ψ als Normierungskonstante. Genauer ist die Normierungskonstante

$$\begin{aligned} Z(\lambda) &:= \sum_i \exp(-\lambda_\nu M_i^{(\nu)}) = e^{-\Psi(\lambda)} & (\text{diskreter Fall}), \\ Z(\lambda) &:= \int dx \exp(-\lambda_\nu M^{(\nu)}(x)) = e^{-\Psi(\lambda)} & (\text{kontinuierlicher Fall}) \end{aligned} \quad (2.11)$$

die sogenannte *Zustandssumme*.

- Die Lagrange-Multiplikatoren λ werden als Funktion der Erwartungswerte \mathbf{M} aufgefasst. Dadurch hängt die kanonische Verteilung implizit von den Erwartungswerten \mathbf{M} ab. Diese Abhängigkeit werden wir im Folgenden genauer charakterisieren.
- Wegen Satz 1.11 ist die Entropiefunktion $\mathbf{M} \mapsto -S(\mathbf{M})$ konvex.

Die Entropie der kanonischen Verteilung ergibt sich aus

$$\begin{aligned} \frac{S(\mathbf{M})}{k} &= \begin{cases} -\sum_i p_i \ln(p_i) & (\text{diskreter Fall}) \\ -\int \rho(x) \ln(\rho(x)) dx & (\text{kontinuierlicher Fall}) \end{cases} \\ &= \mathbb{E}[-\Psi(\lambda) + \lambda_\nu M^{(\nu)}] \\ &= \lambda_\nu \mathbf{M}^{(\nu)} - \Psi(\lambda) \\ &\equiv \langle \lambda, \mathbf{M} \rangle - \Psi(\lambda) \end{aligned} \quad (2.12)$$

und daher

$$k\Psi(\lambda) = k\langle \lambda, \mathbf{M} \rangle - S(\mathbf{M}). \quad (2.13)$$

Außerdem folgt aus (2.12) dass

$$\frac{\partial S(\mathbf{M})}{\partial \mathbf{M}^{(\nu)}} = k\lambda_\nu, \quad (2.14)$$

d.h. \mathbf{M} ist stationärer Punkt von $\mathbf{M} \mapsto k\langle\lambda, \mathbf{M}\rangle - S(\mathbf{M})$ weil somit

$$\frac{\partial}{\partial \mathbf{M}^{(\nu)}} \left[k\langle\lambda, \mathbf{M}\rangle - S(\mathbf{M}) \right] = k\lambda_\nu - \frac{\partial S(\mathbf{M})}{\partial \mathbf{M}^{(\nu)}} = 0. \quad (2.15)$$

Nach (1.92) ist daher

$$k\Psi(\lambda) = \tilde{S}(k\lambda) \quad (2.16)$$

sofern das Infimum in der Definition der Legendre-Transformation (1.89) angenommen wird. Dann ist Ψ die Legendre-Transformierte von S . Außerdem ist Ψ als Legendre-Transformierte² einer konkaven Funktion ebenfalls konkav, s. Abschnitt 1.5.1.

Aus der Normierungsbedingung (2.11) folgt weiterhin, dass

$$\Psi(\lambda) = -\ln \sum_i \exp\left(-\lambda_\nu M_i^{(\nu)}\right) \quad (2.17)$$

und daher

$$\begin{aligned} \frac{\partial \Psi(\lambda)}{\partial \lambda_\nu} &= -\frac{\sum_i (-M_i^{(\nu)}) \exp(-\lambda_\nu M_i^{(\nu)})}{\sum_i \exp(-\lambda_\nu M_i^{(\nu)})} \\ &= \mathbb{E}[M^{(\nu)}] = \mathbf{M}^{(\nu)}. \end{aligned} \quad (2.18)$$

Daraus folgt, dass

$$k\tilde{\Psi}(\mathbf{M}) = S(\mathbf{M}), \quad (2.19)$$

was (1.90) bestätigt. Die Lagrange-Dualität zwischen S und Ψ kann auch direkt aus $S(\mathbf{M}) + k\Psi(\lambda) = k\langle\lambda, \mathbf{M}\rangle$ gefolgert werden. Unsere Herleitung ist allerdings etwas expliziter.

Wir halten fest, dass außerdem gilt

$$-\frac{\partial \ln(Z)}{\partial \lambda_\nu} = \mathbf{M}^{(\nu)}, \quad (2.20)$$

d.h. die logarithmischen Ableitungen der Zustandssumme liefern Erwartungswerte $\mathbf{M}^{(\nu)}$.

6. VL

Abschnitt 1.2, Punkt (vii) zu 1-Formen wurden ebenfalls in der 6. VL behandelt.

Zum Abschluss dieses Abschnitts sei noch erwähnt, dass man Gleichung (2.14) häufig auch wie folgt schreibt (zur Notation siehe auch Punkt (vii) in Abschnitt 1.2).

Verallgemeinerte Gibbssche Fundamentalgleichung

$$dS = k\lambda_\nu d\mathbf{M}^{(\nu)} \quad (\text{mit ESK}). \quad (2.21)$$

Ähnlich hilfreich ist die Gleichung

$$d \ln(Z) = -\mathbf{M}^{(\nu)} d\lambda_\nu \quad (\text{mit ESK}). \quad (2.22)$$

2.1.1 Beispiel: Kanonische Verteilung

Der Phasenraum von N unterscheidbaren Teilchen, die sich einer Region $R \subset \mathbb{R}^3$ befinden, ist

$$\Gamma_u = R^N \times \mathbb{R}^{3N} \quad (2.23)$$

²mit der in der statistischen Physik üblichen Vorzeichenkonvention der Legendre-Transformation, so dass konkave Funktionen auf konvex abgebildet werden

mit zugehörigen Phasenraumkoordinaten

$$\xi = (q_1, \dots, q_{3N}, p_1, \dots, p_{3N})^\top. \quad (2.24)$$

Wir nehmen an, dass die Teilchen durch eine Hamiltonfunktion der Form

$$H(\xi) = \frac{p^2}{2m} + V(q) \quad (2.25)$$

beschrieben sind, wobei m die Masse der einzelnen Teilchen und $p^2 := \sum_{i=1}^{3N} p_i^2$ ist. Nun betrachten wir die Phasenraumfunktion H als einzige Zufallsvariable,

$$M^{(1)}(\xi) = H(\xi), \quad (2.26)$$

und nehmen an, dass sie den Energieerwartungswert

$$\mathbb{E}[H] = U \quad (2.27)$$

hat. Den zugehörigen Lagrange-Parameter bezeichnen wir mit

$$\beta \equiv \lambda_1. \quad (2.28)$$

Dann ist die verallgemeinerte kanonische Verteilung durch die WDF

$$\rho(\xi) = \frac{1}{Z_u} e^{-\beta H(\xi)} \quad (2.29)$$

gegeben, wobei die Zustandssumme Z_u in den üblichen Einheiten als

$$Z_u = \int_{\mathbb{R}^{6N}} \frac{d^{3N}q d^{3N}p}{h^{3N}} e^{-\beta H(\xi)} \quad (2.30)$$

gegeben ist, wobei wir die Einschränkung $q \in \mathbb{R}^N$ in dem Potential $V(q)$ berücksichtigen und mit h das Plancksche Wirkungsquantum bezeichnen.

Wenn nun die Teilchen ununterscheidbar sind, müssen wir zwei Phasenraumkoordinaten ξ_1, ξ_2 als äquivalent betrachten, d.h. $\xi_1 \sim \xi_2$ falls ξ_1 aus einer Permutation der N Teilchen aus ξ_2 hervorgeht. Da es $N!$ viele solcher Permutationen gibt, wählen wir die Zustandssumme von N ununterscheidbaren Teilchen als

$$Z = \frac{Z_u}{N!} = \int_{\mathbb{R}^{6N}} d\xi e^{-\beta H(\xi)}, \quad (2.31)$$

mit dem Phasenraum-Volumenelement

$$d\xi := \frac{d^{3N}q d^{3N}p}{h^{3N} N!} \quad (2.32)$$

wobei wir den Faktor $N!$ mithilfe der Quantenmechanik noch besser begründen werden. Die zugehörige Verteilung ist die *kanonische Verteilung* auch (*Boltzmann-Verteilung* genannt) und ist durch

$$\rho(\xi) = \frac{1}{Z} e^{-\beta H(\xi)} \quad (2.33)$$

gegeben.

Wir definieren die *Temperatur* T über

$$\frac{\partial S}{\partial U} =: \frac{1}{T}, \quad (2.34)$$

was durch Verwendung physikalisch sinnvoller expliziter Hamiltonfunktionen und später mit der Phänomenologie der Thermodynamik (Kapitel 4) gerechtfertigt wird. Mit

(2.14) gilt dann

$$\frac{\partial S}{\partial U} = k\beta = \frac{1}{T} \quad (2.35)$$

und somit

$$\beta = \frac{1}{kT}, \quad (2.36)$$

wobei wir $k = k_B$ als Einheitenkonstante wählen.

2.1.2 Beispiel: Großkanonische Verteilung

Bei der großkanonischen Verteilung handelt es sich um eine Erweiterung der kanonischen Verteilung, wo die Teilchenzahl $N = M^{(2)}$ ebenfalls als Zufallsvariable betrachtet wird. Das führt zu folgenden Modifikationen:

- Phasenraum $\Gamma = \bigcup_{n=0}^{\infty} \mathbb{R}^{6n}$ mit der Notation $\xi_n \in \mathbb{R}^{6n}$
- $M^{(2)} = N$ ist die Teilchenzahl als Phasenraumfunktion so dass $N(\xi_n) = n$. Normalerweise werden hier sowohl die Zufallsvariable N als auch die natürliche Zahl n beide mit N bezeichnet.
- $\lambda_2 =: -\beta\mu$ (Konvention), wobei μ *chemisches Potential* heißt
- Die Zustandssumme wird mit \mathcal{Z} bezeichnet und ist durch

$$\mathcal{Z} = e^{-\Psi} = \sum_{n=0}^{\infty} \int_{\mathbb{R}^{6n}} d\xi_n e^{-\beta(H(\xi_n) - \mu N(\xi_n))} \quad (2.37)$$

mit $d\xi_n = \frac{d^{3n}q d^{3n}p}{h^{3n} n!}$ gegeben.

- Die *großkanonische Verteilung* ist dann durch

$$\rho(\xi) = \frac{1}{\mathcal{Z}} e^{-\beta(H(\xi) - \mu N(\xi))} \quad (2.38)$$

gegeben.

- Mit der *Marginalverteilung*

$$p_n := \frac{1}{\mathcal{Z}} \int_{\mathbb{R}^{6n}} d\xi_n e^{-\beta(H(\xi_n) - \mu n)} \quad (2.39)$$

von N ist z.B.

$$\mathbb{E}[N] = \sum_{n=0}^{\infty} n p_n. \quad (2.40)$$

2.1.2.1 Ideales Gas

Ein *ideales Gas* ist durch die Hamiltonfunktion

$$H(\xi) = \frac{p^2}{2m} + V_{\text{Wand}}(q) \quad (2.41)$$

definiert, wobei das Potential $V_{\text{Wand}}(q)$ eine harte Wand modelliert, die eine zusammenhängende kompakte Region R mit Volumen V einschließt, d.h.

$$V_{\text{Wand}}(q) = \begin{cases} 0 & \text{falls } q \in R, \\ \infty & \text{sonst.} \end{cases} \quad (2.42)$$

Das ideale Gas entspricht der Idealisierung punktförmiger aber hart stoßender Teilchen.

Übung: Berechnen Sie die großkanonische Verteilung und Zustandssumme für ein ideales Gas.

2.1.3 Ein Dissipations-Fluktuations-Satz

Satz 2.1 (Dissipations-Fluktuations-Satz):

Die Zufallsvariablen $M^{(\nu)}$ der verallgemeinerten kanonischen Verteilung erfüllen

$$\text{Cov}(M^{(\nu)}, M^{(\alpha)}) = -\frac{\partial^2 \Psi(\lambda)}{\partial \lambda_\nu \partial \lambda_\alpha} = -\frac{\partial M^{(\nu)}}{\partial \lambda_\alpha}. \quad (2.43)$$

Bemerkungen:

- $\text{Cov}(M^{(\nu)}, M^{(\alpha)})$ misst die Fluktuationen von $M^{(\nu)}$ und $M^{(\alpha)}$ um den Erwartungswert. Der Satz sagt uns, dass sie mit den Änderungen der Erwartungswerte, auch Suszeptibilitäten genannt, übereinstimmen.
- Die durch $\text{Cov}(M^{(\nu)}, M^{(\alpha)})$ definierte Matrix heißt auch Kovarianzmatrix und ist immer positiv-semidefinite (kurze Übungsaufgabe). Das bestätigt, dass Ψ konkav ist (kurze Übungsaufgabe).

Beweis. Die Kummulantenerzeugende (s. Abschnitt 1.3.2.4) lässt sich auf unsere Zufallsvektoren M verallgemeinern und ist dann durch

$$\ln \varphi(q) = \ln \mathbb{E}[e^{i\langle q, M \rangle}] \quad (2.44)$$

gegeben. Durch Rechnungen mit den entsprechenden Reihenentwicklungen kann man zeigen, dass

$$\begin{aligned} -\left. \frac{\partial^2 \ln \varphi(q)}{\partial q_\nu \partial q_\alpha} \right|_{q=0} &= \mathbb{E}[M^{(\nu)} M^{(\alpha)}] - \mathbb{E}[M^{(\nu)}] \mathbb{E}[M^{(\alpha)}] \\ &= \text{Cov}(M^{(\nu)}, M^{(\alpha)}), \end{aligned} \quad (2.45)$$

was sich durch direktes differenzieren von $\ln(\varphi)$ überprüfen lässt (Übungsblatt). Die Kummulantenerzeugende der kanonischen Verteilung ist

$$\begin{aligned} \ln \varphi(q) &= \ln \sum_i p_i \exp(i\langle q, M_i \rangle) \\ &= \ln \sum_i p_i \exp(\Psi(\lambda) - \langle \lambda - iq, M_i \rangle) \\ &= \Psi(\lambda) + \ln \sum_i \exp(-\langle \lambda - iq, M_i \rangle) \\ &= \Psi(\lambda) - \Psi(\lambda - iq), \end{aligned} \quad (2.46)$$

wobei wir (2.11) im letzten Schritt benutzt haben. Eine analoge Rechnung liefert die Gleiche Identität für den kontinuierlichen Fall. Damit ist

$$\begin{aligned} \text{Cov}(M^{(\nu)}, M^{(\alpha)}) &= \left. \frac{\partial^2 \Psi(\lambda - iq)}{\partial q_\nu \partial q_\alpha} \right|_{q=0} \\ &= -\frac{\partial^2 \Psi(\lambda)}{\partial \lambda_\nu \partial \lambda_\alpha} \end{aligned} \quad (2.47)$$

und Gleichung (2.18) beendet den Beweis. \square

2.2 Begründung der Gleichgewichtsthermodynamik

7. VL

Wir werden das Jaynessche Prinzip der maximalen Entropie auf physikalische Vielteilchensysteme anwenden, was uns verallgemeinerte kanonische Verteilungen (Abschnitt 2.1) liefern wird. Dabei betrachten wir Situationen, wo die Zufallsvariable $M^{(\nu)}$ als Phasenraumvariablen, die makroskopischen Größen entsprechen, gewählt werden. Dies wird uns zur Gleichgewichtsthermodynamik (Kapitel 4) führen. Zuerst stellen wir jedoch die folgende Frage.

Grundlegende Frage der statistischen Physik

Wann wird ein System tatsächlich durch die verallgemeinerte kanonische Verteilung korrekt beschrieben?

Diese Frage ist so alt wie die Thermodynamik selbst ihre Details, insbesondere in Hinblick auf die Quantenmechanik, beschäftigen noch heutige Wissenschaftler [15–17]. Ein Hauptaugenmerk liegt dabei auf Systemen, die sich im thermischen Gleichgewicht befinden. Hier werden wir nur kurz die bekanntesten Argumente vorstellen, mit denen ein solches Gleichgewicht begründet wird und dass dann die verallgemeinerte kanonische Verteilung eine korrekte Beschreibung liefert. Gleichzeitig werden wir die Gleichgewichtsthermodynamik begründen.

Wie bereits in den Abschnitten 2.1.1 und 2.1.2 angedeutet, wenden wir das Jaynessche Prinzip wie folgt auf Hamiltonsche Vielteilchensysteme mit Phasenraum Γ und Hamiltonfunktion $H : \Gamma \rightarrow \mathbb{R}$ an. Wir bezeichnen die “makroskopisch relevanten” Observablen (klassisch gegeben durch Phasenraumfunktionen) mit $M^{(1)}, \dots, M^{(m)} : \Gamma \rightarrow \mathbb{R}$. Wir nehmen an, dass ihre Erwartungswerte

$$\mathbb{E}[M^{(\nu)}] = M^{(\nu)} \quad \nu \in [m] \quad (2.48)$$

vorgegeben sind und gelangen über das Jaynessche Prinzip zur verallgemeinerten kanonischen Verteilung

$$\rho(\xi) = e^{\Psi(\lambda) - \langle \lambda, M(\xi) \rangle} \quad (2.49)$$

mit Lagrange-Multiplikatoren $\lambda = \lambda(M)$ und Zustandssumme

$$Z(\lambda) = e^{-\Psi(\lambda)} = \int d\xi e^{-\langle \lambda, M(\xi) \rangle}, \quad (2.50)$$

s. Abschnitt 2.1.

Das Ziel ist, mit (2.49) die makroskopische Gleichgewichtsthermodynamik des Vielteilchensystems zu beschreiben. Daher machen wir die folgende Annahme.

Erhaltungsgrößen-Annahme

Die Observablen $M^{(\nu)} : \Gamma \rightarrow \mathbb{R}$ sind Erhaltungsgrößen, d.h.

$$\frac{dM(\xi(t))}{dt} = 0, \quad (2.51)$$

wobei $t \mapsto \xi(t)$ durch Anfangsbedingungen bestimmt ist.

Es reicht aus, wenn diese Annahme approximativ erfüllt ist, d.h. dass sich die makroskopischen Erwartungswerte M viel langsamer verändern, als die Zeit, die das System braucht um, *ins makroskopische Gleichgewicht zu relaxieren*. Diesen Prozess werden wir in Abschnitt 2.2.2 etwas genauer diskutieren. Vorher werden wir besser motivieren, dass die Verteilung (2.49) im Kontext von Vielteilchensystemen Sinn ergibt und insbesondere *stationär* ist.

2.2.1 Der Satz von Liouville

Wir betrachten ein Hamiltonisches System mit kanonischen Orts- und Impulskoordinaten

$$\xi = (q_1, \dots, q_N, p_1, \dots, p_N). \quad (2.52)$$

Dann ist die Dynamik gegeben durch die kanonischen Gleichungen

$$\dot{q}_k = \frac{\partial H}{\partial p_k}, \quad \dot{p}_k = -\frac{\partial H}{\partial q_k}. \quad (2.53)$$

Nun nehmen wir an, dass die Phasenraumkoordinaten ξ gemäß einer WDF $\rho(\xi)$ verteilt sind. Dann kann man $\rho(\xi)\dot{\xi}$ als Stromdichte und ρ als Dichte im Phasenraum interpretieren. Tatsächlich erfüllen diese Größen die Kontinuitätsgleichung

$$\dot{\rho} + \operatorname{div}(\rho \dot{\xi}) = 0, \quad (2.54)$$

was man mit der üblichen Argumentation zeigen kann.

Mit

$$\begin{aligned} \operatorname{div}(\rho \dot{\xi}) &= \langle \operatorname{grad}(\rho), \dot{\xi} \rangle + \rho \operatorname{div}(\dot{\xi}) \\ &= \langle \operatorname{grad}(\rho), \dot{\xi} \rangle \\ &= \sum_{k=1}^N \left(\frac{\partial \rho}{\partial q_k} \dot{q}_k + \frac{\partial \rho}{\partial p_k} \dot{p}_k \right) \end{aligned} \quad (2.55)$$

erhalten wir dann totale Zeitableitung der WDF als

$$\begin{aligned} \frac{d\rho}{dt} &= \frac{\partial \rho}{\partial t} + \sum_{k=1}^N \left(\frac{\partial \rho}{\partial q_k} \dot{q}_k + \frac{\partial \rho}{\partial p_k} \dot{p}_k \right) \\ &= \dot{\rho} + \operatorname{div}(\rho \dot{\xi}). \end{aligned} \quad (2.56)$$

Zusammen mit der Kontinuitätsgleichung (2.54) erhalten wir das folgende Ergebnis.

Satz 2.2 (Satz von Liouville):

Sei $\rho : \Gamma \rightarrow [0, 1]$ eine WDF auf dem Phasenraum Γ eines konservativen Hamiltonischen Systems. Dann gilt

$$\frac{d\rho}{dt} = 0. \quad (2.57)$$

Bemerkungen:

- Der Satz besagt, dass sich WDFen auf dem Phasenraum unter Dynamik nicht ändern. Man nennt eine solche Verteilung auch *stationär*.
- Der Satz gilt allgemein und nicht nur für unsere Verteilung (2.49).
- Die Phasenraumzellen können sich trotzdem dramatisch mit der Zeit verformen können, solange ihr Volumen gleich bleibt. Die Wahrscheinlichkeitsdichte verhält sich also ähnlich, wie die Dichte einer inkompressiblen bewegten Flüssigkeit.

2.2.2 Gleichgewichtsverteilungen und thermodynamische Variablen

Wir betrachten ein mikroskopisches System, das durch eine zeitliche Entwicklung $\xi(t)$ im Phasenraum charakterisiert ist. Außerdem nehmen wir an, dass die zugänglichen makroskopischen Observablen, d.h. Phasenraumfunktionen durch $M = (M^{(1)}, \dots, M^{(m)})^\top$ gegeben sind. Dann werden makroskopische Beobachtungen durch das Zeitmittel

$$\overline{M} = \frac{1}{\tau} \int_0^\tau dt M(\xi(t)) \quad (2.58)$$

gegeben, wobei τ die Beobachtungsdauer ist. Wenn diese viel größer ist, als die Zeit mikroskopischer Änderungen, dann ist zu erwarten, dass das Zeitmittel (2.58) konstant ist, d.h. unabhängig vom Zeitfenster $[0, T]$ solange dieses nur groß genug ist. In diesem Fall sagt man, dass M *äquilibriert* ist.

Eine *Gleichgewichtsverteilung* ist eine Verteilung ρ , so dass die folgende *Ergodizitätsannahme* erfüllt ist:

$$\overline{M} = \int d\xi M(\xi) \rho(\xi). \quad (2.59)$$

In diesem Fall nennt man (M, ρ) auch ein *Ensemble*.

Es gibt zwei Argumentationswege, die Ergodizitätsannahme (2.59) zu zeigen:

- (i) Eine Anwendung des Jaynesasche Prinzips auf den Phasenraum ergibt: **Zugängliche Phasenraumregionen gleichen Volumens sind gleich Wahrscheinlich.** Daraus folgt (2.59). Man beachte, dass der Satz von Liouville essenziell für dieses Argument ist, da Phasenraumvolumina zeitlich nicht veränderlich sein dürfen.

Wenn die Erwartungswerte von $\mathbb{E}[M^{(\nu)}]$ bekannt sind, liefert das Jaynessche Prinzip die entsprechende verallgemeinerte kanonische Verteilung.

- (ii) Eine Argumentation, die viel älter als die Informationstheorie ist, basiert auf der folgende *Ergodenhypothese*. Ihr Ursprung geht auf Boltzmann (1898) zurück.

In einem konservativen System ist die Energie $H(\xi)$ eine Erhaltungsgröße, d.h. jede durch $H(\xi) = E$ definierte Energieschale ist unter der Dynamik erhalten.

Die *Ergodenhypothese* besagt: **In ausreichen langer Zeit wird jede Zelle auf der Energieschale durchlaufen**, wobei eine *Zelle* ein Bereich positiven Volumens bzw. in diesem Fall positiver Hyperfläche ist.

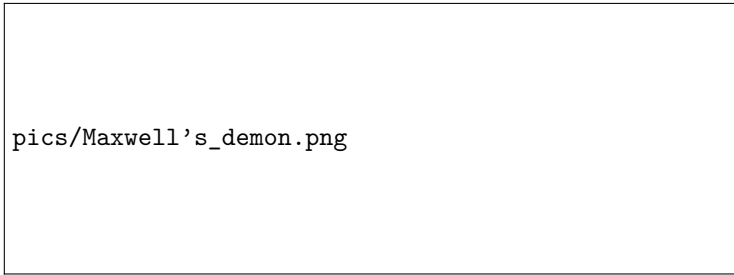
Der Satz von Liouville impliziert, dass die Zeit, welche das System in Zellen gleicher Größe verbringt gleich ist. Damit folgt (2.59), allerdings nur für $\tau \rightarrow \infty$ und bestimmte *nicht-integrable Systeme*³ (im Sinne der analytischen Mechanik), was von Birkhoff und von Neumann in den 1930ern mathematisch bewiesen wurde. Hierbei spielen bestimmte weitere Eigenschaften von *strong mixing* eine Rolle.

Für einen Beobachter, der nur den Energiewert E kennt aber keine sonstige Kenntnis über $\xi(0)$ hat, ist die Zustandsverteilung durch die *mikrokanonische Verteilung* $\delta_{H(\xi)-E}$ gegeben. Aus dieser Verteilung lassen sich auch sämtliche Ensemble herleiten, welche wir aber direkt über die verallgemeinerte kanonische Verteilung herleiten werden.

In den meisten interessanten Situationen sind die Observablen $M^{(\nu)}$ nicht wirklich in einem Gleichgewicht, sondern verändern sich nur sehr viel langsamer als die Zeitskala, auf der die mikroskopische Äquilibration stattfindet. Wenn eine solche Trennung der relevanten Zeitskalen (Beobachtungszeit vs. mikroskopische Veränderungen) vorliegt spricht man auch von einem *Quasigleichgewicht*. Ähnlich spricht man bei langsamer Veränderung von Phasenraumfunktionen auch von Quasierhaltungsgrößen. Beispiele sind Energie, Volumen und Teilchenzahl, die sich langsam im Vergleich zur mikroskopischen Dynamik verändern. In der Gleichgewichtsthermodynamik nimmt man grundsätzlich an, dass sich die Quasierhaltungsgrößen so langsam ändern, dass das System durchgehend durch den entsprechenden Gleichgewichtszustand (verallgemeinerte kanonische Verteilung) beschrieben werden kann.

Allgemein nennt man ein System mit Erhaltungsgrößen $M^{(\nu)}$ *thermodynamisch*, falls es durch die verallgemeinerte kanonische Verteilung beschrieben werden kann. Die damit verbundenen Variablen $M^{(\nu)}$ und λ_ν nennt man *thermodynamische Variablen*.

³Ein konservatives Hamiltonisches System mit N Freiheitsgraden heißt *integrabel*, falls es N unanghängige Erhaltungsgrößen hat. (z.B. wie bei einem gekoppelten harmonischen Oszillator).



pics/Maxwell's_demon.png

Abbildung 2.1: Maxwells Dämon kann die Teilchen verschiedener Art „sehen“ und somit die Mischungsentropie zweier Gase mithilfe einer Klappe verringert.

2.2.3 Zeit und Irreversibilität

Eine Dynamik $\xi(t)$ heißt *reversibel*, falls sich aus ihr bei Zeitumkehr wieder ein physikalisch möglicher Prozess ergibt, d.h. dass $\xi(-t)$ nicht fundamentalen physikalischen Gesetzen widerspricht. Wichtige Beispiele für irreversible Prozesse sind Wärmeleitung, Diffusion und Mischung von Fluiden.

Mikroskopische Hamiltonsche Dynamik ist reversibel, wenn keine weiteren Einschränkungen gemacht werden, während die makroskopische Dynamik (z.B. Äquilibration) i.A. irreversibel ist.

Eine Funktion auf einem physikalischen System heißt *Zustandsfunktion* wenn sie nur vom aktuellen Zustand des Systems und nicht von seiner Historie abhängt.

Die Existenz von irreversiblen Prozessen wird durch folgende Aussage impliziert: *Für jedes isolierte thermodynamische System existiert eine Zustandsfunktion, die mit der Zeit nicht abnehmen kann.* Diese Größe wird nach dem Vorschlag von R. Clausius (~1860) als *Entropie* bezeichnet.

Wir stellen fest, dass das Jaynessche Prinzip bereits die Irreversibilität der durch eine verallgemeinerte kanonische Verteilung beschriebenen System impliziert:

2. Hauptsatz der Thermodynamik

Die Entropie S eines vollständig isolierten Systems kann mit der Zeit nicht abnehmen.

Der 2. Hauptsatz schließt die Existenz von Perpetua mobilia (2. Art) aus. Es kann z.B. bei zwei isolierten Systemen unterschiedlicher Temperatur keine Wärme vom kälteren zum wärmeren System übergehen, wie wir sehen werden. Damit ist etwa ein Schiffsantrieb, der dem Meer Wärme entzieht und damit direkt den Schiffsmotor antreibt nicht möglich.

Als weiteres Beispiel könnte in einem Gedankenexperiment ein sogenannter Maxwell'scher Dämon (1871) Information über ein System ausnutzen, um Mischungsentropie zu verringern, s. Abbildung 2.1. Viel später wurde jedoch gezeigt, dass der Klappmechanismus des Dämons entweder einen „Antrieb“ braucht und damit Entropie in Form von Wärme erzeugt oder in beide Richtungen schwingen muss, womit keine Trennung der Gase möglich ist.

Als einen grundlegenden Fortschritt bezüglich der Verbindung zwischen Informationstheorie und Thermodynamik hat Landauer folgendes gezeigt: Nichtreversible Verarbeitung von Informationen (Löschung von Bits) erzeugt thermodynamische Entropie, die als Wärme in Erscheinung tritt.

8. VL

Abschnitte 1.3.1.5, 1.3.1.6, und die Eigenschaft (iv) (Abschnitt 1.4.1) der Entropie wurden ebenfalls in dieser Vorlesung behandelt.

2.2.4 Extensive und intensive Größen

Vielteilchensysteme lassen sich zusammensetzen, d.h. wenn wir ein System Γ_I und ein zweites System Γ_{II} haben ist das zusammengesetzte System $\Gamma_I \times \Gamma_{II}$ (kartesisches Produkt). Ähnliches gilt für Phasenraumfunktionen: Wenn M_I und M_{II} Phasenraumfunktionen auf den beiden Einzelsystemen sind, ist $M_I + M_{II}$ eine Phasenraumfunktion auf dem zusammengesetzten System $\Gamma_I \times \Gamma_{II}$.

Definition 2.3 (Extensive Größen):

Wir nennen eine thermodynamische Variable M *extensiv* falls sie additive unter Zusammensetzung von Systemen ist, d.h.

$$M_{I+II} = M_I + M_{II}, \quad (2.60)$$

wobei M_{I+II} die entsprechende Größe des zusammengesetzten Systems und M_I und M_{II} die entsprechenden Größen der Einzelsysteme sind.

Beispiele für extensive Variablen sind Entropie S , Teilchenzahl N , Volumen V , innere Energie U , Magnetisierung M und elektrische Ladung Q .

Strikt gesehen, ist die Energie in relevanten Situationen keine extensive Größe, da die Hamiltonfunktion eines zusammengesetzten Systems eine zusätzliche Wechselwirkungsenergie enthält. Diese Wechselwirkungsenergie ist jedoch meist viel kleiner, als Gesamtenergie. Daher betrachten wir die Energie ebenfalls als extensive Größe, was approximativ meist richtig ist.

Definition 2.4 (Intensive Größen):

Wir nennen eine thermodynamische Variable λ *intensiv*, falls sie bei Zusammensetzung zweier Systeme, die untereinander im Gleichgewicht sind, gleich bleibt, d.h.

$$\lambda_{I+II} = \lambda_I = \lambda_{II}, \quad (2.61)$$

wobei λ_{I+II} die entsprechende Größe des zusammengesetzten und λ_I und λ_{II} die entsprechenden Größen der Einzelsysteme sind.

Beispiele für intensive Variablen sind chemisches Potential μ , Druck p , Temperatur T , magnetisches Feld B , und elektrisches Feld. Streng genommen ist die Temperatur für wechselwirkende Teilchen keine intensive Größe. Wenn die Wechselwirkungen ausreichend schnell mit der Entfernung abfallen, verhält sich die Temperatur jedoch auf ausreichend großen Längenskalen intensiv. Eine solche Aussage kann auch für Quantensysteme bewiesen werden [18].

Allgemein gilt: **Wenn die Phasenraumfunktionen $M : \Gamma \rightarrow \mathbb{R}^m$ extensiv sind so sind die zugehörigen Lagrange-Multiplikatoren λ intensiv**, da im Fall (2.60) für die entsprechenden kanonischen Verteilungen

$$\rho(\xi_I, \xi_{II}) = \rho_I(\xi_I) \rho_{II}(\xi_{II}) \quad (2.62)$$

gilt, was aus der Eigenschaft (iv) (Abschnitt 1.4.1) der Entropie und dem Jaynesschen Prinzip folgt (Verständnisaufgabe). Daher wird λ_ν dann auch die *intensive zu $M^{(\nu)}$ konjugierte Größe* genannt.

Außerdem gilt dann für die Entropie

$$S(\alpha M) = \alpha S(M) \quad (2.63)$$

für alle $\alpha \geq 0$ (kleine α können z.B. für die Teilchenzahl offenbar problematisch sein, was aber nicht wichtig ist, da wir uns für „große“ System interessieren). Eine Funktion, die die Eigenschaft (2.63) erfüllt, nennt man auch eine *homogene Funktion ersten Grades*.

Bemerkung für die aufmerksamen Lesenden: Es folgt, dass die Legendretransformation der Entropie über

$$\tilde{S}(k\lambda) = \begin{cases} 0 & \text{falls } \langle \lambda, M \rangle \geq S(M) \\ -\infty & \text{sonst,} \end{cases} \quad \forall M, \quad (2.64)$$

gegeben ist, vgl. (2.16) und Abschnitt 1.5.1.1. Die Rücktransformation ergibt aber trotzdem wieder S . In diesem Fall ergibt $\tilde{S}(k\lambda)$ allerdings nicht die Normierungsfunktion Ψ . Wir werden in der Thermodynamik die Legendre-Transformation nur auf einige Variablen von $S(M)$ anwenden, wo sie wieder mittels stationärer Punkte ausgerechnet werden kann.

Die intensiven Größen bestimmen eindeutig den durch die verallgemeinerte kanonische Verteilung gegebenen Gleichgewichtszustand. Damit definiert „sich im Gleichgewicht befinden“ eine Äquivalenzrelation:

Nullter Hauptsatz der Thermodynamik

Wenn zwei Systeme mit einem dritten System im Gleichgewicht sind, so sind sie auch untereinander im Gleichgewicht.

Aus diesem Hauptsatz kann man schlussfolgern, dass man ein physikalisches System nicht auf den absoluten Nullpunkt kühlen kann.

2.2.5 Der thermodynamische Limes

Wir betrachten zwei extensive Erhaltungsgrößen $M^{(\nu)}$ und $M^{(\nu')}$ mit zugehörigen konjugierten intensiven Variablen λ_ν und $\lambda_{\nu'}$ eines thermodynamischen Systems. Wir vergrößern nun das System um einen Faktor $\alpha > 1$ und bezeichnen die zu $M^{(\nu)}$ und $M^{(\nu')}$ entsprechenden Größen mit $M_\alpha^{(\nu)}$ und $M_\alpha^{(\nu')}$. Wir bemerken, dass die zu $M_\alpha^{(\nu)}$ und $M_\alpha^{(\nu')}$ zugehörigen intensiven nicht von α abhängen und bezeichnen sie daher mit λ_ν und $\lambda_{\nu'}$. Per Definition von Extensivität haben wir

$$M_\alpha^{(\nu)} = \alpha M^{(\nu)} \quad (2.65)$$

und ähnlich für ν' . Damit ist

$$\frac{\partial M_\alpha^{(\nu)}}{\partial \lambda_{\nu'}} = \alpha \frac{\partial M^{(\nu)}}{\partial \lambda_{\nu'}} \quad (2.66)$$

und ähnlich für vertauschte ν und ν' . Die relativen Schwankungen der extensiven Variablen sind damit nach dem Dissipations-Fluktuations-Satz (Satz 2.1) als

$$\frac{\text{Cov}(M_\alpha^{(\nu)}, M_\alpha^{(\nu')})}{M_\alpha^{(\nu)} M_\alpha^{(\nu')}} = -\frac{1}{\alpha} \frac{1}{M^{(\nu)} M^{(\nu')}} \frac{\partial M^{(\nu)}}{\partial \lambda_{\nu'}} \xrightarrow{\alpha \rightarrow \infty} 0 \quad (2.67)$$

gegeben, d.h. **die relativen Schwankungen/Fluktuationen verschwinden im Limes großer Systemgrößen.**

Beispiel: Die relativen Energieschwankungen

$$\frac{\Delta U}{U} = \sqrt{\frac{\text{Var}[H]}{U^2}} = \sqrt{\frac{\text{Cov}(H, H)}{U^2}} \quad (2.68)$$

verschwinden für große Systeme (H ist die Hamiltonfunktion).

Bemerkung: Der zentrale Grenzwertsatz (Satz 1.6) kann im Fall nicht-wechselwirkender Systeme ebenfalls verwendet werden, um zu zeigen, dass $\text{Var}[M_\alpha^{(\nu)}]/M^{(\nu)2}$ wie $1/\alpha$ nach Null geht für $\alpha \rightarrow \infty$.

Daraus folgt insbesondere, dass sobald ein System durch verschiedenen verallgemeinerten kanonischen Verteilungen (etwa kanonisch und großkanonisch) korrekt beschrieben wird diese Verteilungen zu äquivalent sein müssen.

2.3 Die klassischen Ensemble

9. VL

1-Formen (Abschnitt 1.6) wurden ebenfalls in dieser Vorlesung behandelt.

In diesem Abschnitt wählen wir als Zufallsvariablen der verallgemeinerten kanonischen Verteilung (Abschnitt 2.1) $M^{(\nu)}$ extensive physikalische Größen. Damit werden die Lagrange-Multiplikatoren λ_ν intensive physikalische Größen. Diese wollen wir hier identifizieren und physikalische Situationen schildern, wo die speziellen Verteilungen Anwendungen finden.

Die Energie H kommt in jedem der klassischen Ensemble als extensive Größe vor. Damit ist es hilfreich, die verallgemeinerte kanonische Verteilung (Abschnitt 2.1) als

$$\rho(\xi) = \frac{1}{Z} e^{-\beta H(\xi) - \lambda_j M^{(j)}} \quad (\text{ESK}) \quad (2.69)$$

umzuschreiben, wobei wir mit römische Buchstaben (hier j) andeuten, dass nicht über die Energievariablen summiert wird. Die Größe

$$\Phi := -\frac{\ln(Z)}{\beta} = -kT \ln(Z) \quad (2.70)$$

nennt man auch *Potential*; wir werden sehen, dass dem Potential ein vollständiger Satz von thermodynamischen Variablen mittels Ableitungen berechnen lässt. Man nennt $(\beta, \lambda_j/\beta)$ die *natürlichen Variablen* von Ψ , dem Potential Φ und der Zustandssumme Z . Man die VKV mithilfe des Potentials Φ auch als

$$\rho(\xi) = e^{\beta(\Phi - H(\xi) - \lambda_j/\beta M^{(j)})} \quad (2.71)$$

schreiben.

In Abschnitt 2.1.1 haben wir die *innere Energie* als

$$\mathbb{E}[H] = U \quad (2.72)$$

kennengelernt und die *inverse Temperatur* als zugehörige intensive Variable

$$\beta = \lambda_1 = \frac{1}{kT} \quad (2.73)$$

identifiziert. Die Entropie ist dann als

$$S(U, M) = \frac{1}{T} U + k \lambda_j M^{(j)} - k \Psi \quad (2.74)$$

gegeben. Die Gibbssche Fundamentalgleichung (2.21) können wir dann als

$$dU = T dS - \frac{\lambda_j}{\beta} dM^{(j)} \quad (\text{ESK}) \quad (2.75)$$

umschreiben. Wir nennen $M^{(\nu)}$ die *natürlichen Variablen* von S und $(S, M^{(j)})$ die natürlichen Variablen der inneren Energie U .

Häufig schreibt man die Abhängigkeit einer Funktion von β (z.B. Zustandssumme)

auch in eine Abhängigkeit von $T = \frac{1}{k\beta}$ um. Dabei ist es hilfreich zu bemerken, dass

$$\frac{\partial f(\beta)}{\partial \beta} = \frac{\partial f(T)}{\partial T} \frac{\partial T}{\partial \beta} = -\frac{1}{k\beta^2} \frac{\partial f(T)}{\partial T}. \quad (2.76)$$

für eine Funktion f , die von der inversen Temperatur β abhängt.

Die Nebenbedingungen

$$\mathbb{E}[M^{(\nu)}] = M^{(\nu)} \quad (2.77)$$

nennt man auch *Makronebenbedingungen*. Meist hängt die Verteilung noch von zusätzlichen Parametern ab, z.B. die kanonische Verteilung hängt noch implizit von der Teilchenzahl und dem Volumen ab (s. nächster Abschnitt). Bei der Wahl solcher Parameter spricht man auch von *Mikronebenbedingungen*. Häufig werden diese Parameter als zusätzliche Variablen der Zustandssumme etwa in der Form

$$Z = Z(\beta, \lambda; \text{Mikronebenbedingungen}) \quad (2.78)$$

mit aufgeführt. Ähnliches gilt für Ψ und das Potential Φ .

Nun kommen wir zu den klassischen Verteilung der statistischen Mechanik, die in jedem Lehrbuch stehen und auch ausführlich auf Wikipedia erklärt sind.

2.3.1 Das kanonische Ensemble

Das *kanonische Ensemble* (ρ, H) modelliert Situation, wo das System an ein Wärmebad (auch Wärmereservoir genannt) mit der Temperatur T gekoppelt und mit diesem im Gleichgewicht ist. Ein *Wärmebad* ist ein System fester Temperatur, die sich durch die Kopplung an das betrachtete System nicht ändert. Die Temperatur vom Wärmebad ist dabei zwar fest aber beliebig, d.h. sie kann auch sehr kalt sein. Man spricht dann auch manchmal von einem Kältebad.

Die kanonische Verteilung hatten wir bereits in Abschnitt 2.1.1 diskutiert. Im kanonischen Ensemble sind das Volumen V und die Teilchenzahl N feste Parameter. N ist durch den Phasenraum $\Gamma = \mathbb{R}^{6N}$ festgelegt ist und der Potentialterm in der Hamiltonfunktion ist außerhalb von des Volumens unendlich groß. Alternativ kann man den Phasenraum als $\Gamma = R^N \times \mathbb{R}^{3N}$ wählen, wobei $R \subset \mathbb{R}^3$ mit $\text{vol}(R) = V$. Somit ist die einzige verbleibende Makronebenbedingung die der Energie, $\mathbb{E}[H] = U$. Die dazu konjugierte intensive Variable ist die inverse Temperatur

$$\beta = \frac{1}{k} \frac{\partial S}{\partial U}. \quad (2.79)$$

Damit können wir die Zustandssumme Z (und auch Ψ und F unten) als Funktion von (β, V, N) auffassen. Daher Schreiben wir auch $Z = Z(\beta) = Z(\beta; V, N)$.

Die Entropie des kanonischen Ensembles ist

$$S(U) = k(\beta U - \Psi) = \frac{U}{T} + k \ln(Z) \quad (2.80)$$

wobei wir (2.11) verwendet haben. Damit können wir die innere Energie als

$$U(S) = TS - \frac{\Psi(\beta)}{\beta} \quad (2.81)$$

ausdrücken. Außerdem gilt

$$dS = \frac{1}{T} dU \quad \Leftrightarrow \quad T dS = dU. \quad (2.82)$$

Daher ist

$$\frac{\partial U}{\partial S} = T. \quad (2.83)$$

Die (*Helmholtzsche*) *freie Energie* ist als

$$F(T; V, N) := U - TS = -kT \ln Z(T; V, N) \quad (2.84)$$

definiert und kann mit (2.83) als die negative Legendretransformierte (Abschnitt 1.5.1) von U identifiziert werden; gleichzeitig ist sie das Potential der kanonischen Verteilung. Die freie Energie F ist die maximale Arbeit, die sich aus dem System bei konstanter Temperatur und konstantem Volumen extrahieren lässt (durch Größen, die hier implizit als konstant angenommen wurden). Ausschreiben der Legendre-Transformation sagt uns, dass

$$F(T) = \min_S \{U(S) - TS\}. \quad (2.85)$$

Man kann allgemein zeigen, dass die kanonische Verteilung die freie Energie minimiert (unter allen Phasenraumverteilungen). Mit der freien Energie lässt sich die kanonische Verteilung (2.33) auch als

$$\rho(\xi) = e^{\beta(F - H(\xi))} \quad (2.86)$$

schreiben.

2.3.1.1 Beispiel: Zustandssumme eines idealen Gases

Für ein ideales Gas mit N Teilchen der Masse m im Volumen V ist die Zustandssumme

$$Z = \frac{V^N}{N! h^{3N}} \prod_{i=1}^{3N} \int dp_i \exp\left(-\beta \frac{p_i^2}{2m}\right), \quad (2.87)$$

die sich zu

$$Z = \frac{(V/\lambda^3)^N}{N!}, \quad (2.88)$$

mit *thermischer Wellenlänge*

$$\lambda := \frac{h}{\sqrt{2\pi m k T}} \quad (2.89)$$

berechnen lässt (Übungsaufgabe 4.3).

Übungsaufgabe: Zeigen Sie die *kalorische Zustandsgleichung*

$$U = \frac{3}{2} N k T \quad (2.90)$$

und beweisen Sie die *Sackur-Tetrode-Gleichung*

$$\begin{aligned} \frac{S}{kN} &\approx \ln \left[\frac{V}{N \lambda^3} \right] + \frac{5}{2} \\ &= \ln \left[\frac{V}{N} \left(\frac{4\pi m}{3h^2} \frac{U}{N} \right)^{3/2} \right] + \frac{5}{2} \end{aligned} \quad (2.91)$$

unter Verwendung der Stirling-Approximation $\ln(N!) = N \ln(N) - N + O(\ln N)$.

2.3.2 Gibbs-Ensemble

Wir werden sehen, wie Situation modelliert werden, wo das System sowohl an ein Wärmebad als auch an ein eine Druck-kontrollierende (undurchlässiger Kolben) Umgebung gekoppelt und mit dieser im Gleichgewicht ist.

Dafür betrachten wir neben der Energie $M^{(1)} = H$ das Volumen V als Zufallsvariable $M^{(2)}$, so dass $\mathbb{E}[M^{(2)}] = V$. Typischerweise ist diese Zufallsvariable (quasi)-konstant, also scharf verteilt. Die Phänomenologie lehrt uns, dass die Arbeitsdifferential als

$$\delta W = -p dV \quad (2.92)$$

gegeben ist, wobei p der Druck (Kraft pro Fläche) ist. Wir erinnern uns, dass solche 1-Formen entlang von Kurven integriert werden können, s. (1.108), was die Arbeit über das obige Arbeitsdifferential definiert.

Übungsaufgabe: Verdeutlichen Sie diese Gesetzmäßigkeit anhand eines Gases, dessen Druck mithilfe eines Kolbens kontrolliert wird. Argumentieren Sie dabei mit der Strecke, um die der Kolben bewegt wird.

Das motiviert die allgemeine Definition des *Drucks* als

$$p := -\mathbb{E}\left[\frac{\partial H}{\partial V}\right] = -\frac{\partial U}{\partial V} \quad (2.93)$$

Aus der Gibbsschen Fundamentalgleichung (2.75) lesen wir daher

$$p = \frac{\lambda_2}{\beta} \quad (2.94)$$

ab, so dass

$$dU = T dS - p dV. \quad (2.95)$$

Die innere Energie setzt sich hier also als

$$dU = \delta Q + \delta W \quad (2.96)$$

zusammen, wobei $\delta W = -p dV$ die *Arbeit* definiert und

$$\delta Q := T dS \quad (2.97)$$

die *Wärme*. Man beachte, dass Arbeits- und Wärmedifferential i.A. nicht exakt sind, was noch eine wichtige Rolle spielen wird. Allerdings ist $1/T$ ein integrierender Faktor für die Wärme δQ .

Die *Gibbssche Verteilung* ist dann über

$$\rho(\xi) = \frac{1}{Z_G} e^{-\beta(H(\xi)+pV)} \quad (2.98)$$

gegeben. Das Ensemble (ρ, H, V) wird auch als *Druck-Ensemble* oder ebenfalls als *kanonisches Ensemble* bezeichnet.

10. VL

Aus (2.95) oder auch

$$S(U, V) = \frac{1}{T} U + \frac{p}{T} V - k\Psi(\beta, \beta p) \quad (2.99)$$

(vgl. (2.74)) folgt, dass

$$\left(\frac{\partial S}{\partial U}\right)_V = \frac{1}{T}, \quad \left(\frac{\partial S}{\partial V}\right)_U = \frac{p}{T}, \quad (2.100)$$

wobei die Indizes V und U jeweils die Variable bezeichnen, die bei der partiellen Ableitung „konstant gehalten wird“. Außerdem erhalten wir

$$U(S, V) = TS - pV + \frac{\Psi(T, p)}{\beta} \quad (2.101)$$

wobei wir nun Ψ als Funktion von (T, p) auffassen. Mit $T = \left(\frac{\partial U}{\partial S}\right)_V$ und $p = -\left(\frac{\partial U}{\partial V}\right)_S$ erhalten wir die *Gibbs-Energie*

$$G(T, p; N) := U - TS + pV = -kT \ln Z_G(T, p; N) \quad (2.102)$$

als Legendretransformierte von $U(S, V)$. Alternativ kann man $G(T, p)$ auch als Legendretransformierte bzgl. der Variable V der Helmholtz'schen freien Energie $F(T)$ auffassen, welche hier genauso wie zuvor definiert ist.

Die Gibbs-Energie ist das Potential der Gibbs-Verteilung. Sie ist die maximale reversible Arbeit, die sich aus dem System bei konstanter Temperatur und konstantem Druck extrahieren lässt (durch Größen, die hier implizit als konstant angenommen wurden, wie z.B. elektromagnetische Kräfte).

Mit der Gibbs-Energie G lässt sich die Gibbs-Verteilung (2.98) auch als

$$\rho(\xi) = e^{\beta(G - H(\xi) - pV)} \quad (2.103)$$

schreiben.

2.3.2.1 Beispiel: Idealen Gas

Übungsaufgabe 6.3: Zeigen Sie, dass für ein ideales Gas $pV = NkT$ gilt.

2.3.3 Großkanonische Ensemble

Großkanonische Ensemble beschreiben Systeme, welche an ein Wärmebad mit möglichem Teilchenaustausch gekoppelt und mit diesem im thermischen und chemischen Gleichgewicht sind. Dafür kann die großkanonische Verteilung aus Abschnitt 2.1.2 kann auf mehrere Teilchensorten verallgemeinert werden. Die extensiven Größen sind als

$$U = \mathbb{E}[H] \quad (\text{innere Energie}) \quad (2.104)$$

$$N^{(j)} = \mathbb{E}[N^{(j)}] \quad (\text{Teilchenzahl der Sorte } j) \quad (2.105)$$

$$\lambda_j = -\beta\mu_j \quad (\mu_j: \text{zugehöriges chemisches Potential}) \quad (2.106)$$

gegeben. Typischerweise betrachtet man Systeme, deren Hamiltonfunktion die Teilchen auf eine Region des Volumens V einschränkt. Allerdings wird das Volumen als fester Parameter betrachtet und nicht als extensive Variable von der Entropie S . Das liefert

$$\begin{aligned} \rho(\xi) &= \frac{1}{\mathcal{Z}} e^{-\beta(H(\xi) - \mu_j N^{(j)}(\xi))} \\ &= e^{\beta(\Phi - H(\xi) - \mu N)} \end{aligned} \quad (\text{WDF, mit ESK}), \quad (2.107)$$

$$\begin{aligned} \mathcal{Z}(T, \boldsymbol{\mu}; V) &= e^{-\Psi(T, \boldsymbol{\mu}; V)} \\ &= \sum_{\mathbf{n}} \int d\xi_{\mathbf{n}} e^{-\beta(H - \mu_j n^j)} \end{aligned} \quad (\text{Zustandssumme}), \quad (2.108)$$

$$S(U, \mathbf{N}; V) = \frac{U}{T} - \frac{\mu_j}{T} N^{(j)} - k\Psi(T, \boldsymbol{\mu}; V) \quad (\text{Entropie}), \quad (2.109)$$

$$U(S, \mathbf{N}; V) = TS + \mu_j N^{(j)} + \frac{\Psi(T, \boldsymbol{\mu}; V)}{\beta} \quad (\text{innere Energie}), \quad (2.110)$$

$$dU = T dS + \mu_j dN^{(j)} \quad (\text{Gibbs FG zu konst. } V), \quad (2.111)$$

$$\begin{aligned} \Phi(T, \boldsymbol{\mu}; V) &= kT\Psi(T, \boldsymbol{\mu}; V) \\ &= -kT \ln(\mathcal{Z}(T, \boldsymbol{\mu}; V)) \end{aligned} \quad (\text{Großkanonisches Potential}). \quad (2.112)$$

Das Großkanonische Potential ist die Legendre-Transformierte von $U(S, \mathbf{N})$.

Bemerkung: Für nichtwechselwirkende Teilchen ist $(U, \mathbf{N}, V) \mapsto S(U, \mathbf{N}; V)$ eine homogene Funktion ersten Grades. Damit nimmt ihre Legendretransformierte \tilde{S} nur Werte in $\{0, -\infty\}$ an (Abschnitt 1.5.1.1) und ist in der Thermodynamik somit nur von beschränktem Nutzen. Aus diesem Grund bietet es sich an, V als festen Parameter zu betrachten.

2.3.3.1 Beispiel zum chemischen Potential

Wir betrachten Teilchen N in zwei gekoppelten Reservoiren mit chemischen Potentialen $\mu_1 \neq \mu_2$ bei konstantem Volumen V und konstanter innerer Energie U . Wir bezeichnen die Anzahl der Teilchen in den jeweiligen Reservoiren mit $N^{(1)}$ und $N^{(2)}$. Dann ist nach (2.111)

$$0 = dU = T dS + \mu_1 dN^{(1)} + \mu_2 dN^{(2)} \quad (2.113)$$

und mit konstanter Gesamtteilchenzahl $N := N^{(1)} + N^{(2)}$ folgt wegen $dS \geq 0$

$$\begin{aligned} 0 &\geq \mu_1 dN^{(1)} + \mu_2 dN^{(2)} \\ &= \mu_1 dN^{(1)} + \mu_2 (dN - dN^{(1)}) \\ &= -(\mu_2 - \mu_1) dN^{(1)}, \end{aligned} \quad (2.114)$$

d.h.

$$(\mu_2 - \mu_1) dN^{(1)} \geq 0. \quad (2.115)$$

Daraus folgt, dass im Nichtgleichgewicht der Teilchensorten die Teilchenzahl der ersten Teilchensorte $N^{(1)}$ genau dann zunimmt, wenn $\mu_1 < \mu_2$, also sie das kleinere chemische Potential hat. Anders gesagt, die Teilchen gehen vom höheren zum tieferen chemischen Potential.

Übungsaufgabe: Argumentieren Sie mithilfe des kanonischen Ensembles, dass wenn man zwei Wärmebäder mit Temperaturen $T_1 \neq T_2$ in Kontakt bringt, Wärme immer vom wärmeren zum kälteren Bad fließt.

2.3.3.2 Beispiel: Zustandssumme des idealen Gases

Für nur eine Teilchensorte ist die Großkanonische Zustandssumme als

$$\mathcal{Z}(\beta, \mu; V) = \sum_{n=0}^{\infty} Z(\beta; V, n) e^{\beta \mu n} \quad (2.116)$$

schreibbar, wobei $Z(\beta; n)$ die Zustandssumme des idealen Gases im kanonischen Ensemble (2.88) mit n Teilchen (Mikronebenbedingung) ist. Daraus ergibt sich

$$\mathcal{Z} = \exp(zV/\lambda^3) \quad (2.117)$$

mit thermischer Wellenlänge λ aus (2.89) und *Fugazität*

$$z := e^{\beta \mu}. \quad (2.118)$$

Damit kann man

$$U = \frac{3}{2} N k T \quad (2.119)$$

$$N = \frac{V z}{\lambda^3} \quad (2.120)$$

ausrechnen (Übungsaufgabe 5.2). Außerdem folgt

$$\begin{aligned} S(U, \mathbf{N}; V) &= \frac{U}{T} - \frac{\mu}{T} \mathbf{N} - k\Psi \\ &= k\beta U - k\beta\mu \mathbf{N} + \frac{Vz}{\lambda^3}. \end{aligned} \quad (2.121)$$

2.3.4 Das Magnetfeld-Ensemble

Die *Magnetisierung* \mathbf{M} als extensive Variable hat $\lambda_{\mathbf{M}} = -\beta \mathbf{B}$ als dazu konjugierte intensive Variable, wobei \mathbf{B} das Magnetfeld bezeichnet. Eine Änderung der Magnetisierung im B-Feld \mathbf{B} ist mit Arbeit verbunden, was durch das Skalarprodukt $\delta W = \mathbf{B} \cdot d\mathbf{M}$ formalisiert werden kann. Daraus kann man folgern, dass $\lambda_{\mathbf{M}} = -\beta \mathbf{B}$ die konjugierte intensive Variable zur Magnetisierung \mathbf{M} ist.

Die weiteren Details kann man analog zum denen vom Gibb-Ensemble herleiten. Das resultierende Potential ist die *Gibbsche Energie*

$$G(T, \mathbf{B}; V, N) = U - TS - \mathbf{B} \cdot \mathbf{M} = -kT \ln(Z), \quad (2.122)$$

welches die Legendretransformierte von $U(S, \mathbf{M})$ ist.

2.3.5 Das mikrokanonische Ensemble

Der Phasenraum des mikrokanonischen Ensembles ist wie beim kanonischen Ensemble der von N ununterscheidbaren Teilchen in einer Region R vom Volumen V . Allerdings wird hier die Energie als exakt bekannt angenommen. Damit sind die Phasenraumkoordinaten auf die Energieschale

$$\{H(\xi) = U\} := \{\xi \in R^N \times \mathbb{R}^{3N} : H(\xi) = U\} \quad (2.123)$$

eingeschränkt. Da sonst nichts weiter als bekannt angenommen wird liefert das Jaynessche Prinzip die Gleichverteilung auf der Energieschale $\{H(\xi) = U\}$ die *mikrokanonische Verteilung*

$$\begin{aligned} \rho_{\text{MK}}(\xi) &= \frac{1}{D(U)} \delta_{\{H(\xi)=U\}}(\xi) && \text{(Diracmaß)} \\ &= \frac{1}{D(U)} \delta(U - H(\xi)) && \text{(Deltadistributionsschreibweise);} \end{aligned} \quad (2.124)$$

der Normierungsfaktor

$$D(U) = \int d\xi \delta(U - H(\xi)) \quad (2.125)$$

wird auch *Energiedichte* genannt.

Das von der Energieschale $\{H(\xi) = U\}$ eingeschlossene Phasenraumvolumen ist

$$\Omega(U) = \int d\xi \Theta(U - H(\xi)) \quad (2.126)$$

mit Heaviside-Funktion

$$\Theta(x) := \begin{cases} 0 & \text{falls } x < 0 \\ 1 & \text{sonst.} \end{cases} \quad (2.127)$$

Damit ist

$$D(U) = \frac{d\Omega(U)}{dU}. \quad (2.128)$$

Beispiel: Für ein ideales Gas im Volumen V mit Hamiltonfunktion

$$H(\xi) = \sum_{i=1}^{3N} \frac{p_i^2}{2m} + V_{\text{Wand}}(q) \quad (2.129)$$

kann man ausrechnen, dass

$$\Omega(U) = \frac{V^N (2\pi m U)^{3N/2}}{h^{3N} N! (3N/2)!} \quad (\text{ideales Gas}). \quad (2.130)$$

Abschließend sei bemerkt, dass das mikrokanonische Ensemble den „traditionellen“ Ausgangspunkt für die Herleitung des kanonischen Ensembles und des Gibbs-Ensembles. Insbesondere kann die Entropie wie folgt definiert werden.

Thermodynamische Definition der Entropie

Die Entropie eines Systems mit Phasenraumvolumen $\Omega(E)$ ist

$$S = k \ln(\Omega) \quad (2.131)$$

Übungsaufgabe 4.1: Zeigen Sie, dass diese Definition in guter Näherung den gleichen Entropiewert liefert, wie die Shannon-Entropie der diskretisierten mikrokanonischen Verteilung.

11. VL

2.3.6 Wiederholung und Ensembleübersicht

- Wahrscheinlichkeitstheorie: Allgemeine mathematische Theorie, etwa zu WD-Fen und Zufallsvariablen
- Das Jaynessches Prinzip liefert eine allgemeine Methode aus bekannten Erwartungswerten von vorgegebenen Zufallsvariablen eine Wahrscheinlichkeitsverteilung zu schätzen. Das liefert die verallgemeinerte kanonische Verteilung.
- In der statistischen Mechanik spielen die verschiedenen Ensemble (s.u.) eine zentrale Rolle. Sie sind Spezialfälle der verallgemeinerte kanonische Verteilung. Einfache physikalische Überlegungen (wie $\delta W = -p dV$) und die phänomenologische Thermodynamik führen zur physikalischen Interpretation dies Ensemble.

Ensemble ($\rho, M^{(\nu)}$)	Mikro-NB Parameter	Makro-NB $M^{(\nu)}$	Intensive Var. λ_ν	Zustandss. $Z = e^{-\Psi}$	Potential $\Phi = kT\Psi$
Kanonisch	(V, N)	$\mathbb{E}[H] = U$	β	$Z(\beta; V, N)$	$F(T; V, N)$
Gibbs	N	$\mathbb{E}[H] = U,$ $\mathbb{E}[V] = V =: V$	$(\beta, \beta p)$	$Z(\beta, p; N)$	$G(T, p; N)$
Großka- nonisch	V	$\mathbb{E}[H] = U,$ $\mathbb{E}[N] = \mathbf{N}$	$(\beta, -\beta\boldsymbol{\mu})$	$\mathcal{Z}(\beta, \boldsymbol{\mu}; V)$	$\Phi(\beta, \boldsymbol{\mu}; V)$
Magnetfeld	V, N	$\mathbb{E}[H] = U,$ $\mathbb{E}[\mathbf{M}] = \mathbf{M}$	$(\beta, -\beta\mathbf{B})$	$Z(\beta, \mathbf{B}; V, N)$	$G(\beta, \mathbf{B}; V, N)$
Mikroka- nonisch	(U, V, N)	Keine	Keine	Ersatz: $\Omega(U)$	$-kT \ln(\Omega) = -TS(U, V, N)$

3 Quantenstatistik

Wir erinnern uns, dass in der klassischen statischen Mechanik Wahrscheinlichkeiten nur auf Grund von Ignoranz gegenüber mikroskopischen Freiheitsgraden aufgetaucht sind. Es gibt jedoch keine fundamentalen physikalischen Gesetze, die eine deterministische Beschreibung von klassischen Vielteilchensystemen verbieten. Experimente mit Quantensystem hingegen sind aus verschiedenen Gründen intrinsisch probabilistisch. Daher erscheint hier eine vollständig probabilistische Beschreibung besonders plausibel.

Für eine vollständige probabilistische Beschreibung von Quantensystemen ist es notwendig, dass man „klassische Wahrscheinlichkeiten“ und rein durch den Quantencharakter auftauchende Wahrscheinlichkeiten miteinander kombinieren kann. Letztere kommen aus dem Superpositionsprinzip für Quantenzustände, die i.A. nicht scharf gemessen werden können.

Eine wichtige Eigenschaft von klassischen Wahrscheinlichkeiten ist, dass man sie *mischen* kann, d.h. für Wahrscheinlichkeitsmaße $(\mathbb{P}^{(j)})_{j \in [n]}$ und einen Wahrscheinlichkeitsvektor $p \in [0, 1]^n$ ist

$$\sum_{j=1}^n p_j \mathbb{P}^{(j)} \quad (\text{Konvexkombination}) \quad (3.1)$$

wieder ein Wahrscheinlichkeitsmaß. Diese klassische Mischung von Wahrscheinlichkeiten hat auch eine operationelle Bedeutung: Man zieht mit Wahrscheinlichkeit p_j aus der Verteilung $\mathbb{P}^{(j)}$.

Eine analoge klassische Mischung von Quantenzuständen ist auf der Ebene von Zustandsvektoren (d.h. Wellenfunktionen) nicht möglich, da passend normierte Linearkombinationen von Zustandsvektoren Superpositionen von Quantenzuständen ergeben. Daher werden wir Zustandsvektoren (bis auf globale Phasen) in die Menge der Operatoren einbetten, wo dann Mischungsoperationen, welche (3.1) verallgemeinern, neben Superpositionen ebenfalls möglich sind. Dies führt zur Definition von Dichteoperatoren. Wir werden auch sehen, dass sich klassische Wahrscheinlichkeitsverteilungen ebenfalls als Dichteoperatoren schreiben lassen.

3.1 Quantenmechanikgrundlagen

Wiederholen Sie die absoluten Grundlagen der Quantenmechanik:

- Hilbertraum \mathcal{H} . Wichtige Beispiele sind $\mathcal{H} = \mathbb{C}^d$ und

$$\mathcal{H} = \mathcal{L}^2(\mathbb{R}^n) = \left\{ \psi : \mathbb{R}^n \rightarrow \mathbb{C} \mid \|\psi\|^2 := \int |\psi|^2 < \infty \right\}. \quad (3.2)$$

In dem gesamten Kapitel sind \mathcal{H} und \mathcal{H}_i mit $i \in \mathbb{Z}_+$ Hilberträume.

- Zustandsvektoren: $|\psi\rangle \in \mathcal{H}$ mit $\| |\psi\rangle \| = 1$
- Wir bezeichnen den Raum der linearen Operatoren auf \mathcal{H} mit $L(\mathcal{H})$ und mit $\text{rank}(\hat{A})$ den Rang eines Operators $\hat{A} \in L(\mathcal{H})$.
- *Observablen* sind hermitesche/selbstadjungierte lineare Operatoren $\hat{A} : \mathcal{H} \rightarrow \mathcal{H}$ (potentiell mit eingeschränktem Definitionsbereich $\text{dom}(\hat{A}) \subset \mathcal{H}$). Die Menge der Observablen auf \mathcal{H} wird mit $\text{Herm}(\mathcal{H}) := \{ \hat{A} \in L(\mathcal{H}) : \hat{A} = \hat{A}^\dagger \}$ bezeichnet.
- Erwartungswerte (reine Zustände): $\langle \psi | \hat{A} | \psi \rangle$

3.1.1 Dichteoperatoren

Die *Spur* eines Operators $\hat{A} \in L(\mathcal{H})$ ist als

$$\text{Tr}[\hat{A}] := \sum_{j=1}^{\dim(\mathcal{H})} \langle j | \hat{A} | j \rangle \quad (3.3)$$

definiert, wobei $\{|j\rangle\} \subset \mathbb{C}^d$ eine Orthonormalbasis ist. Sie hat folgende Eigenschaften:

- $\text{Tr}[\hat{A}\hat{B}] = \text{Tr}[\hat{B}\hat{A}]$ für alle $\hat{A}, \hat{B} \in L(\mathcal{H})$
- Sie ist unitär invariant, d.h. $\text{Tr}[\hat{U}\hat{A}\hat{U}^\dagger] = \text{Tr}[\hat{A}]$ für alle unitären Operatoren $\hat{U} \in L(\mathcal{H})$, und damit unabhängig von der Basiswahl.
- $\text{Tr}[\hat{A} \otimes \hat{B}] = \text{Tr}[\hat{A}] \text{Tr}[\hat{B}]$ für alle $\hat{A} \in L(\mathcal{H})$ und $\hat{B} \in L(\mathcal{H}')$

12. VL

Wir nennen einen selbstadjungierten Operator $\hat{A} \in \text{Herm}(\mathcal{H})$ *positiv semidefinit*, falls

$$\langle \psi | \hat{A} | \psi \rangle \geq 0 \quad \forall |\psi\rangle \in \mathcal{H}. \quad (3.4)$$

Dafür schreiben wir auch

$$\hat{A} \succeq 0. \quad (3.5)$$

Definition 3.1 (Dichteoperator):

Ein *Dichteoperator* ist ein selbstadjungierter Operator $\hat{\rho}$, der positiv-semidefinit ist und (Spur)-normiert ist, d.h. $\text{Tr}[\hat{\rho}] = 1$. Die Menge der Dichteoperatoren of einem Hilbertraum \mathcal{H} ist

$$\mathcal{S}(\mathcal{H}) := \{\hat{\rho} \in \text{Herm}(\mathcal{H}) : \hat{\rho} \succeq 0, \text{Tr}[\hat{\rho}] = 1\}. \quad (3.6)$$

Eine Matrixdarstellung eines Dichteoperators in $\mathcal{S}(\mathbb{C}^d)$ heißt auch *Dichtematrix*. Einen Dichteoperator nennt man auch *Quantenzustand* oder kurz *Zustand*.

Zentrale Beispiele:

- Sei $|\psi\rangle \in \mathcal{H}$ ein Zustandsvektor. Dann ist

$$\hat{\rho} = |\psi\rangle\langle\psi| \quad (3.7)$$

ein Dichteoperator.

- Sei (p_j) eine Zähldichte und $\{|\psi_j\rangle\} \subset \mathcal{H}$ eine Menge von Zustandsvektoren in einem Hilbertraum \mathcal{H} . Dann nennt man $(p_j, |\psi_j\rangle)_j$ auch ein *Ensemble* von Zustandsvektoren. Das Ensemblemittel

$$\hat{\rho} = \sum_j p_j |\psi_j\rangle\langle\psi_j| \quad (3.8)$$

ist ein Dichteoperator.

Mit (3.7) können durch Zustandsvektoren gegebene Quantenzustände als dargestellt werden. Mit (3.8) mit Orthonormalbasis $\{|\psi_j\rangle\}$ können klassische diskrete Wahrscheinlichkeitsverteilungen als diagonale Dichtematrizen dargestellt werden.

Wir nennen $\hat{\rho} \in \mathcal{S}(\mathcal{H})$ einen *reinen Zustand* falls $\hat{\rho}$ von der Form (3.7) ist, d.h. $\text{rank}(\hat{\rho}) = 1$.

Bemerkung zum Fall $\dim(\mathcal{H}) = \infty$: Dichteoperatoren fallen in die Kategorie der *Spurklasseoperatoren*, welche spezielle *kompakte Operatoren* sind und damit ein reines Punktspektrum haben. So lassen sich WDFen beschreiben, die unter Quantisierung zu Zuständen führen, die nur diskrete Messwerte haben.

Aufgaben:

- Zeigen Sie, dass ein Operator $\hat{\rho} \in L(\mathbb{C}^d)$ genau dann ein Dichteoperator ist, falls $\hat{\rho}$ von der Form (3.8) ist.
- Finden Sie zwei unterschiedliche Ensemble von Zustandsvektoren die über (3.8) den gleichen Zustand $\hat{\rho}$ darstellen.
- Zeigen Sie, dass $\hat{\rho} \in L(\mathbb{C}^d)$ genau dann ein reiner Zustand ist, falls $\text{Tr}[\hat{\rho}^2] = 1$ ist.
- Zeigen Sie, dass die Menge der Dichteoperatoren $\mathcal{S}(\mathcal{H})$ eine konvexe Menge ist.

Sei nun $A \in \text{Herm}(\mathcal{H})$ eine Observable und $\hat{\rho} \in \mathcal{S}(\mathcal{H})$ ein Zustand. Dann ist der *Erwartungswert* von A im Zustand $\hat{\rho}$ als

$$\langle A \rangle_{\hat{\rho}} := \text{Tr}[A\hat{\rho}] \quad (3.9)$$

definiert. Für reine Zustände der Form (3.7) folgt

$$\langle A \rangle_{\hat{\rho}} = \langle \psi | \hat{A} | \psi \rangle \quad (3.10)$$

und für Ensemble der Form (3.8) ist

$$\langle A \rangle_{\hat{\rho}} = \sum_j p_j \langle \psi_j | \hat{A} | \psi_j \rangle, \quad (3.11)$$

was einem klassischen Erwartungswert der Zufallsvariable $j \mapsto \langle \psi_j | \hat{A} | \psi_j \rangle$ entspricht. Dichteoperatoren verallgemeinern also tatsächlich sowohl klassische Wahrscheinlichkeitsverteilungen als auch durch Zustandsvektoren gegebene Quantenzustände.

Aufgabe: Für $\mathcal{H} = \mathbb{C}^2$ betrachten wir den *maximal gemischten Zustand*

$$\hat{\rho}_1 := \frac{1}{2} (|0\rangle\langle 0| + |1\rangle\langle 1|) \quad (3.12)$$

in und den Superpositionszustand

$$\hat{\rho}_2 := \frac{1}{2} (|0\rangle + |1\rangle)(\langle 0| + \langle 1|), \quad (3.13)$$

wobei $(|0\rangle, |1\rangle) \subset \mathbb{C}^2$ die kanonische Basis ist. Finden sie eine Observable, welche $\hat{\rho}_1$ und $\hat{\rho}_2$ unterscheidet.

3.1.2 Zeitentwicklung

Nun wollen wir sehen, dass die Beschreibung von Quantenzuständen durch Dichteoperatoren mit der durch die Schrödingergleichung

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = H |\psi(t)\rangle \quad (3.14)$$

gegebene Zeitentwicklung kompatibel ist. Wie gewöhnlich bezeichnet H hier den Hamiltonoperator. Für $\hat{\rho}(t) = |\psi(t)\rangle\langle\psi(t)|$ (Fall (3.7)) impliziert die Schrödingergleichung die *Von-Neumann-Gleichung*

$$i\hbar \dot{\hat{\rho}} = [\hat{H}, \hat{\rho}] \quad (3.15)$$

(schnelle *Übungsaufgabe*), wobei $[\hat{A}, \hat{B}] := \hat{A}\hat{B} - \hat{B}\hat{A}$ den *Kommutator* definiert. Da diese Differentialgleichung (3.15) linear in $\hat{\rho}$ ist und jeder Quantenzustand $\hat{\rho}$ als Linearkombination (3.8) von reinen Quantenzuständen geschrieben werden kann gilt sie auch für alle Quantenzustände.

Für einen zeitkonstanten Hamiltonoperator H wird die Schrödingergleichung durch

$$|\psi(t)\rangle = \hat{U}(t) |\psi(t=0)\rangle \quad (3.16)$$

mit dem *Zeitentwicklungsoperator*

$$\hat{U}(t) := e^{-i\hat{H}t/\hbar} \quad (3.17)$$

gelöst; die Exponentialfunktion eines Operators ist über den Spektralkalkül definiert (Abschnitt 1.7). Ähnlich wird die von Neumann-Gleichung durch

$$\hat{\rho}(t) = \hat{U}(t)\hat{\rho}(t=0)U(t)^\dagger \quad (3.18)$$

gelöst. Damit lässt sich der Erwartungswert einer Observablen $\hat{A} \in \text{Herm}(\mathcal{H})$ als

$$\begin{aligned}\langle \hat{A} \rangle_{\hat{\rho}(t)} &= \text{Tr}[\hat{U}(t)\hat{\rho}(t=0)\hat{U}(t)^\dagger \hat{A}] \\ &= \text{Tr}[\hat{\rho}(t=0)\hat{A}(t)]\end{aligned}\quad (3.19)$$

mit

$$\hat{A}(t) = \hat{U}(t)^\dagger \hat{A} \hat{U}(t) \quad (\text{Heisenberg-Bild}) \quad (3.20)$$

schreiben.

3.1.3 Zusammengesetzte Systeme

Zusammengesetzte Systeme sind sowohl Vielteilchensysteme als auch einzelne Teilchen mit mehreren Freiheitsgraden, wie etwa Bewegungs-, Rotations-, Schwingungs- und Spinfreiheitsgrade. In beiden Fällen ist jeder Freiheitsgrad i mit einem Hilbertraum \mathcal{H}_i assoziiert und der Hilbertraum aller Freiheitsgrade ist als *Tensorprodukt*

$$\mathcal{H} = \bigotimes_{i=1}^N \mathcal{H}_i \quad (3.21)$$

gegeben. Nun wollen wir kurz die benötigten Details zum Tensorprodukt zusammenfassen. Der Übersicht halber betrachten wir den Fall $N = 2$.

Das *Skalarprodukt* zweier *Produktvektoren* $|\phi_1\rangle \otimes |\phi_2\rangle, |\psi_1\rangle \otimes |\psi_2\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ ist über

$$(\langle \phi_1 | \otimes \langle \phi_2 |) (|\psi_1\rangle \otimes |\psi_2\rangle) := \langle \phi_1 | \psi_1 \rangle \langle \phi_2 | \psi_2 \rangle \quad (3.22)$$

definiert. Das Skalarprodukt zweier beliebiger Vektoren ist entsprechend als dessen sesquilineare Fortsetzung definiert. Diese lässt sich bzgl Basen Orthonormalbasen $\{|i\rangle_1\} \subset \mathcal{H}_1$ und $\{|j\rangle_2\} \subset \mathcal{H}_2$ berechnen. Hier haben wir die *Produktbasis*

$$|i, j\rangle := |i\rangle_1 |j\rangle_2 := |i\rangle_1 \otimes |j\rangle_2, \quad (3.23)$$

welche eine Orthonormalbasis für $\mathcal{H}_1 \otimes \mathcal{H}_2$ ist. Für einen Vektor $|\psi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ ist die Basisentwicklung damit

$$|\psi\rangle = \sum_{i,j} \psi_{i,j} |i, j\rangle \quad (3.24)$$

mit Entwicklungskoeffizienten

$$\psi_{i,j} = \langle i, j | \psi \rangle. \quad (3.25)$$

Das Skalarprodukt zweier Vektoren $|\phi\rangle, |\psi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ mit Entwicklungskoeffizienten $\{\phi_{i,j}\}$ und $\{\psi_{i,j}\}$ ist

$$\langle \phi | \psi \rangle = \sum_{i,j} \phi_{i,j}^* \psi_{i,j}, \quad (3.26)$$

wobei $z^* = \text{Re}(z) - i \text{Im}(z)$ das komplex konjugierte einer Zahl $z \in \mathbb{C}$ bezeichnet.

Ein *Produktoperator* $\hat{A}_1 \otimes \hat{A}_2 \in \text{L}(\mathcal{H}_1 \otimes \mathcal{H}_2) \cong \text{L}(\mathcal{H}_1) \otimes \text{L}(\mathcal{H}_2)$ kann als die lineare Fortsetzung von

$$(\hat{A}_1 \otimes \hat{A}_2) |\psi_1\rangle \otimes |\psi_2\rangle := (\hat{A}_1 |\psi_1\rangle) \otimes (\hat{A}_2 |\psi_2\rangle) \quad (3.27)$$

definiert werden.

Lokale Observablen

Eine Observable auf dem System 1 ist eine Observable der Form

$$\hat{A} = \hat{A}_1 \otimes \mathbb{1}. \quad (3.28)$$

Damit ist die natürliche Erweiterung von Observablen auf größere Systeme gegeben.

Produktzustände

Produktzustände sind von der Form

$$\hat{\rho} = \hat{\rho}_1 \otimes \hat{\rho}_2 . \quad (3.29)$$

Sie beschreiben unabhängige Zustandspräparationen auf zwei Systemen.

Kurze Aufgaben: Zeigen Sie, dass

- $(\hat{A}_1 \otimes \hat{A}_2)(\hat{B}_1 \otimes \hat{B}_2) = (\hat{A}_1 \hat{B}_1) \otimes (\hat{A}_2 \hat{B}_2)$
- $\hat{\rho}_1 \otimes \hat{\rho}_2$ ist ein Dichteoperator, wenn $\hat{\rho}_1$ und $\hat{\rho}_2$ Dichteoperatoren sind.
- Finden Sie einen Dichteoperator in $\mathcal{S}(\mathbb{C}^2 \otimes \mathbb{C}^2)$ der nicht von der Produktform $\hat{\rho}_1 \otimes \hat{\rho}_2$ ist.

13. VL

Der Spektralkalkül (Abschnitt 1.7) wurde ebenfalls in dieser Vorlesung behandelt.

3.1.4 Reduzierte Zustände/Marginale

Nun betrachten wir Zustände in $\mathcal{S}(\mathcal{H}_1 \otimes \mathcal{H}_2)$.

Partielle Spur

Die *partielle Spur* $\text{Tr}_2 : \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_2) \rightarrow \mathcal{L}(\mathcal{H}_1)$ ist die durch

$$\text{Tr}_2[\hat{X}_1 \otimes \hat{X}_2] = \hat{X}_1 \text{Tr}[\hat{X}_2] \quad (3.30)$$

definierte lineare Abbildung. Der *reduzierte Zustand* (oder das *Marginal*) von $\hat{\rho} \in \mathcal{S}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ auf \mathcal{H}_1 ist $\text{Tr}_2[\hat{\rho}] \in \mathcal{S}(\mathcal{H}_1)$.

Die partielle Spur Tr_1 und der reduzierte Zustand auf \mathcal{H}_2 sind analog definiert.

Es gilt

$$\text{Tr}[\hat{\rho}(\hat{A}_1 \otimes \mathbb{1})] = \text{Tr}[\text{Tr}_2[\hat{\rho}]\hat{A}_1] , \quad (3.31)$$

d.h. der reduzierte Zustand $\text{Tr}_2[\hat{\rho}]$ kann alle *lokalen Erwartungswerte* auf System 1 beschreiben.

Kurze Übungsaufgaben:

- Zeigen Sie, dass der reduzierte Zustand tatsächlich ein Dichteoperator ist.
- Beweisen Sie (3.31).
- Sei $\hat{\rho} \in \mathcal{S}(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2})$ ein diagonaler Dichteoperator. Zeigen Sie, dass $\text{Tr}_2[\hat{\rho}]$ ebenfalls ein diagonaler Dichteoperator ist. Weiterhin seien p und $p^{(1)}$ die durch die jeweiligen Diagonalen gegebenen Wahrscheinlichkeitsvektoren. Erklären Sie, warum $p^{(1)}$ als Marginalverteilung von p aufgefasst werden kann (vgl. Abschnitt 1.3.1.6).

3.2 Informationsmaße für Quantenzustände

Die Shannon-Entropie hat eine natürliche Verallgemeinerung auf Quantenzustände, die Von-Neumann-Entropie genannt wird. Entropiemaße spielen eine wichtige Rolle in der Quanteninformationstheorie [19].

Wie zuvor setzen wir $0 \ln(0) := 0$.

Definition 3.2 (Von-Neumann-Entropien):

Die *Von-Neumann-Entropie* eines Dichteoperators $\hat{\rho}$ ist

$$S(\hat{\rho}) := -k \operatorname{Tr}[\hat{\rho} \ln(\hat{\rho})], \quad (3.32)$$

wobei die Einheitenkonstante k dem Kontext entsprechend gewählt wird.

Eigenschaften:

- (i) Die Entropie eines Dichteoperators $\hat{\rho}$ mit Spektralzerlegung $\rho = \sum_i \lambda_i |\lambda_i\rangle\langle\lambda_i|$ ist

$$S(\rho) = -k \sum_i \lambda_i \ln(\lambda_i) = S(\lambda), \quad (3.33)$$

wobei der Vektor der Eigenwerte λ eines Dichteoperators immer eine Zähldichte ist.

- (ii) Für alle $\hat{\rho} \in \mathcal{S}(\mathcal{H})$ gilt

$$0 \leq S(\hat{\rho}) \leq k \ln(\dim(\mathcal{H})), \quad (3.34)$$

was wir unten kurz beweisen. Für reine Zustände $\hat{\rho} = |\psi\rangle\langle\psi|$ gilt

$$S(|\psi\rangle\langle\psi|) = 0 \quad (3.35)$$

und für den *maximal gemischten Zustand* $\hat{\rho} = \frac{1}{\dim(\mathcal{H})}$

$$S(\mathbb{1}/\dim(\mathcal{H})) = k \ln(\dim(\mathcal{H})). \quad (3.36)$$

Diese Schranken werden von den entsprechenden Schranken von der Shannon-Entropie geerbt (Abschnitt 1.4.1).

- (iii) Die Entropie ist *additiv auf Produktzuständen*,

$$S(\hat{\rho} \otimes \hat{\sigma}) = S(\hat{\rho}) + S(\hat{\sigma}) \quad (3.37)$$

für alle $\hat{\rho} \in \mathcal{S}(\mathcal{H}_1)$ und $\hat{\sigma} \in \mathcal{S}(\mathcal{H}_2)$ (kurze Übungsaufgabe) und im Allgemeinen *subadditiv*,

$$S(\hat{\rho}) \leq S(\operatorname{Tr}_2[\hat{\rho}]) + S(\operatorname{Tr}_1[\hat{\rho}]) \quad (3.38)$$

für alle $\hat{\rho} \in \mathcal{S}(\mathcal{H}_1 \otimes \mathcal{H}_2)$, was wir in Lemma 1 unten kurz beweisen.

- (iv) Die Entropie ist *konkav* [19, Theorem 5.23], d.h.

$$S\left(\sum_{i=1}^n p_i \hat{\rho}_i\right) \geq \sum_{i=1}^n p_i S(\hat{\rho}_i) \quad (3.39)$$

für alle $(\hat{\rho}_i) \subset \mathcal{S}(\mathcal{H})$ und Wahrscheinlichkeitsvektoren $p \in [0, 1]^n$.

Den *Kern* eines Operators $\hat{X} \in \mathcal{L}(\mathcal{H})$ bezeichnen wir mit

$$\ker(\hat{X}) := \{|\psi\rangle \in \mathcal{H} : \hat{X} |\psi\rangle = 0\}. \quad (3.40)$$

Ähnlich, wie die Von-Neumann-Entropie die Shannon-Entropie verallgemeinert, verallgemeinert die *relative Entropie* die Kullback-Leibler-Divergenz (Zusatzinfo, Abschnitt 1.4.2).

Definition 3.3 (Relative Entropie):

Die *relative Entropie* von zwei Dichteoperatoren $\hat{\rho}, \hat{\sigma} \in \mathcal{S}(\mathcal{H})$ ist

$$S(\hat{\rho}||\hat{\sigma}) := \begin{cases} k \operatorname{Tr}[\hat{\rho} \ln(\hat{\rho}) - \hat{\rho} \ln(\hat{\sigma})] & \text{falls } \ker(\hat{\sigma}) \subset \ker(\hat{\rho}), \\ \infty & \text{sonst,} \end{cases} \quad (3.41)$$

wobei die Einheitenkonstante k dem Kontext entsprechend gewählt wird.

Eigenschaften:

- Die relative Entropie von Dichteoperatoren $\hat{\rho}$ und $\hat{\sigma}$ mit Spektralzerlegungen $\hat{\rho} = \sum_i \lambda_i |\lambda_i\rangle\langle\lambda_i|$ und $\hat{\sigma} = \sum_j \xi_j |\xi_j\rangle\langle\xi_j|$ ist

$$S(\hat{\rho}||\hat{\sigma}) = k \sum_i \lambda_i \left(\ln(\lambda_i) - \sum_j |\langle\xi_j|\lambda_i\rangle|^2 \ln(\xi_j) \right) \quad (3.42)$$

(kurze Übungsaufgabe).

- Im Allgemeinen ist $S(\hat{\rho}||\hat{\sigma}) \neq S(\hat{\sigma}||\hat{\rho})$.

Mit dem folgenden Lemma kann die relative Entropie als nicht-symmetrische Abstandsfunktion auf $\mathcal{S}(\mathcal{H})$ betrachtet werden.

Lemma 3.4 (Ungleichung von Klein):

Für alle $\hat{\rho}, \hat{\sigma} \in \mathcal{S}(\mathcal{H})$ gilt (mit $k > 0$)

$$S(\hat{\rho}||\hat{\sigma}) \geq 0 \quad (3.43)$$

und

$$S(\hat{\rho}||\hat{\sigma}) = 0 \quad \Leftrightarrow \quad \hat{\rho} = \hat{\sigma}. \quad (3.44)$$

Beweis. Wir verwenden vollständige orthonormierte Eigenbasen $(|\lambda_i\rangle)$ und $(|\xi_j\rangle)$ von $\hat{\rho}$ und $\hat{\sigma}$, d.h. $\sum_i |\lambda_i\rangle\langle\lambda_i| = \sum_j |\xi_j\rangle\langle\xi_j| = \mathbb{1}$. Wir nehmen an, dass $\ker(\hat{\sigma}) \subset \ker(\hat{\rho})$,

da sonst $S(\hat{\rho}||\hat{\sigma}) = \infty > 0$. Dann ist

$$\frac{1}{k} S(\hat{\rho}||\hat{\sigma}) = \sum_i \lambda_i \left(\ln(\lambda_i) - \sum_j |\langle \xi_j | \lambda_i \rangle|^2 \ln(\xi_j) \right) \quad (3.45)$$

$$= \sum_i \lambda_i \left(\left(\sum_j |\langle \xi_j | \lambda_i \rangle|^2 \right) \ln(\lambda_i) - \sum_j |\langle \xi_j | \lambda_i \rangle|^2 \ln(\xi_j) \right) \quad (3.46)$$

$$= \sum_{i,j} \lambda_i |\langle \xi_j | \lambda_i \rangle|^2 (\ln(\lambda_i) - \ln(\xi_j)) \quad (3.47)$$

$$= - \sum_{i,j} \lambda_i |\langle \xi_j | \lambda_i \rangle|^2 \ln \left(\frac{\xi_j}{\lambda_i} \right) \quad (3.48)$$

$$\geq - \sum_{i,j} \lambda_i |\langle \xi_j | \lambda_i \rangle|^2 \left(1 - \frac{\xi_j}{\lambda_i} \right) \quad (3.49)$$

$$= \sum_{i,j} |\langle \xi_j | \lambda_i \rangle|^2 (\xi_j - \lambda_i) \quad (3.50)$$

$$= \sum_j \left(\sum_i |\langle \xi_j | \lambda_i \rangle|^2 \right) \xi_j - \sum_i \left(\sum_j |\langle \xi_j | \lambda_i \rangle|^2 \right) \lambda_i \quad (3.51)$$

$$= 1 - 1 = 0 \quad (3.52)$$

In (3.46), sowie (3.52) haben wir $\sum_i |\langle \xi_j | \lambda_i \rangle|^2 = \sum_j |\langle \xi_j | \lambda_i \rangle|^2 = 1$ (Normierung, vollständige Basen) verwendet, in (3.49), dass die Taylorentwicklung 1. Ordnung (Tangente) von konkaver Funktion obere Schranke liefert. Außerdem wurde die Normierung der Dichteoperatoren ($\sum_i \lambda_i = \sum_j \xi_j = 1$) in (3.52) verwendet.

Die Ungleichung (3.49) ist nun mit Gleichheit erfüllt, wenn für alle i, j gilt dass entweder $|\langle \xi_j | \lambda_i \rangle|^2 = 0$ oder $\ln(\xi_j/\lambda_i) = 0$, d.h. wenn $\hat{\rho} = \hat{\sigma}$. \square

Als einfache Anwendung des Lemmas zeigen wir, dass die obere Schranke in (3.34) gilt. Mit

$$\begin{aligned} 0 &\leq S(\hat{\rho}||\mathbb{1}/\dim(\mathcal{H})) \\ &= -k \operatorname{Tr}[\hat{\rho} \ln(\mathbb{1}/\dim(\mathcal{H}))] - S(\hat{\rho}) \\ &= k \ln(\dim(\mathcal{H})) - S(\hat{\rho}) \end{aligned} \quad (3.53)$$

folgt die obere Schranke (3.34).

Man kann argumentieren, dass die Subadditivität der Von-Neumann-Entropie aus der Subadditivität der Shannon-Entropie folgt oder wieder Klein's Ungleichung verwenden:

Lemma 1 (Subadditivität der Entropie). *Für alle $\hat{\rho} \in \mathcal{S}(\mathcal{H}_1 \otimes \mathcal{H}_2)$*

$$S(\hat{\rho}) \leq S(\operatorname{Tr}_2[\hat{\rho}]) + S(\operatorname{Tr}_1[\hat{\rho}]). \quad (3.54)$$

Beweis. Mit der Wahl

$$\hat{\sigma} = \operatorname{Tr}_2[\hat{\rho}] \otimes \operatorname{Tr}_1[\hat{\rho}] \quad (3.55)$$

ist

$$0 \leq S(\hat{\rho}||\hat{\sigma}) = -k \operatorname{Tr}[\hat{\rho} \ln(\operatorname{Tr}_2[\hat{\rho}] \otimes \operatorname{Tr}_1[\hat{\rho}])] - S(\hat{\rho}). \quad (3.56)$$

Wegen $\ln(\hat{A} \otimes \hat{B}) = \ln(\hat{A}) \otimes \mathbb{1} + \mathbb{1} \otimes \ln(\hat{B})$ liefert dies

$$\begin{aligned} S(\hat{\rho}) &\leq -k \operatorname{Tr}[\hat{\rho} (\ln(\operatorname{Tr}_2[\hat{\rho}]) \otimes \mathbb{1})] - k \operatorname{Tr}[\hat{\rho} (\mathbb{1} \otimes \ln(\operatorname{Tr}_1[\hat{\rho}]))] \\ &= -k \operatorname{Tr}[\operatorname{Tr}_2[\hat{\rho}] \ln(\operatorname{Tr}_2[\hat{\rho}])] - k \operatorname{Tr}[\operatorname{Tr}_1[\hat{\rho}] \ln(\operatorname{Tr}_1[\hat{\rho}])] \\ &= S(\operatorname{Tr}_2[\hat{\rho}]) + S(\operatorname{Tr}_1[\hat{\rho}]), \end{aligned} \quad (3.57)$$

wobei wir (3.31) benutzt haben. \square

14. VL

Für die Herleitung der verallgemeinerten kanonischen Verteilung für den quantenmechanischen Fall benötigen wir die folgende Aussage.

Lemma 3.5 (Pinching-Ungleichung):

Seien $(\hat{P}_j) \subset \text{Herm}(\mathcal{H})$ orthogonale Projektoren mit

$$\sum_j \hat{P}_j = \mathbb{1}. \quad (3.58)$$

Wir definieren die *Pinching-Abbildung* $\mathcal{P} : \text{L}(\mathcal{H}) \rightarrow \text{L}(\mathcal{H})$ über

$$\mathcal{P}(X) := \sum_j \hat{P}_j X \hat{P}_j. \quad (3.59)$$

Dann ist $\mathcal{P}(\hat{\rho}) \in \mathcal{S}(\mathcal{H})$ und

$$S(\hat{\rho}) \leq S(\mathcal{P}(\hat{\rho})) \quad (3.60)$$

für alle $\hat{\rho} \in \mathcal{S}(\mathcal{H})$. Außerdem ist $S(\hat{\rho}) = S(\mathcal{P}(\hat{\rho}))$ genau dann wenn $\hat{\rho} = \mathcal{P}(\hat{\rho})$.

Beweis. Man kann direkt überprüfen, dass $\mathcal{P}(\hat{\rho})$ ein Dichteoperator ist. Weiterhin gilt

$$\begin{aligned} S(\hat{\rho} \|\mathcal{P}(\hat{\rho})) &= k \text{Tr} \left[\hat{\rho} \ln(\hat{\rho}) - \hat{\rho} \ln \left(\sum_j \hat{P}_j \hat{\rho} \hat{P}_j \right) \right] \\ &= -k \text{Tr} \left[\hat{\rho} \sum_k \hat{P}_k \ln \left(\sum_j \hat{P}_j \hat{\rho} \hat{P}_j \right) \hat{P}_k \right] - S(\hat{\rho}) \\ &= -k \text{Tr} [\mathcal{P}(\hat{\rho}) \ln(\mathcal{P}(\hat{\rho}))] - S(\hat{\rho}) \\ &= S(\mathcal{P}(\hat{\rho})) - S(\hat{\rho}) \end{aligned} \quad (3.61)$$

und wegen der Nichtnegativität (3.43) ist

$$S(\hat{\rho}) \leq S(\mathcal{P}(\hat{\rho})). \quad (3.62)$$

Falls $\mathcal{P}(\hat{\rho}) \neq \hat{\rho}$ folgt mit (3.44) auch $S(\mathcal{P}(\hat{\rho})) > S(\hat{\rho})$, was auch die letzte Behauptung beweist. \square

3.3 Analogietabelle-Tabelle: Klassische Mechanik vs. Quantenmechanik

Die Beschreibung der klassischen Mechanik und der Quantenmechanik hat weitgehende Analogien, die wir hier kurz auflisten wollen.

	Klassische Mechanik	Quantenmechanik
Konfigurationsraum	Phasenraum Γ , e.g., $\Gamma = \mathbb{R}^{6N}$	Hilbertraum \mathcal{H} e.g., $\mathcal{H} = \mathcal{L}^2(\mathbb{R}^{3N})$
Koordinaten	Kanonische K. p_1, \dots, q_{3N}	Orthogonalbasis

Zustände	WDFen $\rho \geq 0, \int d\xi \rho(\xi) = 1$	Dichteoperatoren $\mathcal{S}(\mathcal{H})$ $\hat{\rho} \succeq 0, \text{Tr}[\hat{\rho}] = 1$
Observablen	Phasenraumfunktion $M : \Gamma \rightarrow \mathbb{R}$	S.a. Operator $\hat{A} = \hat{A}^\dagger$
Erwartungswerte	Integral $\langle f \rangle_\rho = \int d\xi \rho(\xi) M(\xi)$	Spur $\langle \hat{A} \rangle_{\hat{\rho}} = \text{Tr}[\hat{\rho} \hat{A}]$
Reine Zustände	Diracmaß $\rho = \delta_{\xi_0}$	Zustandsvektor $\hat{\rho} = \psi\rangle\langle\psi $
Ein Teilchen	$\Gamma = \{(q, p)^\top : q, p \in \mathbb{R}^3\}$	$\mathcal{H} = \mathcal{L}^2(\mathbb{R}^3)$
Zusammensetzung	$\Gamma = \Gamma_1 \times \Gamma_2$	$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$
Lokale Observablen	$M(\xi_1, \xi_2) = M_1(\xi_1)$	$\hat{A} = \hat{A}_1 \otimes \mathbb{1}$
Produkt Observablen	$(M_1 \times M_2)(\xi_1, \xi_2) = M_1(\xi_1) M_2(\xi_2)$	$\hat{A}_1 \otimes \hat{A}_2$
Unabhängigkeit	$\rho(\xi_1, \xi_2) = \rho_1(\xi_1) \rho_2(\xi_2)$	$\hat{\rho} = \hat{\rho}_1 \otimes \hat{\rho}_2$
Marginale	$\rho_1(\xi_1) = \int d\xi_2 \rho(\xi_1, \xi_2)$	$\hat{\rho}_1 = \text{Tr}_2[\hat{\rho}]$
Energie	Hamiltonfunktion H	Hamiltonoperator H
Dynamik	$\dot{\rho}_t = \{\rho_t, H\}$ (Poisson-Klammer) $\dot{p}_i = -\frac{\partial H}{\partial q_i}, \dot{q}_i = \frac{\partial H}{\partial p_i}$ $\rho_t(\xi) = \rho(\mathcal{F}_{-t}(\xi))$	$\dot{\hat{\rho}}_t = -\frac{i}{\hbar} [\hat{H}, \hat{\rho}]$ $i\hbar \dot{\psi}(t)\rangle = \hat{H} \psi(t)\rangle$ $U_t = e^{-iHt/\hbar}$
Unabhängige Dynamik	$H(\xi_1, \xi_2) = H_1(\xi_1) + H_2(\xi_2)$	$\hat{H} = \hat{H}_1 \otimes \text{id} + \text{id} \otimes \hat{H}_2$
Entropie	Shannon $S(p) = -k \sum_i p_i \ln(p_i)$ Differenziell $S(\rho) = -k \int d\xi \rho(\xi) \ln \rho(\xi)$	Von Neumann $S(\hat{\rho}) = -k \text{Tr}[\hat{\rho} \ln(\hat{\rho})]$
VKV (ν : ESK)	$\rho(\xi) = \frac{1}{Z} e^{-\lambda_\nu M^{(\nu)}(\xi)}$ (WDF)	$\hat{\rho} = \frac{1}{Z} e^{-\lambda_\nu \hat{M}^{(\nu)}} \in \mathcal{S}(\mathcal{H}),$ $\hat{M}^{(\nu)} \in \text{Herm}(\mathcal{H})$

3.4 Der verallgemeinerte kanonisch Dichteoperator

Das Prinzip der maximalen Entropie (Abschnitt 2.1) können wir nun auf Quantensystem erweitern, was ebenfalls von Jaynes initialisiert wurde [20]. Um ein Quantenvielteilchensystem zu beschreiben, nehmen wir an, dass es durch Observablen $\hat{M}^{(\nu)} \in \text{Herm}(\mathcal{H})$ mit $\nu \in [m]$ beschrieben wird. Um zu einem stationären Gleichgewichtszustand zu gelangen nehmen wir an, dass diese Observablen Erhaltungsgrößen sind, d.h.

$$[\hat{M}^{(\nu)}, \hat{H}] = 0 \quad (3.63)$$

für alle ν , wobei $\hat{H} \in \text{Herm}(\mathcal{H})$ der Hamiltonoperator ist. Weiterhin nehmen wir wieder an, dass deren Erwartungswerte

$$\langle \hat{M}^{(\nu)} \rangle = M^{(\nu)} \quad (3.64)$$

bekannt sind.

Außerdem machen wir eine Annahme, welche die Analyse der Zustände maximaler Entropie massiv erleichtern wird und in vielen wichtigen Situationen erfüllt ist: Wir nehmen an, dass die Observablen auch untereinander kommutieren,

$$[\hat{M}^{(\nu)}, \hat{M}^{(\mu)}] = 0 \quad (3.65)$$

für alle ν und μ . Wir bemerken, dass diese Annahme automatisch erfüllt ist, falls der Hamiltonoperator keine entarteten Eigenwerte hat. Es sei auch bemerkt, dass eine Verallgemeinerung von Zuständen maximaler Entropie zu nichtkommutierenden Erhaltungsgrößen kürzlich formuliert wurde [21].

Da nun alle auftretenden Operatoren kommutieren finden wir eine Orthonormalbasis aus Energieeigenzuständen $\{|j\rangle\} \subset \mathcal{H}$ in der sie alle gleichzeitig diagonal sind. Damit lassen sich die Argumente, die zur verallgemeinerten kanonischen Verteilung geführt haben weitgehend auf den Quantenfall übertragen.

Der *diagonale* Dichteoperator, der die Entropie unter den Nebenbedingungen $\hat{\rho} \in \mathcal{S}(\mathcal{H})$ und (3.64) maximiert, ist von der Form

$$\hat{\rho} = \frac{1}{Z} e^{-\lambda_\nu \hat{M}^{(\nu)}} \quad (\text{ESK}), \quad (3.66)$$

d.h. hat die Diagonaleinträge

$$\hat{\rho}_{j,j} = \frac{1}{Z} \exp\left(-\sum_\nu \lambda_\nu \hat{M}_{j,j}^{(\nu)}\right), \quad (3.67)$$

was aus Abschnitt 2.1 folgt (die Einschränkung, dass $\hat{\rho}$ diagonal sein soll reduziert das Entropie-Maximierungsproblem auf das klassische Entropie-Maximierungsproblem). Diesen Dichteoperator nennt man auch *verallgemeinertes Gibbs-Ensemble*.

Satz 3.6 (Quantenmechanische VKV):

Sei \hat{H} ein Hamiltonoperator auf \mathcal{H} und seien $(\hat{M}^{(\nu)})_{\nu \in [m]} \subset \text{Herm}(\mathcal{H})$ paarweise kommutierende Erhaltungsgrößen. Der Zustand $\hat{\rho} \in \mathcal{S}(\mathcal{H})$, der die Entropie (3.32) unter den Nebenbedingungen

$$\langle \hat{M}^{(\nu)} \rangle = M^{(\nu)} \quad \forall \nu \in [m] \quad (3.68)$$

maximiert ist

$$\hat{\rho} = \frac{1}{Z} e^{-\lambda_\nu \hat{M}^{(\nu)}} \quad (\text{ESK}), \quad (3.69)$$

wobei $Z = \text{Tr}[e^{-\lambda_\nu \hat{M}^{(\nu)}}]$ und $\lambda \in \mathbb{R}^m$ implizit durch (3.68) gegeben ist.

Beweis. Es gilt $\hat{\rho} \in \mathcal{S}(\mathcal{H})$ und $\langle \hat{M}^{(\nu)} \rangle_{\hat{\rho}} = M^{(\nu)}$ für alle ν (kurze Übungsaufgabe). Es bleibt zu zeigen, dass $\hat{\rho}$ die Entropie maximiert. Dafür wählen wir eine gemeinsame orthonormale Eigenbasis $(|j\rangle) \in \text{L}(\mathcal{H})$ von \hat{H} und $(\hat{M}^{(\nu)})$, d.h.,

$$\hat{H} = \sum_j E_j |j\rangle\langle j| \quad \hat{M}^{(\nu)} = \sum_j \mu_j^{(\nu)} |j\rangle\langle j| \quad (3.70)$$

mit Eigenwerten $(E_j), (\mu_j^{(\nu)}) \subset \mathbb{R}$. Wir definieren nun eine Pinching-Abbildung $\mathcal{P} : \text{L}(\mathcal{H}) \rightarrow \text{L}(\mathcal{H})$ über

$$\mathcal{P}(\hat{X}) := \sum_j |j\rangle\langle j| \hat{X} |j\rangle\langle j|. \quad (3.71)$$

Dann ist

$$\mathcal{P}(\hat{\rho}) = \hat{\rho}. \quad (3.72)$$

Mit Lemma 3.5 folgt, dass ein entropiemaximierender Dichteoperator diese Gleichung erfüllen muss. Unter dieser Einschränkung liefert die Entropiemaximierung den verallgemeinerten kanonischen Zustand (3.69), s. Abschnitt 2.1. \square

Der Normierungsfaktor

$$Z(\lambda) = \text{Tr}[e^{-\lambda_\nu \hat{M}^{(\nu)}}] \quad (3.73)$$

heißt auch hier wieder *Zustandssumme*. Die Entropie des verallgemeinerten kanonischen Dichteoperators $\hat{\rho}$ ist wieder

$$S(\mathbf{M}) := S(\hat{\rho}) = k\langle \lambda, \mathbf{M} \rangle - k\Psi(\lambda) \quad (3.74)$$

mit

$$\Psi = -\ln(Z) \quad (3.75)$$

(schnelle Übungsaufgabe) und Lagrange-Parametern λ_ν , die wieder als

$$\frac{\partial S(\mathbf{M})}{\partial \mathbf{M}^{(\nu)}} = k\lambda_\nu \quad (3.76)$$

gegeben sind. Vollständig analog zum klassischen Fall folgt auch wieder

$$-\frac{\partial \ln(Z)}{\partial \lambda_\nu} = \mathbf{M}^{(\nu)}. \quad (3.77)$$

3.5 Dynamische Äquilibration und Thermalisierung (Zusatzinfo)

Wie in Abschnitt 2.2.2 wollen wir motivieren, dass in der Natur häufig Gleichgewichtszustände auftreten, die durch eine verallgemeinerte kanonische Verteilungen gegeben sind. Dazu betrachten wir einen gemäß der Von-Neumann-Gleichung (3.15) zeitentwickelten Zustand $\hat{\rho}(t) \in \mathcal{S}(\mathcal{H})$ und definieren den *zeitgemittelten Zustand* als

$$\hat{\omega} := \overline{\hat{\rho}(t)} := \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \hat{\rho}(t) dt. \quad (3.78)$$

Wir wollen den zeitgemittelten Zustand zunächst genauer charakterisieren. Dazu schreiben wir den Hamiltonoperator bzgl. der Energieeigenprojektoren \hat{P}_j als

$$\hat{H} = \sum_j E_j \hat{P}_j \quad (3.79)$$

mit $E_j \neq E_i$ für $j \neq i$. Damit definieren wir wieder eine Pinching-Abbildung $\mathcal{P} : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H})$ über

$$\mathcal{P}(\hat{X}) := \sum_j \hat{P}_j \hat{X} \hat{P}_j. \quad (3.80)$$

Proposition 3.7 (Zeitmittel):

Der zeitgemittelte Zustand ist

$$\overline{\hat{\rho}(t)} = \mathcal{P}(\hat{\rho}(0)). \quad (3.81)$$

Beweis. Wir schauen uns das Block-Matrixelement (i, j) in der Energieeigenbasis an:

$$\begin{aligned} \hat{P}_i \overline{\hat{\rho}(t)} \hat{P}_j &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \hat{P}_i e^{-i\hat{H}t/\hbar} \hat{\rho}(0) e^{i\hat{H}t/\hbar} \hat{P}_j dt \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \hat{P}_i e^{-i\hat{H}t/\hbar} \hat{P}_i \hat{\rho}(0) \hat{P}_j e^{i\hat{H}t/\hbar} \hat{P}_j dt \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T e^{-i(E_i - E_j)t/\hbar} dt \hat{P}_i \hat{\rho}(0) \hat{P}_j \\ &= \delta_{i,j} \hat{P}_i \hat{\rho}(0) \hat{P}_j, \end{aligned} \quad (3.82)$$

d.h. der zeitgemittelte Zustand $\overline{\hat{\rho}(t)}$ ist block-diagonal in der Energieeigenbasis. Daraus folgt die Behauptung. \square

Wir bemerken, dass falls eine Dynamik in dem Sinne äquilibriert, dass Erwartungswerte von relevanten Observablen äquilibrieren, dann muss der Gleichgewichtszustand die gleichen Erwartungswerte haben wie der zeitgemittelte Zustand $\overline{\hat{\rho}(t)}$. Falls die Observablen zu einem System Σ gehören, liegt die folgende Definition nahe.

Definition 3.8 (Äquilibration im Zeitmittel):

Sei $\hat{H} \in \text{Herm}(\mathcal{H}_\Sigma \otimes \mathcal{H}_\mathcal{E})$ ein Hamiltonoperator, der auf einem System Σ und Umgebung \mathcal{E} wirkt. Wir nennen einen zeitentwickelten Zustand $\text{Tr}_\mathcal{E}[\hat{\rho}(t)]$ auf Σ (im Zeitmittel) äquilibriert, falls $\hat{\rho}(t)$ für die meisten t auf Σ nicht unterscheidbar vom zeitgemittelten Zustand $\hat{\rho}(t)$ ist, d.h.,

$$\text{Tr}_\mathcal{E}[\hat{\rho}(t)] \approx \text{Tr}_\mathcal{E}[\mathcal{P}(\hat{\rho}(0))] . \quad (3.83)$$

Beispiele für nicht äquilibrierende Systeme

- Seien $|E\rangle, |E'\rangle \in \mathcal{H}_\Sigma \otimes \mathcal{H}_\mathcal{E}$ Energieeigenzustände mit Energien $E < E'$ und $|\psi(0)\rangle := (|E\rangle + |E'\rangle)/\sqrt{2}$. Dann ist

$$|\psi(t)\rangle = e^{-iEt/\hbar} \left(|E\rangle + e^{-i(E'-E)t/\hbar} |E'\rangle \right) / \sqrt{2} \quad (3.84)$$

und daher

$$|\psi(t)\rangle\langle\psi(t)| = \frac{1}{2} (|E\rangle\langle E| + |E'\rangle\langle E'| + e^{i(E'-E)t/\hbar} |E\rangle\langle E'| + e^{-i(E'-E)t/\hbar} |E'\rangle\langle E|) . \quad (3.85)$$

Daraus folgt, dass bestimmte Matrixeinträge des Systemzustands $\text{Tr}_\mathcal{E}[|\psi(t)\rangle\langle\psi(t)|]$ mit der Frequenz $(E' - E)/\hbar$ oszillieren und somit keine Äquilibration stattfindet.

- Sei $\hat{H} = \hat{H}_1 \otimes \mathbb{1} + \mathbb{1} \otimes \hat{H}_2$ ein Hamiltonoperator eines nichtwechselwirkenden Systems. Dann kann ein Nichtgleichgewichtszustand auf einem System nicht äquilibrieren.

Vollständig wechselwirkende System mit einem Anfangszustand $\hat{\rho}(0)$, der einen signifikanten Überlapp mit vielen Energieeigenzuständen hat (große *effektive Dimension* von $\hat{\rho}(0)$) äquilibrieren mathematisch beweisbar in folgendem Sinne [22]. Wenn $\dim(\mathcal{H}_\mathcal{E})$ groß ist so ist das Zeitmittel des Abstands von $\text{Tr}_\mathcal{E}[\hat{\rho}(t)]$ und dem zeitgemittelten Zustand $\text{Tr}_\mathcal{E}[\hat{\omega}]$ ist verschwindend klein, d.h.

$$\|\text{Tr}_\mathcal{E}[\hat{\rho}(t)] - \text{Tr}_\mathcal{E}[\hat{\omega}]\| \rightarrow 0 \quad (3.86)$$

für $\dim(\mathcal{H}_\mathcal{E}) \rightarrow \infty$.

Wir bemerken, dass für jede Observable $\hat{A} \in \text{Herm}(\mathcal{H})$ die Äquivalenz

$$[\hat{H}, \hat{A}] = 0 \quad \Leftrightarrow \quad \mathcal{P}(\hat{A}) = \hat{A} \quad (3.87)$$

gilt, d.h. \hat{A} ist genau dann eine Erhaltungsgröße wenn $\mathcal{P}(\hat{A}) = \hat{A}$.

Die Zustände, welche die gleichen Erhaltungsgrößen-Erwartungswerte wie der Anfangszustand haben, sind

$$\{\sigma \in \mathcal{S}(\mathcal{H}) : \text{Tr}[\hat{\sigma}\hat{A}] = \text{Tr}[\hat{\rho}(0)\hat{A}] \quad \forall \hat{A} = \mathcal{P}(\hat{A})\} . \quad (3.88)$$

Unter diesen Zuständen ist der zeitgemittelte Zustand $\hat{\sigma} = \overline{\hat{\rho}(t)}$ derjenige ist, welcher die Entropie maximiert (Lemma 3.5).

Kurz: Entropiemaximierung gegeben alle Erhaltungsgrößen liefert den gleichen Zustand wie Zeitmittelung, also den Gleichgewichtszustand falls er denn existiert.

Man bemerke, dass (3.87) exponentiell mindesten $\dim(\mathcal{H})$ viele Erhaltungsgrößen charakterisiert. Das sind exponentiell viele in der Teilchenzahl! Als nächstes betrachten wir Situationen, in denen ein Gleichgewichtszustand $\hat{\omega}$ durch wenige relevante makroskopische Erwartungswerte (3.64) bereits eindeutig bestimmt ist.

Definition 3.9 (Thermalisierung / Äquilibration ins TD GG):

Wir sagen, dass ein System zu einer Verallgemeinerten kanonischen Verteilung (3.69) *thermalisiert* (d.h. in ein thermodynamisches Gleichgewicht äquilibriert) falls es (i) äquilibriert und (ii) der Gleichgewichtszustand $\hat{\omega}$ mit der Verteilung (3.69) übereinstimmt.

Es gibt verschiedene Modelle und Annahmen, die von Äquilibration zu Thermalisierung führen. Eine Übersicht über verschiedene Ansätze kann in [15, Kapitel 7] gefunden werden.

3.6 Ununterscheidbare Teilchen

Wir betrachten ein Quantensystem aus N ununterscheidbaren Teilchen. Den 1-Teilchen Hilbertraum bezeichnen wir mit \mathcal{H} . Dann ist der N -Teilchen Hilbertraum $\mathcal{H}^{\otimes N}$. Zunächst betrachten wir $N = 2$. Mit $\mathbb{F} \in U(\mathcal{H} \otimes \mathcal{H})$ bezeichnen wir den Flip-Operator, der über

$$\mathbb{F} |\psi\rangle |\phi\rangle := |\phi\rangle |\psi\rangle \quad (3.89)$$

definiert ist. Sei nun $|\psi\rangle \in \mathcal{H}_2$ ein Zustandsvektor von 2 ununterscheidbaren Teilchen. Ununterscheidbarkeit bedeutet, dass

$$\mathbb{F} |\psi\rangle = e^{i\varphi} |\psi\rangle \quad (3.90)$$

für eine von der Teilchenart abhängigen Phase φ . Wegen

$$\mathbb{F}^2 = \mathbb{1} \quad (3.91)$$

sind die $+1$ und -1 die einzigen Eigenwerte von \mathbb{F} und daher

$$e^{i\varphi} = 1 \quad \text{oder} \quad e^{i\varphi} = -1. \quad (3.92)$$

Im Fall $e^{i\varphi} = 1$ nennt man $|\psi\rangle$ *symmetrisch* und im Fall $e^{i\varphi} = -1$ *antisymmetrisch*. Diese Überlegung lässt sich auf N Teilchen verallgemeinern. Mit S_N bezeichnen wir die symmetrische Gruppe, d.h. die Gruppe der Permutationen auf N Elementen. Für eine Permutation $\sigma \in S_N$ bezeichnen wir mit $\hat{U}_\sigma \in U(\mathcal{H}^{\otimes N})$ die unitäre Darstellung mit

$$\hat{U}_\sigma |\psi_1\rangle \dots |\psi_N\rangle := |\psi_{\sigma(1)}\rangle \dots |\psi_{\sigma(N)}\rangle. \quad (3.93)$$

Falls $|\psi\rangle \in \mathcal{H}_N$ ununterscheidbare Teilchen beschreibt ist

$$\hat{U}_\sigma |\psi\rangle = |\psi\rangle \quad (3.94)$$

oder

$$\hat{U}_\sigma |\psi\rangle = \text{sign}(\sigma) |\psi\rangle \quad (3.95)$$

für alle $\sigma \in S_N$, wobei

$$\text{sign}(\sigma) := \begin{cases} +1 & \text{falls } \sigma \text{ ein Produkt einer geraden Anzahl Transpositionen ist,} \\ -1 & \text{falls } \sigma \text{ ein Produkt einer ungeraden Anzahl Transpositionen ist,} \end{cases} \quad (3.96)$$

(eine Transposition ist eine Vertauschung von nur 2 der N Elemente). Wenn ein Vektor $|\psi\rangle \in \mathcal{H}_N$ (3.94) erfüllt heißt er *symmetrisch* und falls er (3.95) erfüllt *antisymmetrisch*.

Definition 3.10 (Bosonen und Fermionen):

Ununterscheidbare Teilchen, die durch symmetrische Zustandsvektoren beschrieben werden nennt man *Bosonen* und ununterscheidbare Teilchen, die durch antisymmetrische Zustandsvektoren beschrieben werden *Fermionen*.

15. VL

Diese Vorlesung wurde von Lennart Bittel gehalten.

In der relativistischen Quantenfeldtheorie wird die folgende wichtige physikalische Aussage bewiesen, welche die diskutierten Symmetrieeigenschaften mit dem Spin der Teilchen verbindet.

Satz 3.11 (Spin-Statistik-Satz):

Zustandsvektoren, die identische Teilchen mit *ganzzahligem* Spin beschreiben, sind *symmetrisch* unter Vertauschung der Teilchen (Bosonen).

Zustandsvektoren, die identische Teilchen mit *halbzahligem* Spin beschreiben, sind *antisymmetrisch* unter Vertauschung der Teilchen (Fermionen).

Wir bemerken, dass die Mengen von symmetrischen und antisymmetrischen Vektoren zwei Untervektorräume von $\mathcal{H}^{\otimes N}$ bilden. Diese Untervektorräume bezeichnen wir mit \mathcal{H}_N^+ und \mathcal{H}_N^- . Die Orthogonalprojektoren auf den symmetrischen und den antisymmetrischen Unterraum bezeichnen wir jeweils mit $\hat{P}_N^+, \hat{P}_N^- \in \text{Herm}(\mathcal{H}^{\otimes N})$. Die Projektoren sind gegeben durch

$$\hat{P}_N^+ |\psi_1\rangle \dots |\psi_N\rangle = c_\psi^+ \sum_{\sigma \in S_N} |\psi_{\sigma(1)}\rangle \dots |\psi_{\sigma(N)}\rangle \quad (3.97)$$

und

$$\hat{P}_N^- |\psi_1\rangle \dots |\psi_N\rangle = c_\psi^- \sum_{\sigma \in S_N} \text{sign}(\sigma) |\psi_{\sigma(1)}\rangle \dots |\psi_{\sigma(N)}\rangle, \quad (3.98)$$

wobei $c_\psi^\pm > 0$ Normierungsfaktoren sind, so dass $\hat{P}_{N,\pm}^2 = \hat{P}_{N,\pm}$ (Projektoreigenschaft). Falls $|\psi_j\rangle = |\psi_{j'}\rangle$ für ein Paar $j \neq j'$ folgt interessanterweise dass

$$\hat{P}_N^- |\psi_1\rangle \dots |\psi_N\rangle = 0, \quad (3.99)$$

d.h. $|\psi_1\rangle \dots |\psi_N\rangle$ kann keinen antisymmetrischen Quantenzustand definieren. Diese Einsicht ergibt das folgende Resultat.

Pauli-Prinzip

Zwei identische Fermionen können sich nicht im gleichen Einteilchenzustand befinden.

Um das großkanonische Ensemble auf Fermionen und Bosonen erweitern zu können nehmen wir noch die folgende Konstruktion vor. Mit der direkten Summe über ver-

schiedene Teilchenzahlen

$$\begin{aligned}\mathcal{F}_+ &:= \overline{\bigoplus_{N=0}^{\infty} \mathcal{H}_N^+} \\ \mathcal{F}_- &:= \overline{\bigoplus_{N=0}^{\infty} \mathcal{H}_N^-}\end{aligned}\tag{3.100}$$

bilden wir die Hilberträume für variable Teilchenzahl, die *Fock-Räume* genannt werden; der Strich bedeutet, dass die Räume bzgl. des induzierten Skalarprodukts topologisch abgeschlossen werden, was als mathematisches Detail erwähnt sein soll. Die Fockräume \mathcal{F}_{\pm} sind die Hilberträume für eine variable Anzahl von Bosonen/Fermionen mit Einteilchenhilbertraum \mathcal{H} . Wir bemerken, dass der Hilbertraum \mathcal{H}_0^{\pm} eindimensional ist; der entsprechende Zustand $|\emptyset\rangle \in \mathcal{F}_{\pm}$ heie *Vakuumzustand*.

Nun whlen wir eine Orthonormalbasis $\{|k\rangle\} \subset \mathcal{H}$. Daraus bekommen wir eine Basis für die Fock-Räume \mathcal{F}_{\pm} , indem wir angeben, wie viele Teilchen jeweils im Zustand $|k\rangle$ sind. Wir schreiben

$$|N_1, N_2, \dots\rangle \in \mathcal{F}_{\pm} \quad \text{in Besetzungszahldarstellung} \tag{3.101}$$

für den Zustand mit jeweils N_j Teilchen im Einteilchenzustand $|k\rangle$. Für $\sum_j N_k = N$ ist $|N_1, N_2, \dots\rangle$ als $P_N^{\pm} |k_1\rangle^{\otimes N_1} |k_2\rangle^{\otimes N_2} \dots$ gegeben. Es gilt

$$\begin{aligned}N_k &\in \{0, 1, 2, \dots\} && \text{für Bosonen,} \\ N_k &= \{0, 1\} && \text{für Fermionen,}\end{aligned}\tag{3.102}$$

für alle k , wobei die Einschränkung der Teilchenzahl für Fermionen aus dem Pauli-Prinzip folgt. In dieser Basis werden die *Teilchenzahloperatoren* \hat{N}_j als

$$\hat{N}_j |N_1, N_2, \dots\rangle = N_j |N_1, N_2, \dots\rangle \tag{3.103}$$

und

$$\hat{N} := \sum_j \hat{N}_j \tag{3.104}$$

definiert.

Kurz: Eine Basiswahl für den Einteilchenhilbertraum \mathcal{H} legt also die Besetzungszahldarstellung und die Teilchenzahloperatoren auf \mathcal{F}_{\pm} fest.

3.7 Ideale Quantengase I

Wir betrachten ein Quantenvielteilchensystem aus nicht-wechselwirkenden ununterscheidbaren Teilchen. Dann ist der Hamiltonoperator \hat{H} durch den Einteilchenhamiltonoperator \hat{H}_1 mit Eigenenergien $\{E_j\}$ festgelegt. Als Basis des Einteilchenhilbertraums whlen wir die Eigenbasis $|E_j\rangle$ von \hat{H}_1 .

3.7.1 Großkanonisches Ensemble

Als Beispiele für die VKV im Quantenfall betrachten wir das Großkanonische Ensemble für nicht-wechselwirkende Fermion und Bosonen. Das liefert die Fermi- und Bose-Verteilung für die Teilchenzahl.

Analog zum klassischen großkanonischen Ensemble whlen wir die Makronebenbedingungen

$$\begin{aligned}\langle \hat{H} \rangle &= U, \\ \langle \hat{N} \rangle &= N,\end{aligned}\tag{3.105}$$

was zum großkanonischen Dichteoperator

$$\hat{\rho} = \frac{1}{\mathcal{Z}} e^{-\beta(\hat{H} - \mu \hat{N})} \quad (3.106)$$

führt (Abschnitt 3.4). Die Wahrscheinlichkeit, bei einer Teilchenzahrmessung die Teilchenzahlen $\mathbf{N} = (N_1, N_2, \dots)$ zu messen ist

$$\mathbb{P}[\mathbf{N}] = \langle \mathbf{N} | \hat{\rho} | \mathbf{N} \rangle = \frac{1}{\mathcal{Z}} \exp \left(-\beta \sum_j (E_j - \mu) N_j \right). \quad (3.107)$$

Die großkanonische Zustandssumme ist

$$\mathcal{Z} = \sum_{\mathbf{N}} \exp \left(-\beta \sum_j (E_j - \mu) N_j \right). \quad (3.108)$$

Aufgabe: Zeigen Sie, dass

$$\mathcal{Z} = \prod_j (1 + t_j) \quad \text{für Fermionen,} \quad (3.109)$$

$$\mathcal{Z} = \prod_j \frac{1}{1 - t_j} \quad \text{für Bosonen} \quad (3.110)$$

mit

$$t_j := e^{-\beta(E_j - \mu)}, \quad (3.111)$$

wobei für Bosonen $E_j > \mu$ angenommen werden muss (warum?).

Hier können wir die *Einteilchenzustandssumme*

$$\mathcal{Z}_j = \sum_{N_j} e^{-\beta(E_j - \mu)N_j} \quad (3.112)$$

$$= \begin{cases} 1 + t_j & \text{(Fermionen)} \\ \frac{1}{1 - t_j} & \text{(Bosonen)} \end{cases} \quad (3.113)$$

identifizieren, so dass $\mathcal{Z} = \prod_j \mathcal{Z}_j$.

Die *mittlere Besetzungszahl* im 1-Teilchenzustand $|E_j\rangle$ ist mit (3.112) als

$$\begin{aligned} \langle \hat{N}_j \rangle &= \frac{1}{\beta} \frac{\partial}{\partial \mu} \ln(\mathcal{Z}_j) \\ &= \begin{cases} \frac{t_j}{1 + t_j} = \frac{1}{1/t_j + 1} & \text{(Fermionen),} \\ \frac{t_j}{1 - t_j} = \frac{1}{1/t_j - 1} & \text{(Bosonen)} \end{cases} \end{aligned} \quad (3.114)$$

gegeben. Mit (3.111) erhalten wir die Verteilungen

$$\langle \hat{N}_j \rangle = \frac{1}{\exp\left(\frac{E_j - \mu}{kT}\right) - \sigma}, \quad (3.115)$$

mit

$$\sigma = \begin{cases} -1 & \text{(Fermi-Dirac, Fermionen),} \\ 0 & \text{(Boltzmann, klassisch),} \\ 1 & \text{(Bose-Einstein, Bosonen).} \end{cases} \quad (3.116)$$

Aufgabe: Skizzieren Sie $\langle \hat{N}_j \rangle$ in Abhängigkeit von E_j für verschiedene Temperaturen.

Ähnlich erhalten wir für die innere Energie (Gleichung (2.20) mit $\lambda_N = -\beta\mu$)

$$U = - \left(\frac{\partial \ln(\mathcal{Z})}{\partial \beta} \right)_{\beta\mu} = \sum_k E_j N_j. \quad (3.117)$$

3.7.1.1 Verbindung zur klassischen statistischen Mechanik

Wir nehmen an, dass die nicht-wechselwirkenden spinlose Teilchen in einem Würfel mit Kantenlänge L eingesperrt und jeweils durch den Einteilchenhamiltonoperator $\hat{H} = \frac{\hat{p}^2}{2m}$ mit 1-Teilchen Impulsoperator $\hat{\mathbf{p}} = (\hat{p}_x, \hat{p}_y, \hat{p}_z)^\top$ beschrieben sind. In der Quantenmechanik wird gezeigt, dass die Impulseigenwerte durch

$$\mathbf{p} = \frac{2\pi\hbar}{L} \boldsymbol{\nu} = \frac{h}{L} \boldsymbol{\nu} \quad \text{mit} \quad \boldsymbol{\nu} \in \mathbb{Z}^3 \quad (3.118)$$

gegeben sind; wir erinnern uns dass die entsprechenden Eigenfunktion von $\hat{\mathbf{p}}$ für periodische Randbedingungen durch $\psi_{\mathbf{p}}(\mathbf{x}) = L^{-3/2} e^{i\mathbf{p} \cdot \mathbf{x} / \hbar}$ gegeben sind. Damit kann z.B. die Einteilchenzustandssumme als

$$\sum_{\mathbf{p}} e^{-\beta(E_{\mathbf{p}} - \mu)} \approx \frac{L^3}{h^3} \int d^3p e^{-\beta(E_{\mathbf{p}} - \mu)} \quad (3.119)$$

approximiert werden, wenn L groß und β klein ist. Ein ähnliches, wenn auch leicht technischeres Argument kann für absorbierende Randbedingungen, wo ψ auf dem Rand des Würfels verschwindet, gemacht werden.

Das Argument kann auch auf die N -Teilchen Zustandssumme erweitert werden, wobei bei der Integration der Faktor $1/N!$ wegen der Ununterscheidbarkeit der Teilchen berücksichtigt werden muss. Zusammen mit

$$V = L^3 = \int d^3x V_{\text{Wand}}(\mathbf{x}) \quad (3.120)$$

und Approximation (3.119) begründet dies die Definition des klassischen Phasenraumvolumenelements (2.32).

Die Energieeigenwerte nicht-wechselwirkender Teilchen sind für die gegebenen Impulseigenwerte (3.118) als

$$E_{\mathbf{p}} = \frac{\mathbf{p}^2}{2m} \quad (3.121)$$

gegeben.

Aufgabe: Zeigen Sie, dass Einsetzen von $E_{\mathbf{p}}$ in (3.108) im klassischen Limes wieder die klassische Zustandssumme ergibt.

3.7.1.2 Spin

(Wird in einer noch kommenden Vorlesung behandelt.)

Die Theorie quantenmechanischer Drehimpulse impliziert, dass ein Teilchen mit Gesamtspin $s \in \mathbb{Z}/2$ hat $2s + 1$ Spin-Freiheitsgrade mit zugehörigen Zustandsvektoren

$$\{|m_s\rangle\}_{m_s=-s, -s+1, \dots, s} \subset \mathcal{H}_s, \quad (3.122)$$

die eine Orthonormalbasis des Spin-Hilbertraums $\mathcal{H}_{\text{Spin}}$ bilden, d.h. $\mathcal{H}_{\text{Spin}} \cong \mathbb{C}^{2s+1}$. Einbeziehung weiterer Freiheitsgrade (z.B. Impuls) liefert den Einteilchenhilbertraum

$$\mathcal{H}_1 = \mathcal{H}_{\text{Spin}} \otimes \mathcal{H}_{\text{Impuls}}. \quad (3.123)$$

Nun nehmen wir an, dass ein Spin- s -Teilchen durch einen Hamiltonoperator der Form

$$H = \mathbb{1}_{\text{Spin}} \otimes H_{\text{Impuls}} \quad (3.124)$$

gegeben ist mit $H \in \text{Herm}(\mathcal{H}_{\text{Impuls}})$, d.h. die Energie des Teilchens ist unabhängig von seinem Spin, was in vielen Situationen eine gute Näherung ist. Nun betrachten wir 2 Teilchen mit Einteilchenhilbertraum \mathcal{H}_1 . Falls s ungerade ist handelt es sich um Fermionen. Damit sind sie durch antisymmetrische Zustandvektoren $\psi \in \mathcal{H}_2^-$ beschrieben, d.h.

$$\psi((m_{s1}, \mathbf{p}_1), (m_{s2}, \mathbf{p}_2)) = -\psi((m_{s2}, \mathbf{p}_2), (m_{s1}, \mathbf{p}_1)). \quad (3.125)$$

Eine verallgemeinerte Version davon gilt natürlich auch für N Fermionen und eine Analoge Überlegung kann man für Teilchen mit ganzzahligem Spin, also Bosonen anstellen.

In jedem Fall führen wegen (3.124) die $2s + 1$ Spin-Freiheitsgrade zu einer (zusätzlichen) Entartung $g_s = 2s + 1$ eines jeden Impuls- und damit auch Energie-Eigenzustands.

4 Thermodynamik

16. VL

Abschnitt 1.6.2 mit dem Satz von Euler wurde ebenfalls in dieser Vorlesung behandelt.

Eine zentrales ursprüngliches Ziel der Thermodynamik ist die Beschreibung von Dampfmaschinen und wie man optimal Wärme in mechanische Arbeit umwandeln kann. Heutzutage stehen die allgemeineren Kreisprozessen im Mittelpunkt, wo auch chemische Reaktionen eine Rolle spielen können. Auch spezifische Eigenschaften von Stoffen werde analysiert und klassifiziert, wie etwa mithilfe von Suzeptibilitäten. Diese können helfen, z.B. verschiedene Aggregatzustände besser zu verstehen.

Die Thermodynamik kann auf Grundlage ihrer Hauptsätze als in sich weitgehend abgeschlossene Theorie aufgebaut werden. Hier werden wir jedoch ausgehend von der statistischen Mechanik die wichtigsten Konzepte der Thermodynamik einführen. Dabei gehen wir wieder von Systemen aus, die durch Gleichgewichtszustände beschrieben sind. Die makroskopischen Eigenschaften dieser Zustände können wiederum durch wenige Zustandsvariablen beschrieben werden. Die sind typischerweise die innere Energie U , das Volumen V und die Teilchenzahl N (bzw. \mathbf{N} bei mehrkomponentigen Systemen). Wie in der Literatur üblich unterscheiden wir nun die Zufallsvariable N und ihren Erwartungswert nicht mehr, d.h. $\mathbb{E}[\mathbf{N}] = \mathbf{N}$ und ähnlich für das Volumen V . Wir nehmen an, dass die betrachteten Systeme so groß sind, dass sämtliche Fluktuationen vernachlässigt werden können (die Ensemble aus Abschnitt 2.3 sind dann äquivalent). Weiterhin betrachten wir *quasi-statische* Situationen, d.h. Messungen und Änderungen an den thermodynamischen Systemen dauern sehr lange im Vergleich zu der Zeit, die das jeweilige System braucht, um ins Gleichgewicht zu äquilibrieren. So können auch zeitliche Änderungen von thermodynamischen Variablen betrachtet werden. Häufig wird eine Variable konstant gehalten. Entsprechende häufig auftretende Prozesse sind:

Prozess	konstante Größe
Isobar	Druck p
Isochor	Volumen V
Isothermisch	Temperatur T
Isentropisch	Entropie S
Adiabatisch	Wärme $\delta Q = 0$

Die einfachsten Kreisprozesse setzen sich aus ebensolchen Prozessen zusammen. Zunächst werden wir einen geschlossenen Ausdruck für die innere Energie $U(S, V, \mathbf{N})$ finden, der die entsprechenden Relationen aus den verschiedenen Ensembles aus Abschnitt 2.3 zusammenführt. Anschließend werden wir die Liste der thermodynamischen Potentiale (s. Abschnitt 2.3) vervollständigen. Damit lassen sich die thermodynamischen Variablen auf verschiedenste Weise als Ableitungen schreiben. Dies führt zu zusätzlichen Relationen diesen Ableitungen (Maxwell-Relationen) und zu den systemspezifischen *Suszeptibilitäten* (z.B. Wärmekapazität). Letztere führen mithilfe des Dissipations-Fluktuations-Satzes 2.1 zu Stabilitätskriterien, die im Gleichgewicht automatisch erfüllt sind und Einschränkungen an die Suszeptibilitäten liefern.

Darauf aufbauend werden wir verschiedene reversible und irreversible Prozesse kennenlernen (z.B. Carnot, Gay-Lussac Versuch) und diese zusammen mit den Hauptsätzen der Thermodynamik diskutieren.

Wenn nicht explizit anders spezifiziert gehen wir in diesem Kapitel von Systemen aus, die durch die extensiven thermodynamischen Variablen S , U , V , und N und den zugehörigen intensiven Variablen T , p und μ beschrieben sind.

4.1 Euler-Gleichung und Gibbs-Duhem-Relation

Im Gibbs-Ensemble ist $dU = TdS - pdV$ das Differential der inneren Energie $U(S, V; N)$ aus (2.100); im großkanonischen Ensemble ist hingegen $dU = TdS + \mu dN$ das Differential von $U(S, N; V)$ aus (2.110). Die jeweiligen Rollen von V und N als Makro- und Mikronebenbedingungen sind also jeweils vertauscht. Die Äquivalenz der Ensembles impliziert, dass allgemeinert

$$dU = TdS - pdV + \mu dN \quad (4.1)$$

gelten muss. Allgemeiner können wir aus der Extensivität der Größen S , U , V und N die folgende Gleichung für U herleiten.

Euler-Gleichung

Die innere Energie ist

$$U(S, V, \mathbf{N}) = TS - pV + \mu_j N^{(j)} \quad (\text{ESK}). \quad (4.2)$$

Beweis. Die innere Energie ist eine homogene Funktion ersten Grades, d.h.

$$U(\alpha S, \alpha V, \alpha \mathbf{N}) = \alpha U(S, V, \mathbf{N}). \quad (4.3)$$

Der Satz von Euler (Satz 1.13), die Fundamentalgleichung (2.111) und $\frac{\partial U}{\partial V} = -p$ liefern daher die behauptete Gleichung. \square

Die Euler-Gleichung impliziert

$$U - TS + pV - \mu N = 0 \quad (\text{ESK}), \quad (4.4)$$

d.h. die intensiven Größen $T = T(S, V, N)$, $p = p(S, V, N)$ und $\mu = \mu(S, V, N)$ sind nicht unabhängig sondern aus zwei von ihnen kann man den Wert der dritten erhalten.

Die Euler-Gleichung ergibt das Differential

$$dU = T dS + S dT - p dV - V dp + \mu_j dN^{(j)} + N^{(j)} d\mu_j \quad (\text{ESK}), \quad (4.5)$$

was mit (4.1) die folgende Gleichung liefert.

Gibbs-Duhem-Relation

$$S dT - V dp + N^{(j)} d\mu_j = 0 \quad (\text{ESK}) \quad (4.6)$$

Wir erinnern uns an die Definition (2.112) des großkanonischen Potentials

$$\Phi = kT\Psi = U - TS - \mu_j N^{(j)}, \quad (4.7)$$

wobei wir (2.110) verwendet haben. Die Fundamentalbeziehung (4.1) und die Euler-Gleichung (4.2) liefern die Fundamentalbeziehung für Φ :

$$\begin{aligned} d\Phi &= d(U - TS - \mu_j N^{(j)}) \quad (\text{ESK}) \\ &= TdS - p dV + \mu_j dN^{(j)} - TdS - SdT - \mu_j dN^{(j)} + N^{(j)} d\mu_j \\ &= -SdT - p dV - N^{(j)} d\mu_j. \end{aligned} \quad (4.8)$$

Außerdem liefert die Euler-Gleichung

$$\Phi(T, \boldsymbol{\mu}; V) = -p(T, \boldsymbol{\mu}) V; \quad (4.9)$$

wobei wir noch zeigen müssen, dass $p(T, \boldsymbol{\mu})$ nicht von V abhängt. Die Fundamentalbeziehung (4.8) impliziert

$$\left(\frac{\partial \Phi}{\partial V} \right)_{T, \boldsymbol{\mu}} = -p. \quad (4.10)$$

Andererseits ist mit der Produktregel

$$\left(\frac{\partial \Phi}{\partial V} \right)_{T, \boldsymbol{\mu}} = - \left(\frac{\partial \Phi}{\partial V} \right)_{T, \boldsymbol{\mu}} V - p, \quad (4.11)$$

was

$$\left(\frac{\partial \Phi}{\partial V} \right)_{T, \boldsymbol{\mu}} = 0 \quad (4.12)$$

impliziert.

Ähnlich erhalten wir mit der Euler-Gleichung für die Gibbs-Energie (2.102)

$$G(T, p; \mathbf{N}) = \mu_j(T, p, \mathbf{N}) N^{(j)} \quad (\text{ESK}), \quad (4.13)$$

wobei für eine Teilchensorte

$$\left(\frac{\partial \mu}{\partial N} \right)_{T, p} = 0 \quad (4.14)$$

ist.

Die anderen thermodynamischen Potentiale mit ihren Fundamentalgleichungen können zueinander analog behandelt werden, wie wir es hier beispielhaft für das großkanonische Potential gesehen haben. Wir werden sie in Abschnitt 4.2 vervollständigt auflisten.

17. VL

4.1.1 Beispiel: Ideales Gas

Wir erinnern uns, dass für das großkanonische Ensemble

$$\frac{\partial \ln(\mathcal{Z})}{\partial \mu} = \frac{1}{\mathcal{Z}} \sum_N \int d\xi_N \beta N e^{-\beta(H(\xi) - \mu N)} = \beta \mathbb{E}[N] = \beta N \quad (4.15)$$

gilt. Mit dem großkanonischen Potential $\Phi = -kT \ln(\mathcal{Z})$ aus (2.112) ergibt das

$$\frac{\partial \Phi(T, \mu; V)}{\partial \mu} = -N. \quad (4.16)$$

Die großkanonische Zustandssumme eines idealen Gases wurde in der Übung als

$$\mathcal{Z} = \exp(zV/\lambda^3), \quad z = e^{\beta\mu}, \quad \lambda = \frac{h}{\sqrt{2\pi mkT}}, \quad (4.17)$$

berechnet, vgl. (2.117). Daraus folgt, dass

$$\frac{\partial \Phi}{\partial \mu} = -kT \frac{V}{\lambda^3} \frac{\partial z}{\partial \mu} = \beta \Phi, \quad (4.18)$$

d.h.

$$\Phi = -NkT. \quad (4.19)$$

Mit (4.9) folgt die *thermische Zustandsgleichung*

$$pV = NkT. \quad (4.20)$$

4.1.2 Erster Hauptsatz der Thermodynamik

Die Fundamentalgleichung

$$dU = TdS - p dV + \mu_j dN^{(j)} \quad (\text{ESK}). \quad (4.21)$$

aus (4.1) ist der *erster Hauptsatz der Thermodynamik* für quasi-statische Prozesse. Die Energieänderung setzt sich aus den folgenden drei Teilen zusammen:

- Die Wärmeänderung

$$\delta Q = TdS \quad (4.22)$$

ist eine nichtexakte 1-Form (Abschnitt 1.6); der zweite Hauptsatz besagt, dass für isolierte Systeme $dS \geq 0$ ist.

- Die mechanische Arbeit

$$\delta W = -p dV \quad (4.23)$$

entspricht dem dem System zugeführten („systemegoistische“ Zählung) durch mechanische Arbeit bedingte Energieanteil und ist ebenfalls nicht exakt.

- Die Energieerhöhung durch Teilchenzufuhr ist durch μdN gegeben. Das chemische Potential μ ist die Energieerhöhung pro Teilchen bei konstanter Entropie und konstantem Volumen.
- Erweiterungen mit anderen Energieformen wie elektrischer und magnetischer Arbeit sowie Verformungsarbeit sind ebenfalls möglich, werden hier aber nicht betrachtet.

4.2 Thermodynamische Potentiale

In Abschnitt 2.3 haben wir das kanonische Ensemble, das Gibbs-Ensemble und das großkanonische Ensemble mit den jeweiligen Zustandssummen Z (kanonisch), Z_G (Gibbs) und \mathcal{Z} (großkanonisch) kennengelernt. Außerdem haben wir schon einige thermodynamischen Potentiale kennengelernt. Dem ist noch hinzuzufügen, dass die innere Energie U und die Entropie S ebenfalls als *Potentiale* bezeichnet werden, da sie in bestimmten Situationen die zur Verfügung stehende Energie quantifizieren. Da die Ensemble in der Thermodynamik äquivalent sind, werden wir Makro- und Mikroneben-

bedingungen nicht mehr strikt trennen und führen Variablen meist in gleichbleibender Reihenfolge auf.

	Potential/ Variablen	Zusammenhang zu anderen Variablen	Fundamental- gleichung	Relationen bzgl. Zustandssumme
Innere Energie	$U(S, V, N)$	$U = TS - pV + \mu N$	$dU = TdS - pdV + \mu dN$	$U = TS - kT \ln(Z)$
Freie Energie	$F(T, V, N)$	$F = U - TS$	$dF = -SdT - pdV + \mu dN$	$F = -kT \ln(Z)$
Gibbs-Energie	$G(T, p, N)$	$G = U - TS + pV$	$dG = -SdT + Vdp + \mu dN$	$G = -kT \ln(Z_G) = \mu N$
Großkanon. Pot.	$\Phi(T, V, \mu)$	$\Phi = U - TS - \mu N$	$d\Phi = -SdT - pdV - Nd\mu$	$\Phi = -kT \ln(Z) = -pV$
Enthalpie (neu)	$H(S, p, N)$	$H := U + pV$	$dH = TdS + Vdp + \mu dN$	

Bemerkung:

- Z : kanonisch Zustandssumme, Z_G : Zustandssumme aus dem Gibbs-Ensemble, \mathcal{Z} : Großkanonische Zustandssumme
- Mit der Gleichung $U = TS - pV + \mu N$ ist der funktionelle Zusammenhang

$$U(S, V, N) = T(S, V, N)S - p(S, V, N)V + \mu(S, V, N)N \quad (4.24)$$

gemeint. Analoge Aussagen gelten für die anderen Spalten und Zeilen der Tabelle.

- Die Fundamentalbeziehungen für F , G , Φ und H folgen jeweils mit der Euler-Gleichung (4.2), vgl. (4.8).
- Aus den Fundamentalgleichungen lassen sich die Ableitungen der Potentiale ablesen; z.B. folgt aus $dU = TdS + \dots$ mit (1.3) dass $\frac{\partial U}{\partial S} = T$ ist.
- Die Enthalpie ist das einzige wirklich neu eingeführte Potential. Bei isobaren Prozessen (p konstant) ohne Teilchenaustausch (N konstant) ist

$$dH = TdS = \delta Q, \quad (4.25)$$

also die vom System aufgenommene Wärme.

- Die Tabelle gilt auch für den mehrkomponentigen Fall, wenn man μN jeweils mit $\sum_j \mu_j N^{(j)}$ ersetzt.

Verständnisaufgaben:

- Listen Sie einige Ableitungen der Potentiale nach ihren natürlichen Variablen auf.
- Stellen Sie die innere Energie U über (i) eine Ableitung der der Freien Energie F und (ii) als Legendre-Transformierte von F dar. Wodurch unterscheiden die sich so berechneten Funktionen U ?
- Stellen sie ähnliche Überlegen wie in (4.25) für einige der anderen Potentiale an.

Satz 4.1 (Konvexität der inneren Energie):

Sei $(\frac{\partial S}{\partial U})_{V,N} > 0$ und S eine konkave und homogene Funktion von (U, V, N) . Dann ist $(S, V, N) \mapsto U(S, V, N)$ ist eine konvexe Funktion.

Beweis. Die Konkavität von S bedeutet, dass für alle (U_1, V_1, N_1) und (U_2, V_2, N_2) und $\alpha \in [0, 1]$

$$\begin{aligned} S(\alpha U_1 + (1 - \alpha)U_2, \alpha V_1 + (1 - \alpha)V_2, \alpha N_1 + (1 - \alpha)N_2) \\ \geq \alpha S(U_1, V_1, N_1) + (1 - \alpha)S(U_2, V_2, N_2) \end{aligned} \quad (4.26)$$

ist. Wir setzen nun $S_i := S(U_i, V_i, N_i)$ für $i = 1, 2$ und für jedes $X \in \{S, V, N\}$ setzen wir $X = \alpha X_1 + (1 - \alpha)X_2$.

Die Funktion $U(S, V, N)$ erhält man aus der Entropie S indem man $S(U, V, N)$ nach U auflöst (das ist nach dem Satz über implizite Funktionen lokal immer möglich wenn $\frac{\partial S}{\partial U} \neq 0$). Wegen der Homogenität von S ist dann

$$\begin{aligned} S(U(S, V, N), V, N) &= \alpha S(U(S_1, V_1, N_1), V_1, N_1) \\ &\quad + (1 - \alpha)S(U(S_2, V_2, N_2), V_2, N_2) \\ &\leq S(\alpha U(S_1, V_1, N_1) + (1 - \alpha)U(S_2, V_2, N_2), V, N) \end{aligned} \quad (4.27)$$

Wegen $\frac{\partial S}{\partial U} > 0$ und obiger Relation kann $S(U, V, N)$ daher als

$$U(S, V, N) \leq \alpha U(S_1, V_1, N_1) + (1 - \alpha)U(S_2, V_2, N_2) \quad (4.28)$$

(lokal) invertiert werden. \square

Bemerkungen:

- Da die Entropie S im Gleichgewicht maximal unter den durch U, V, N gegebenen Nebenbedingungen ist und hier als die konkave differentielle bzw. Von-Neumann-Entropie erklärt war, ist sie laut Satz 1.11 auch eine konkave Funktion.
- Für Systeme von nicht-wechselwirkenden Teilchen ist die Entropie homogen.
- $(\frac{\partial S}{\partial U})_{V, N} > 0$ ist eine physikalisch motivierte Bedingung: Bei höheren Energien sind mehr Zustände zugänglich, womit die Entropie auch mit der Energie zunimmt.

Im abstrakten allgemeinen Fall (2.74) fassen wir $S(U, M)$ und $U(S, M)$ als Potentiale auf. Die anderen Potentiale erhält man, indem man bzgl. einige der Variablen S und $M^{(1)}, \dots, M^{(m)}$ Legendre-transformiert; hier mit der konvexen Version der Legendretransformation, was aber für die konkrete Berechnung keinen Unterschied zur konkaven Version macht, s. Abschnitt 1.5.1. Die entsprechenden neuen Variablen sind dann die entsprechenden Variablen aus $\{T, \lambda_1, \dots, \lambda_m\}$.

Auf diese Weise hatten wir in Abschnitt 2.3 die Potentiale F , G und Φ dargestellt. Aus dU in der obigen Liste können wir

$$\left(\frac{\partial U}{\partial V}\right)_{V, N} = -p \quad (4.29)$$

ablesen (andere Ableitungen der Potentiale können auf ähnliche Weise abgelesen werden). Dies impliziert, dass die Enthalpie die Legendre-Transformierte der inneren Energie $U(S, V, N)$ bzgl. der Variable V ist. Streng genommen ist die Enthalpie damit eine Funktion von $(S, -p; N)$, doch das Minuszeichen wird meistens in der Notation unterdrückt; eine solche Vorzeichenunterdrückung werden wir in der nachfolgenden Liste mit einem $+$ kennzeichnen.

18. VL

(10.12.19)

Zusammengefasst erhalten wir die folgenden Zusammenhänge über die Legendre-Transformation.

Funktion f	Neu Variable(n)	LT von f
$U(S, V, N)$	$\frac{\partial U}{\partial S} = T$	$-F(T, V, N)$
$U(S, V, N)$	$\frac{\partial U}{\partial V} = -p$	$-H(S, +p, N)$
$U(S, V, N)$	$\frac{\partial U}{\partial S} = T, \frac{\partial U}{\partial V} = -p$	$-G(T, +p, N)$
$U(S, V, N)$	$\frac{\partial U}{\partial S} = T, \frac{\partial U}{\partial N} = \mu$	$-\Phi(T, V, \mu)$
$-F(T, V, N)$	$-\frac{\partial F}{\partial T} = S$	$U(S, V, N)$
$F(S, V, N)$	$-\frac{\partial F}{\partial V} = p$	$-G(T, +p, N)$
$-G(T, p, N)$	$-\frac{\partial G}{\partial T} = S$	$-H(S, p, N)$

Bemerkungen:

- Hier haben wir ausschließlich die Definition der Legendre-Transformation für konvexe Funktionen verwendet, s. Abschnitt 1.5.1 und vgl. [4, Kapitel 5.2]. In der konkreten Berechnung macht dies jedoch keinen Unterschied zur konkaven Version, da in beiden Fällen der Kritische Punkt von (1.91) gleich sind.
- Damit sind die thermodynamischen Potentiale konvex in ihren extensiven Variablen und konkav in den intensiven Variablen.
- Man kann in den hier betrachteten Spezialfällen verschiedene Variablen nacheinander transformieren und so das gleiche Ergebnis bekommen, wie wenn man die mehrdimensionale Legendre-Transformation ausführt; z.B. ist kann G als Legendre-Transformierte von U bzgl. (S, V) oder von F bzgl. V dargestellt werden.
- Auch die Rücktransformationen sind entsprechend möglich; z.B. ist U die Legendre-Transformierte von F bzgl. S . Man beachte, dass die oben angegebene Liste nicht vollständig ist.

Kurze Übungsaufgabe: Fügen Sie der Liste entsprechende Zeilen hinzu.

4.2.1 Maxwell-Relationen

Die Gibbs-Duhem-Relation liefert eine Beziehung zwischen den 1. Ableitungen der Fundamentalbeziehungen. Die Maxwellrelationen hingegen liefern Beziehungen zwischen den 2. Ableitungen. Aus der Tatsache, dass zweite partielle Ableitungen von glatten Funktionen vertauschen (Satz von Schwarz) erhalten wir aus den 2. Ableitungen der Potentiale Relationen für die Ableitungen der jeweiligen thermodynamischen Variablen, z.B.:

$$-\left(\frac{\partial p}{\partial S}\right)_{V,N} = \left(\frac{\partial T}{\partial V}\right)_{S,N} = \left(\frac{\partial^2 U}{\partial S \partial V}\right)_N, \quad (4.30)$$

$$\left(\frac{\partial \mu}{\partial S}\right)_{V,N} = \left(\frac{\partial T}{\partial N}\right)_{S,V} = \left(\frac{\partial^2 U}{\partial S \partial N}\right)_V, \quad (4.31)$$

$$\left(\frac{\partial \mu}{\partial V}\right)_{S,N} = -\left(\frac{\partial p}{\partial N}\right)_{S,V} = \left(\frac{\partial^2 U}{\partial V \partial N}\right)_S, \quad (4.32)$$

$$\left(\frac{\partial p}{\partial T}\right)_{V,N} = \left(\frac{\partial S}{\partial V}\right)_{T,N} = -\left(\frac{\partial^2 F}{\partial T \partial V}\right)_N, \quad (4.33)$$

$$-\left(\frac{\partial S}{\partial p}\right)_{T,N} = \left(\frac{\partial V}{\partial T}\right)_{p,N} = \left(\frac{\partial^2 G}{\partial T \partial p}\right)_N. \quad (4.34)$$

Solche Relationen werden *Maxwell-Relationen* genannt. Sie sind die *Integrabilitätsbedingungen* für die Existenz der Potentiale. Umgekehrt stellen sie wichtige Zusammenhänge zwischen verschiedenen thermodynamischen Variablen her (hier zwischen Paaren von konjugierten Paaren von Variablen).

Das *Guggenheim-Quadrat* ist ein populäres Merkschema für die Maxwell-Relationen, s. [Wikipedia](#).

Im nächsten Abschnitt sehen wir, dass einige der Maxwell-Relationen eine direkte Aussage über die sogenannte Suszeptibilitäten machen.

Übungsaufgabe: Leiten Sie einige weitere Maxwell-Relationen her.

4.2.1.1 Hilfreiche Identitäten aus der Analysis

Der *Satz von der Umkehrabbildung* aus der Analysis besagt, dass falls die Ableitung $d_a f \in L(\mathbb{R}^n)$ einer glatten Funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ an einem Punkt $a \in \mathbb{R}^n$ invertierbar ist, die Abbildung f auch lokal invertierbar ist und

$$(d_a f)^{-1} = d_{f(a)}(f^{(-1)}) \quad (4.35)$$

gilt.

Wir betrachten nun reelle Variablen x, y, z , die in dem Sinne abhängig sind, dass $z = z(x, y)$ als Funktion von x und y aufgefasst werden kann. Aus (4.35) können wir die folgende Relation folgern.

Lemma 4.2:

Sei $z(x, y)$ eine glatte Funktion mit $(\partial z / \partial x)_y \neq 0$. Dann gilt lokal, dass

$$\left(\frac{\partial z}{\partial x} \right)_y = \frac{1}{\left(\frac{\partial x}{\partial z} \right)_y}. \quad (4.36)$$

Alternativer „Beweis“ von (4.36). Wir setzen y konstant und nehmen an, dass $z(\cdot, y)$ glatt und invertierbar ist. Dann ist

$$dx = \frac{\partial x}{\partial z} dz = \frac{\partial x}{\partial z} \frac{\partial z}{\partial x} dx \quad (4.37)$$

und damit

$$\frac{\partial x}{\partial z} \frac{\partial z}{\partial x} = 1. \quad (4.38)$$

□

Der *Satz über die implizite Funktionen* besagt, dass unter der Bedingung, dass die Ableitung einer Funktion $f(\cdot, \mathbf{y})$ für einen Punkt \mathbf{y} invertierbar ist, man für einen Wert \mathbf{z} eine eindeutige Funktion $\mathbf{y} = g_{\mathbf{z}}(\mathbf{x})$ finden kann, so dass lokal $f(\mathbf{x}, g_{\mathbf{z}}(\mathbf{x})) = \mathbf{z}$ ist und

$$\frac{\partial g_{\mathbf{z}}}{\partial x_j}(\mathbf{x}) = - \left(D_{g_{\mathbf{z}}(\mathbf{x})} f(\mathbf{x}, \cdot) \right)^{-1} \nabla_{\mathbf{x}} f(\cdot, g_{\mathbf{z}}(\mathbf{x})) \quad (4.39)$$

gilt, wobei die Ableitungen hier jeweils auf die durch Punkte gekennzeichnete Funktionsargumente wirken, s. [Wikipedia \(englische Version\)](#). Mit eindimensionalen Variablen x, y, z und den Wahlen $f(x, y) = z(x, y)$ und $g_z(x) = y(x, z)$ folgt, dass

$$\left(\frac{\partial y}{\partial x} \right)_z = - \frac{\left(\frac{\partial z}{\partial x} \right)_y}{\left(\frac{\partial z}{\partial y} \right)_x} \quad (4.40)$$

gilt, was wir mithilfe von (4.36) in der folgenden, sich leichter zu merkenden Form schreiben können.

Lemma 4.3:

Lokal gilt

$$\left(\frac{\partial x}{\partial z}\right)_y \left(\frac{\partial y}{\partial x}\right)_z \left(\frac{\partial z}{\partial y}\right)_x = -1 \quad (4.41)$$

unter der Bedingung, dass zwei der drei partiellen Ableitungen nicht verschwinden (die dritte verschwindet dann automatisch nicht).

Alternativer Beweis von (4.41). Es gilt

$$\begin{aligned} dy &= \left(\frac{\partial y}{\partial x}\right)_z dx + \left(\frac{\partial y}{\partial z}\right)_x dz \\ &= \left(\frac{\partial y}{\partial x}\right)_z \left(\left(\frac{\partial x}{\partial y}\right)_z dy + \left(\frac{\partial x}{\partial z}\right)_y dz \right) + \left(\frac{\partial y}{\partial z}\right)_x dz \\ &= dy + \left(\frac{\partial y}{\partial x}\right)_z \left(\frac{\partial x}{\partial z}\right)_y dz + \left(\frac{\partial y}{\partial z}\right)_x dz \end{aligned} \quad (4.42)$$

und der Vergleich der Koeffizientenfunktionen von dz ergibt

$$\left(\frac{\partial y}{\partial x}\right)_z \left(\frac{\partial x}{\partial z}\right)_y = - \left(\frac{\partial y}{\partial z}\right)_x, \quad (4.43)$$

wobei wir die Annahme verwendet haben, dass $\left(\frac{\partial x}{\partial y}\right)_z \neq 0$. Die Identität (4.36) liefert (4.41). \square

4.2.1.2 Weitere Maxwell-Relationen

Mit (4.36) lassen sich die Maxwell-Relationen umschreiben, z.B. folgt aus (4.30)

$$- \left(\frac{\partial S}{\partial p}\right)_{V,N} = - \frac{1}{\left(\frac{\partial p}{\partial S}\right)_{V,N}} = \frac{1}{\left(\frac{\partial T}{\partial V}\right)_{S,N}} = \left(\frac{\partial V}{\partial T}\right)_{S,N}, \quad (4.44)$$

was bis zu drei weitere Maxwell-Relationen ergibt. Ähnlich kann man (4.41) verwenden, um weitere Maxwell-Relationen aufzustellen.

4.2.2 Suszeptibilitäten und thermodynamische Stabilität

Eine *Suszeptibilität* ist ein Maß für die Änderung einer extensiven Größe bei Veränderung einer intensiven Größe.

In diesem Abschnitt halten wir die Teilchenzahl konstant, d.h.

$$dN = 0 \quad (\text{in diesem Abschnitt}). \quad (4.45)$$

Wir können damit den thermodynamischen Zustand mit zwei Variablen beschreiben.

- Da die Temperatur T und der Druck p als Ableitungen der inneren Energie U geschrieben werden können, können wir sie als Funktion

$$T = T(S, V) \quad \text{und} \quad p = p(S, V) \quad (4.46)$$

auffassen.

- Typischerweise verhalten sie diese Funktionen streng monoton in ihren Variablen. Damit können sie invertiert werden, d.h. als Funktionen

$$S = S(T, p) \quad \text{und} \quad V = V(T, p) \quad (4.47)$$

aufgelöst werden (Satz von der Umkehrabbildung aus der Analysis). Alternative können wir $-S$ und V als Ableitung der Gibbs-Energie $G(T, p)$ auffassen.

- Ähnlich bekommen wir die Koordinaten

$$S = S(T, V) \quad \text{und} \quad p = p(T, V) \quad (4.48)$$

indem wir die freie Energie $F(T, V)$ partiell ableiten.

Die Parametrisierung (4.47) über (T, p) liefert

$$\frac{dV}{V} = \frac{1}{V} \left(\frac{\partial V}{\partial T} \right)_p dT + \frac{1}{V} \left(\frac{\partial V}{\partial p} \right)_T dp. \quad (4.49)$$

Relative Volumenänderung dV/V und die Wärmezufuhr $\delta Q = TdS$ führen zu

- *Thermischer Ausdehnungskoeffizient*

$$\alpha := \frac{1}{V} \left(\frac{\partial V}{\partial T} \right)_{p,N} = \frac{1}{V} \frac{\partial^2 G}{\partial T \partial p} \quad (4.50)$$

- *Isotherme Kompressibilität*

$$\kappa_T := -\frac{1}{V} \left(\frac{\partial V}{\partial p} \right)_{T,N} = -\frac{1}{V} \frac{\partial^2 G}{\partial p^2} \quad (4.51)$$

Da G in p konkav ist folgt

$$\kappa_T \geq 0. \quad (4.52)$$

Die Eigenschaft (4.52) nennt man auch *mechanische Stabilität*, da sich der Stoff sonst immer weiter ausdehnen und der Druck immer mehr zunehmen würde.

Mit $dV = (\partial_T V)_{p,N} dT + (\partial_p V)_{T,N} dp$ gilt also

$$\frac{dV}{V} = \alpha dT - \kappa_T dp. \quad (4.53)$$

- *Wärmekapazität bei konstantem Druck*

$$C_p = T \left(\frac{\partial S}{\partial T} \right)_{p,N} = -\frac{1}{T} \frac{\partial^2 G}{\partial T^2}. \quad (4.54)$$

Mit $dH = TdS + \mu dN + Vdp$ haben wir

$$\frac{1}{T} \left(\frac{\partial H}{\partial T} \right)_{p,N} dT = \frac{1}{T} dH = dS = \left(\frac{\partial S}{\partial T} \right)_{p,N} dT, \quad (4.55)$$

daher ist auch

$$C_p = \left(\frac{\partial H}{\partial T} \right)_{p,N}. \quad (4.56)$$

Außerdem gilt

$$\delta Q = TdS = C_p dT, \quad (4.57)$$

d.h. C_p gibt Wärmemenge an, die erforderlich ist, um den Körper um 1 Kelvin zu erwärmen, wenn p und N festgehalten wird. Auf einem Weg mit konstantem

Druck p gilt dann

$$Q(T) - Q(T_0) = \int_{T_0}^T \delta Q = \int_{T_0}^T dT C_p(T). \quad (4.58)$$

Wir werden auch feststellen, dass

$$C_p \geq C_V \geq 0. \quad (4.59)$$

Ausgehend von den Koordinaten (4.48) definieren wir die folgenden weiteren Suszeptibilitäten.

- *Wärmekapazität bei konstantem Volumen V :*

$$C_V := T \left(\frac{\partial S}{\partial T} \right)_{V,N} = \left(\frac{\partial U}{\partial T} \right)_{V,N} = -\frac{1}{T} \frac{\partial^2 F}{\partial T^2}, \quad (4.60)$$

wobei im zweiten Schritt

$$dS = (dU + pdV - \mu dN) / T = (1/T) dU \quad (4.61)$$

verwendet wurde.

Aus der Konkavität von F in T folgt

$$C_V \geq 0. \quad (4.62)$$

Die Eigenschaft (4.62) nennt man auch *thermische Stabilität*, da sonst Wärmezufuhr die Temperatur des Körpers verringern und der Stoff so sämtliche Wärme „aufsaugen“ würde.

- *Isochorer Spannungskoeffizient bei konstantem Volumen V :*

$$\beta_{\text{sp}} := \frac{1}{p} \left(\frac{\partial p}{\partial T} \right)_{V,N} = -\frac{1}{p} \frac{\partial^2 F}{\partial T \partial V} \quad (4.63)$$

Die spezifischen Wärmen c_V, c_P folgen durch Division mit Masse der Substanz.

Ausgehend von den Koordinaten (p, S) ist die *adiabatische (isentropische) Kompressibilität* als

$$\kappa_S := -\frac{1}{V} \left(\frac{\partial V}{\partial p} \right)_{S,N} = -\frac{1}{V} \frac{\partial^2 H}{\partial p^2} \quad (4.64)$$

definiert. Da H in p konkav ist gilt

$$\kappa_S \geq 0. \quad (4.65)$$

19. VL

- Evaluation bis 16.12.19.

Die Maxwellrelationen führen zu *allgemeinen Beziehungen zwischen den Suszeptibili-*

täten (Übungsaufgabe):

$$\frac{C_p}{C_V} = \frac{\kappa_T}{\kappa_S} \quad (4.66)$$

$$C_p - C_V = \frac{TV\alpha^2}{\kappa_T} \quad (\text{ÜA. 10.1c}) \quad (4.67)$$

$$\kappa_T - \kappa_S = \frac{TV\alpha^2}{C_p} \quad (4.68)$$

$$p\beta_{\text{sp}} = \frac{\alpha}{\kappa_T} \quad (4.69)$$

Mit (4.67) folgt tatsächlich, dass $C_p \geq C_V$.

Übungsaufgabe: Berechnen Sie einige dieser Suszeptibilitäten für ein ideales Gas, d.h. für die thermische und kalorische Zustandsgleichung $pV = NkT$ und $U = \frac{3}{2}NkT$. Beispielsweise sind

$$\alpha = \frac{1}{T}, \quad (4.70)$$

$$\kappa_T = \frac{1}{p}, \quad (4.71)$$

$$C_V = \frac{3}{2}Nk. \quad (4.72)$$

Mit (4.67) folgt, dass

$$C_p - C_V = Nk \quad (4.73)$$

und daher

$$C_p = \frac{5}{2}Nk. \quad (4.74)$$

4.2.2.1 Adiabatangleichung

Wir betrachten ein ideales Gas,

$$U = \frac{3}{2}NkT, \quad pV = NkT, \quad (4.75)$$

mit konstanter Teilchenzahl N . Wegen $(\frac{\partial U}{\partial V})_T = 0$ ist dann

$$\begin{aligned} dS &= \frac{1}{T} (dU + pdV) \\ &= C_V \frac{dT}{T} \frac{p}{T} dV \\ &= C_V \frac{dT}{T} + Nk \frac{dV}{V} \end{aligned} \quad (4.76)$$

Im folgenden bezeichnen C_1, C_2, \dots Konstanten. Mit $C_p - C_V = Nk$ ergibt das

$$S(T, V) = C_V \ln(T) + (C_p - C_V) \ln(V) + C_1, \quad (4.77)$$

was man per Differentiation überprüfen kann. Die ideale Gasgleichung ergibt

$$S = C_V \ln(p) + C_p \ln(V) + C_2 \quad (4.78)$$

mit der Konstanten $C_2 = C_1 - C_V \ln(Nk)$. Nun setzen wir S konstant. Dann ist

$$\ln(p) + \kappa \ln(V) = C_3 \quad (4.79)$$

mit dem *Isotropenexponenten* $\kappa = C_p/C_V$. Anders ausgedrückt gilt die *Adiabatengleichung*

$$pV^\kappa = C_4, \quad (4.80)$$

wobei wegen (4.67) $\kappa > 1$ gilt. Damit ist $p(V)$ gegeben als

$$p(V) \propto \begin{cases} \frac{1}{V^\kappa} & \text{im adiabatischen Fall,} \\ \frac{1}{V} & \text{im isothermischen Fall.} \end{cases} \quad (4.81)$$

4.2.3 Gleichgewichtsbedingungen

Wir betrachten ein thermodynamisches System im Gleichgewicht und ändern nun einige der thermodynamischen Variablen. Wie werden nun die anderen Variablen darauf reagieren? Was ist der neue Gleichgewichtszustand? Diesen Frage wollen wir in diesem Abschnitt nachgehen. Dazu führen wir eine thermodynamische Größe ein, die *Exergie*, die bei Äquilibration nicht zunehmen kann.

Der Inhalt des Kapitels baut auf [1, Kapitel 2.3] auf.

4.2.3.1 Exergie

Wir betrachten Systeme Σ , Σ_E , und Σ_W , die nicht im Gleichgewicht sind. Weiterhin betrachten wir Σ_E als Umgebung/Bad von Σ mit Temperatur T_E und Druck p_E , die sich nicht ändern da wir Σ_E als sehr groß annehmen. Ein weiteres System Σ_W dient als Arbeitsreservoir für Σ , d.h. es kann Arbeit $W' \geq 0$ von Σ nach Σ_W adiabatisch transferiert werden. (Man mache sich zu der Gesamtsituation eine schematische Skizze).

Nun betrachten wir eine Zustandsänderung von Σ und Σ_E mit entsprechenden Energie-, Volumen- und Entropieänderungen $\Delta U, \Delta U_E, \Delta V, \Delta V_E$, und $\Delta S, \Delta S_E$ (Endwert minus Anfangswert). Wir nehmen an, dass Σ_E durchgehend im Gleichgewicht ist, d.h. dass die Zustandsänderung von Σ_E quasistatisch verläuft. Die Teilchenzahlen sollen dabei konstant bleiben. Insgesamt erhalten wir die Bilanz

$$\begin{aligned} \Delta V + \Delta V_E &= 0, \\ \Delta U + \Delta U_E &= -W'. \end{aligned} \quad (4.82)$$

Damit ist die von Σ_E and Σ abgegebene Arbeit

$$W = p_E \Delta V_E \quad (4.83)$$

und

$$Q = -T_E \Delta S_E \quad (4.84)$$

die abgegebene Wärme. Damit ist

$$\Delta U_E = -W - Q = -p_E \Delta V_E + T_E \Delta S_E \quad (4.85)$$

und so

$$\begin{aligned} \Delta S_E &= \frac{1}{T_E} (\Delta U_E + p_E \Delta V_E) \\ &= \frac{1}{T_E} (-\Delta U - W' - p_E \Delta V), \end{aligned} \quad (4.86)$$

wobei im zweiten Schritt (4.82) verwendet wurde. Da $\Sigma \times \Sigma_E$ als adiabatisch abgeschlossen angenommen wurden, d.h. $TdS + T_E dS_E = 0$, gilt

$$\Delta S + \Delta S_E \geq 0, \quad (4.87)$$

d.h. die Entropie nimmt zu, wenn $\Sigma \times \Sigma_E$ ins Gleichgewicht äquilibrieren. Um genau zu sein betrachten wir den Fall $T \leq T_E$. Dann ist $0 = \frac{T}{T_E} dS + dS_E \leq dS + dS_E$. Der

Fall $T \geq T_E$ folgt analog und damit auch (4.87).

Zusammen ist

$$\Delta S + \frac{1}{T_E} (-\Delta U - W' - p_E \Delta V) \geq 0 \quad (4.88)$$

und so ist

$$W' \leq -\Delta U + T_E \Delta S - p_E \Delta V =: -\Delta \Lambda. \quad (4.89)$$

Das motiviert die Definition der *Exergie* (availability)

$$\Lambda := U - U_0 - T_E(S - S_0) + p_E(V - V_0), \quad (4.90)$$

als maximal verfügbare Arbeit, wobei (S_0, U_0, V_0) den Gleichgewichtszustand von Σ mit Σ_E beschreibt. Automatisch gilt im Gleichgewicht, dass $\Lambda = 0$. Allgemein haben wir $\Delta \Lambda \leq -W' \leq 0$ und daher $\Lambda \geq 0$.

Übungsaufgabe: Die Situation wird durch Gibbs-Ensemble mit Verteilungen (2.98) beschrieben. Zeigen Sie, dass

$$S(\rho \| \rho_0) = \frac{\Lambda}{T_E} \quad (4.91)$$

ist, wobei ρ der Anfangszustand von Σ und ρ_0 der Gleichgewichtszustand von Σ mit Σ_E ist; $S(\rho \| \rho_0)$ bezeichnet deren relative Entropie (1.84) bzw. (3.43), je nachdem, ob klassisch oder quantenmechanisch gerechnet wird.

Die Beziehung (4.91) zeigt aus informationstheoretischer Sicht, dass $\Lambda \geq 0$ und im Gleichgewicht $\Lambda = 0$ ist.

4.2.3.2 Gleichgewichtsbedingungen

Zur Einführung der Exergie unter Berücksichtigung der Teilchenzahl verwenden wir den quantenmechanischen Rahmen, d.h. Dichteoperatoren, Obervablen und die relative Quantenentropie. Alle Argumenten und Rechnungen sind jedoch analog auch auf kontinuierliche klassische Verteilungen (mit Abschnitt 1.4) anwendbar. Wir betrachten nun zwei verallgemeinerte kanonische Dichteporen (Abschnitt 3.4)

$$\begin{aligned} \hat{\rho} &= \exp(\Psi - \lambda_\nu \hat{M}^{(\nu)}), \\ \hat{\rho}_0 &= \exp(\Psi_0 - \lambda_\nu^0 \hat{M}^{(\nu)}) \quad (\text{ESK}) \end{aligned} \quad (4.92)$$

auf dem gleichen Hilbertraum und mit den gleichen Observablen $\hat{M}^{(\nu)}$. Dann ist ihre relative Entropie (3.41)

$$\begin{aligned} S(\hat{\rho} \| \hat{\rho}_0) &= k \operatorname{Tr} [\hat{\rho} (\ln(\hat{\rho}) - \ln(\hat{\rho}_0))] \\ &= -S(\hat{\rho}) + S(\hat{\rho}_0) + \lambda_\nu^0 (\mathbf{M}^{(\nu)} - \mathbf{M}_0^{(\nu)}) \quad (\text{ESK}) \end{aligned} \quad (4.93)$$

mit $\mathbf{M}^{(\nu)} = \operatorname{Tr}[\hat{\rho} \hat{M}^{(\nu)}]$ und $\mathbf{M}_0^{(\nu)} = \operatorname{Tr}[\hat{\rho}_0 \hat{M}^{(\nu)}]$.

Mit der üblichen Wahl $\mathbf{M} = (U, V, \mathbf{N})$ und $\boldsymbol{\lambda} = (\beta, \beta p, -\beta \boldsymbol{\mu})$ legt dies die folgende Verallgemeinerung der *Exergie* nahe:

$$\Lambda := T_0 S(\rho \| \rho_0) = U - U_0 + p_0(V - V_0) - \mu_{0,j}(N^{(j)} - N_0^{(j)}) - T_0(S - S_0). \quad (4.94)$$

Wegen (1.85) gilt

$$\Lambda \geq 0. \quad (4.95)$$

Zwei thermodynamische Systeme, die sich jeweils im Gleichgewichtszustand ρ und ρ_0 befinden sind genau dann untereinander im Gleichgewicht, wenn $\rho = \rho_0$, was äquivalent zu $S(\rho \| \rho_0) = 0$, vgl. (3.44). Damit können wir die Exergie heranziehen, um Gleichgewichtssituation zu charakterisieren. Dazu betrachten wir im Folgenden Situationen, wo einige der thermodynamischen Variablen nicht im Gleichgewicht sind,

wobei U_0, V_0, \mathbf{N}_0 , usw. jeweils die thermodynamischen Größen im Gleichgewicht bezeichnen. Einige der Variablen werden nun auf ihrem Gleichgewichtswert konstant gehalten. Über die Exergie (4.94) können wir für entsprechende Situationen Rückschlüsse auf Extremalität des zu den Variablen passenden Potentials ziehen.

Beispiele:

- (i) Isolierte Systeme: Für konstantes U, V, \mathbf{N} ist

$$S - S_0 \leq 0 \quad (4.96)$$

und S ist *im Gleichgewicht maximal*.

- (ii) Isentropisch-isochores System: Für konstantes S, V, \mathbf{N} ist

$$U - U_0 \geq 0 \quad (4.97)$$

und $U(S, V, \mathbf{N})$ ist *im Gleichgewicht minimal*.

20. VL

- (iii) Isotherm-isochores System: Für konstantes T, V, \mathbf{N} ist

$$\begin{aligned} 0 &\leq \Lambda \\ &= \underbrace{(U - TS)}_{=F} - \underbrace{(U_0 - T_0 S_0)}_{=F_0} + S \underbrace{(T - T_0)}_{=0} + p_0 \underbrace{(V - V_0)}_{=0} - \mu_{0j} \underbrace{(N^{(j)} - N_0^{(j)})}_{=0} \\ &= F - F_0 \end{aligned} \quad (4.98)$$

(mit ESK) und $F(T, V, \mathbf{N})$ ist *im Gleichgewicht minimal*.

- (iv) Isotherm-isobares System: Für konstantes T, p, \mathbf{N} ist

$$\begin{aligned} 0 &\leq \Lambda \\ &= \underbrace{(U - TS + pV)}_{=G} - \underbrace{(U_0 - T_0 S_0 + p_0 V_0)}_{=G_0} - V \underbrace{(p - p_0)}_{=0} + S \underbrace{(T - T_0)}_{=0} - \mu_{0j} \underbrace{(N^{(j)} - N_0^{(j)})}_{=0} \\ &= G - G_0 \end{aligned} \quad (4.99)$$

(mit ESK) und $G(T, p, \mathbf{N})$ ist *im Gleichgewicht minimal*.

- (v) Isentropisch-isobares System: Für konstantes S, p, \mathbf{N} ist

$$\begin{aligned} 0 &\leq \Lambda \\ &= \underbrace{(U + pV)}_{=H} - \underbrace{(U_0 + p_0 V_0)}_{=H_0} - V \underbrace{(p - p_0)}_{=0} - T_0 \underbrace{(S - S_0)}_{=0} - \mu_{0j} \underbrace{(N^{(j)} - N_0^{(j)})}_{=0} \\ &= H - H_0 \end{aligned} \quad (4.100)$$

(mit ESK) und $H(S, p, \mathbf{N})$ ist *im Gleichgewicht minimal*; vgl. Übungsaufgabe 9.4.

- (vi) Isotherm-isochores System: Für konstantes $T, V, \boldsymbol{\mu}$ ist

$$\begin{aligned} 0 &\leq \Lambda \\ &= \underbrace{(U - TS - \mu_j N^{(j)})}_{=H} - \underbrace{(U_0 - T_0 S_0 - \mu_{0j} N_0^{(j)})}_{=H_0} - S \underbrace{(T - T_0)}_{=0} - p \underbrace{(V - V_0)}_{=0} + N^{(j)} \underbrace{(\mu_0 - \mu_{0j})}_{=0} \\ &= \Phi - \Phi_0 \end{aligned} \quad (4.101)$$

(mit ESK) und $\Phi(T, V, \boldsymbol{\mu})$ ist *im Gleichgewicht minimal*.

4.2.3.3 Anwendungsbeispiel: Dampfdruck

Wir betrachten zwei Phasen der selben Substanz im Gleichgewicht, N_1 Teilchen in einer Gasphase und N_2 Teilchen in einer flüssigen Phase. Durch Kopplung an ein größeres Bad Σ_E werden p und T konstant gehalten. Gleichzeitig soll die Gesamtteilchenzahl $N = N_1 + N_2$ erhalten bleiben.

Wir suchen nun den *Dampfdruck* $P(T) = p$, bei dem eine Koexistenz von Gas und Flüssigkeit, d.h. $N_1, N_2 > 0$, möglich ist.

Zuerst führen wir hierzu molare Größen ein: $n = N/N_A$ ist die Teilchenzahl in Mol, wobei N_A die Avogadro-Konstante ist und ähnlich für n_1 und n_2 . Mit g, g_1, g_2 bezeichnen wir die *molare Gibbs-Energie*, d.h. die Gibbs-Energie eines Mols. Eine analoge Notation kann für andere extensive Größen verwendet werden.

Im Gleichgewicht ist die gesamte Gibbs-Energie $G(T, p)$ mit $G = n_1 g_1 + n_2 g_2$ minimal. Zulässig Abweichungen vom Gleichgewicht erfüllen

$$\Delta n_1 + \Delta n_2 = 0. \quad (4.102)$$

Die entsprechende Abweichung der Gibbs-Energie muss wegen der Minimalität von G

$$\Delta G = g_1 \Delta n_1 + g_2 \Delta n_2 = (g_1 - g_2) \Delta n_1 \stackrel{!}{=} 0 \quad (4.103)$$

erfüllen (Minimalität von G und Taylor 1. Ordnung um den GG-Wert von n_1), d.h.

$$g_1(T, P(T)) \stackrel{!}{=} g_2(T, P(T)), \quad (4.104)$$

sofern die Dampfdruckkurve $P(T)$ als Funktion lokal existiert (Satz über implizite Funktionen, Abschnitt 4.2.1.1). Mit $dn = dn_1 + dn_2 = 0$ ist die Fundamentalgleichung

$$dg = -sdT + vdp \quad (4.105)$$

und mit $g_1 = g_2$ ist

$$-(s_2 - s_1)dT + (v_2 - v_1)dp = 0. \quad (4.106)$$

Damit ist

$$\frac{dP}{dT} = \frac{s_2 - s_1}{v_2 - v_1} \quad (\text{Clausius-Clapeyron-Gleichung}). \quad (4.107)$$

Mit der *molarer Verdampfungswärme* $q := (s_1 - s_2)T$ schreiben wir das als

$$\frac{dP}{dT} = \frac{q}{(v_1 - v_2)T}. \quad (4.108)$$

Beispiel: Ideales Gas (ein Gas weit weg vom kritischen Punkt). Mit idealer *Gaskonstante* $R = N_A k$ ist $v_1 = \frac{RT}{P(T)} \gg v_2$ und damit haben wir die Differentialgleichung

$$\frac{dP}{dT} = \frac{q}{RT^2} P(T), \quad (4.109)$$

die durch

$$P(T) = C \exp\left(-\frac{q}{RT}\right) \quad (\text{Dampfdruck ideales Gas}) \quad (4.110)$$

gelöst wird.

4.2.3.4 Anwendungsbeispiel: Gibbs-Energie in chemischen Reaktionen

Die Gibbs Energie $G(T, p) = H - TS$ gibt an, ob eine Reaktion „spontan“ abläuft. Ist $\Delta G = \Delta H - T\Delta S < 0$ so ist die Reaktion *exergonisch*. Bei konstanten T und p verringert die Reaktion also die Gibbs-Energie und könnte daher aus thermodynamischer Sicht also stattfinden. Mit $G = H - TS$ kann das so gedeutet werden, dass die Entropiezunahme als die treibende „Kraft“ einer chemischen Reaktion angesehen wer-

den kann, welche die durch H gemessene Wärmeerzeugung überkompensieren muss. (Vgl. Übungsaufgabe 9.3).

4.2.3.5 Anwendungsbeispiel: Massenwirkungsgesetz

Übungsaufgabe 11.3

4.3 Thermodynamische Prozesse

4.3.1 Grundbegriffe

- Im *Gleichgewicht* kann ein System durch U, V, N bzw. entsprechend konjugierte Variablen beschrieben werden. Alle anderen Größen folgen dann aus Zustandsgleichungen.
- Umgekehrt kann ein System, das sich nicht im Gleichgewicht befindet, nicht durch thermodynamische Zustandsvariablen beschrieben werden.
- Ein *thermodynamischer Prozess* ist ein Prozess bei dem mindestens Anfangs- und Endzustände thermodynamische Gleichgewichtszustände sind.
- Ein *reversibler Prozess* ist eine thermodynamische Zustandsänderung, die jederzeit wieder umgekehrt ablaufen könnte, ohne dass das System oder dessen Umgebung dabei bleibende Veränderungen erfahren. Die Definition von Reversibilität ist hier etwas enger als in der statistischen Mechanik (Abschnitt 2.2.3).
- Reversible Prozesse sind quasistatisch (Nichtgleichgewichtsprozesse sind nicht reversibel) und isentropisch (2. Hauptsatz). Bei Untersystemen kann die Entropie zu- oder abnehmen, doch die Gesamtentropie muss gleich bleiben.
- Es gibt quasistatische (also beliebig langsame) irreversible Prozesse (Gay-Lussac-Versuch). Nicht-quasistatische Prozesse sind immer irreversibel.
- Irreversible Prozesse sind also nicht umkehrbar, z.B. ein Übergang von einem Gleichgewichtszustand in einen Nichtgleichgewichtszustand. Erfahrungsgemäss geht das Nichtgleichgewichtssystem dann in einen neuen Gleichgewichtszustand über, aber der Übergang zwischen Anfangs- und Endzustand ist irreversibel (Entropie hat sich erhöht). Das System geht nicht mehr spontan in den Nichtgleichgewichts-Zwischenzustand zurück.
- Ist ein Prozess reversibel und wird die Wärme Q umgesetzt, so gilt wegen der Energieerhaltung

$$TdS_{\text{rev}} = \delta Q_{\text{rev}}, \quad (4.111)$$

wobei „rev“ betont, dass es sich um einen reversiblen Prozess handelt. Man spricht auch von *reversibel umgesetzte Wärme*.

Irreversible Prozesse

4.3.2 Irreversible Expansion eines Gases: Gay-Lussac-Versuch (1807)

Wir hatten gesagt, dass TdS in bestimmten Situationen als *Wärme* δQ interpretiert werden kann. Der Gay-Lussac-Versuch (mit Versuch von 1807) beschreibt die Irreversible Expansion eines Gases und gibt über den Zusammenhang von dS und δQ genaueren Aufschluss.

Der Gay-Lussac-Versuch darf nicht mit dem *Joule-Thomson-Effekt* (später) oder dem *Gesetz von Gay-Lussac* verwechselt werden, die ebenfalls die Ausdehnung von Gasen zum Thema haben.

pics/Gay-Lussac.png

Abbildung 4.1: Im Gay-Lussac-Versuch expandiert ein Gas in ein Vakuum.

Bei dem Versuch expandiert ein Gas adiabatisch (Abbildung 4.1). Der Prozess läuft bei konstanter innerer Energie $U = U_1 + U_2$ ab. Man verwendet ein isoliertes Gefäß mit dem Volumen $V = V_1 + V_2$. Zu Beginn sind die Volumina V_1 und V_2 mit einer Trennwand getrennt, in V_1 befindet sich ein Gas der Temperatur T und V_2 ist evakuiert. Dann wird die Trennwand entfernt und Gas strömt auch nach V_2 . Wie sehen nach Äquilibration die thermodynamischen Zustandsvariablen aus?

Bemerkung: Nach Entfernen der Trennwand setzen (auch bei sehr langsamer, also quasistatischer Prozessführung) irreversible (turbulente) Prozesse ein, die thermodynamisch nicht beschreibbar sind. Nur Anfangs- und Endzustand können thermodynamisch beschrieben werden. Insbesondere hat die Relation $\delta Q = TdS$ hier keine Bedeutung (gilt also nicht).

4.3.2.1 Ideales Gas

Vor Expansion sind Entropie, Druck und innere Energie durch

$$S = Nk \left(\frac{5}{2} + \ln \left[\frac{V_1}{N\lambda^3} \right] \right) \quad (4.112)$$

mit thermischer Wellenlänge $\lambda = h/\sqrt{2\pi mkT}$ [s. Sackur-Tetrode-Gleichung (2.91)], $p = NkT/V_1$ und $U = \frac{3}{2}NkT$ gegeben. Da $U(T, V)$ offensichtlich nicht volumenabhängig ist, herrscht nach der Expansion die gleiche Temperatur, also auch die gleiche thermische Wellenlänge.

Die einzige Änderung der Entropie (4.112) ist durch $V_1 \rightarrow V = V_1 + V_2$ gegeben und der Druck nach der Expansion ist $p = NkT/V$. Die Entropie erhöht sich also um

$$\Delta S = Nk \ln(V/V_1) > 0. \quad (4.113)$$

Die Entropie hat sich also erhöht, ohne dass Wärme zugeführt wurde, d.h. mit $\Delta Q = 0$.

Wir werden sehen, dass der 2. Hauptsatz für thermodynamische Prozesse in einem abgeschlossenen Gesamtsystem besagt, dass

$$TdS \geq \delta Q_{\text{rev}}, \quad (4.114)$$

wobei irreversible Prozesse durch $TdS > \delta Q_{\text{rev}}$ charakterisiert sind. Der hier betrachtete Prozess ist also irreversibel obwohl er quasistatisch geführt werden kann. Die Aufhebung der Zwangsbedingung (Trennwand) führt zu einer höheren Entropie. Eine spontane Umkehrrichtung des Prozesses ist nicht möglich (System nimmt nicht von sich aus Zustand vor Aufhebung der Zwangsbedingung ein).

21. VL

Im ersten Tutorium im neuen Jahr (7.1.) bietet Kathrin Dorn eine wiederholende Diskussion zur Quantenmechanik an. Diese wird nach der üblichen Besprechung des aktuellen Übungsblattes stattfinden.

4.3.2.2 Allgemeines Gas

Für allgemeine Gase kann man folgendes zeigen: (i) dass sich die Temperatur ändern kann und (ii) die Entropie im Gay-Lussac-Versuch immer ansteigt. Dies folgt aus den Gleichungen

$$\left(\frac{\partial T}{\partial V}\right)_U = \frac{p - T\left(\frac{\partial p}{\partial T}\right)_V}{C_V} \quad (4.115)$$

und

$$\left(\frac{\partial S}{\partial V}\right)_U = \frac{p}{T} > 0, \quad (4.116)$$

welche auf dem 11. Übungsblatt hergeleitet werden sollen.

Die Entropie muss also bei Expansion mit konstanter Energie U immer zunehmen. Dies bestätigt, dass das Ausströmen eines Gases immer irreversibel ist.

4.3.3 Mischung von Gasen

In Übungsaufgabe 6.2 wurde für die Mischung zweier idealer unterscheidbarer Gase gezeigt, dass die Entropieänderung

$$\Delta S = N_1 k \ln\left(\frac{V}{V_1}\right) + N_2 k \ln\left(\frac{V}{V_2}\right) \quad (4.117)$$

mit $V = V_1 + V_2$ ist, wobei zu Beginn N_1 . Die Mischungsentropie idealer Gase ist also gleich der Entropieerzeugung zweier entsprechender Gay-Lussac-Versuche!

Im dem Fall von Gasen ununterscheidbarer Teilchen ist

$$\Delta S = Nk \ln\left(\frac{V}{N}\right) - N_1 k \ln\left(\frac{V_1}{N_1}\right) - N_2 k \ln\left(\frac{V_2}{N_2}\right) = N_1 k \ln(p_1) + N_2 k \ln(p_2) - Nk \ln(p) \quad (4.118)$$

mit $N = N_1 + N_2$ und $p_x = N_x kT/V_x$. Das heisst, dass es, wie erwartet, bei gleichen Drücken $p_1 = p_2$ keine Entropieänderung gibt. Dafür ist der $1/N!$ -Faktor im Phasenraumvolumenelement (2.32) notwendig.

Reversible Prozesse

Isotherme und adiabatische Volumenänderungen können zu einem reversiblen Kreisprozess – dem Carnot-Prozess – zusammengesetzt werden.

4.3.4 Isotherme Expansion eines idealen Gases

Der Isotherme Prozess mit $p = NkT/V$ ist die reversible Variante vom Gay-Lussac-Versuch. Hier wird bei einer Expansion (quasistatisch) von V_0 nach V die Energie

$$\begin{aligned} -W &= \int_{V_0}^V p(V) dV \\ &= NkT \int_{V_0}^V \frac{dV}{V} \\ &= NkT \ln(V/V_0) > 0 \end{aligned} \quad (4.119)$$

als Arbeit geleistet (Aufgabe: Zeichnen Sie die Arbeit in einem p - V -Diagramm ein). Da die innere Energie $U = \frac{3}{2}NkT$ konstant bleibt wurde Wärme $Q = -W$ vom

Reservoir an das System abgegeben, was einer Entropieerhöhung um

$$\Delta S = \frac{Q}{T} = -\frac{W}{T} > 0 \quad (4.120)$$

entspricht. Dabei wurde jedoch dem *Gesamtsystem* keine Wärme zugeführt (ideales Gas). Wir werden sehen, dass die Zustandsänderung als Teil eines Carnot-Prozesses reversibel gemacht werden kann: Die Arbeit kann „verwendet“ werden, um das Gas wieder isotherm auf das ursprüngliche Volumen zu komprimieren.

4.3.5 Adiabatische Expansion eines idealen Gases

Aus Abschnitt 4.2.2 wissen wir, dass für ein ideales Gas

$$C_p = \frac{5}{2}Nk, \quad C_V = \frac{3}{2}Nk \quad (4.121)$$

ist. Der Isotrophenexponent $\kappa = C_p/C_V$ ist daher

$$\kappa = \frac{5}{3}. \quad (4.122)$$

Damit ist für einen adiabatischen Prozess

$$p \propto V^{-5/3}. \quad (4.123)$$

Die durch eine Expansion (quasistatisch) generierte Arbeit ist mit $p(V) = p_0(V/V_0)^{-\kappa}$ als

$$\begin{aligned} -W &= \int_{V_0}^V p \, dV \\ &= p_0 V_0^{5/3} \int_{V_0}^V V^{-5/3} \, dV \\ &= \frac{3}{2} NkT_0 \left(1 - (V/V_0)^{-2/3} \right) \end{aligned} \quad (4.124)$$

gegeben, wobei wir $p_0 V_0 = NkT_0$ verwendet haben. Wegen $T \, dS = \delta Q = 0$ bleibt die Entropie konstant und die Abgabe der mechanischen Arbeit geht mit einer Temperaturänderung einher. Im Gay-Lussac-Versuch hingegen wird die Arbeit „dissipiert“ und T bleibt konstant.

Abschließend sei noch bemerkt, dass der Adiabatenkoeffizient κ von der Gasart abhängt. In Übungsaufgabe 11.2 soll κ aus „mechanischen Eigenschaften“ eines Gases bestimmt werden.

22. VL

7.1.2020

4.3.6 Der Carnot-Prozess

Wir betrachten ein System, das an Wärmebäder der Temperaturen $T_2 > T_1$ und an ein Arbeitsreservoir koppelt. Wir nehmen an, dass das System einen reversiblen (also insbesondere quasistatischen) Kreisprozess (periodischer Prozess) beschreibt. An die Reservoirs werden jeweils die Wärmen Q_2, Q_1 und die geleistete Arbeit $-W$ abgegeben. Wir wählen die Vorzeichen so, dass Energien, die an das System übertragen werden, positiv und die Abgegebenen Energiebeträge negativ gezählt werden.

Für einen Zyklus gilt dann

$$\Delta U = W + Q_1 + Q_2 = 0 \quad (4.125)$$

$$\Delta S = \frac{Q_1}{T_1} + \frac{Q_2}{T_2} = 0, \quad (4.126)$$

da U und S Zustandsfunktionen sind und der Prozess quasistatisch durchlaufen wird. Die Entropieänderungen der Wärmebäder ist daher über

$$\Delta S_1 := -\frac{Q_1}{T_1} = \frac{Q_2}{T_2} =: -\Delta S_2 \quad (4.127)$$

gegeben. Der *Wirkungsgrad* des Kreisprozesses ist als

$$\eta := \frac{-W}{Q_2}, \quad (4.128)$$

also dem Verhältnis aus Arbeit und „verbrauchte“ Wärme definiert, wobei man die an das kältere Wärmebad abgegebene Wärme als verbraucht betrachtet. Mit (4.125) ist

$$\eta = \frac{Q_1 + Q_2}{Q_2} = 1 + \frac{Q_1}{Q_2}. \quad (4.129)$$

Aus (4.126) folgt

$$\frac{Q_1}{Q_2} = -\frac{T_1}{T_2} \quad (4.130)$$

und damit der maximale Wirkungsgrad, der mit dem 1. und 2. Hauptsatz kompatibel ist (Carnot-Fall),

$$\eta_C = 1 - \frac{T_1}{T_2} < 1. \quad (4.131)$$

Als reversibler Kreisprozess kann der Carnot-Prozess in zwei Richtungen betrieben werden:

- Vorwärtsbetrieb/Rechtsprozess (im T - S - und p - V -Diagramm): Für $Q_2 > 0$ wird $-W = \eta Q_2 > 0$ als Arbeit geleistet und Wärme $-Q_1 = Q_2 T_1/T_2 > 0$ an das T_1 -Wärmebad abgegeben (Wärmekraftmaschine).
- Rückwärtsbetrieb/Linksprozess: Für $Q_2 < 0$ wird Arbeit $W = \eta Q_2 > 0$ aufgenommen und Wärme $Q_1 = Q_2 T_1/T_2 > 0$ dem T_1 -Wärmebad entzogen und an das T_2 -Wärmebad abgegeben (Wärmepumpe/Kältemaschine).

Der Wirkungsgrad der Wärmepumpe ist als

$$\eta_W := \frac{-Q_2}{W} = 1/\eta \quad (4.132)$$

definiert und gibt im Carnot-Fall

$$\eta_{C,W} = \frac{T_2}{T_2 - T_1} > 1. \quad (4.133)$$

Analog ist der Wirkungsgrad der Kältepumpe als

$$\eta_K := \frac{Q_1}{W} \quad (4.134)$$

definiert und ist im Carnot-Fall

$$\begin{aligned} \eta_{C,K} &= \frac{-Q_2}{W} \frac{T_1}{T_2} = \frac{1}{\eta} \frac{T_1}{T_2} \\ &= \frac{T_1}{T_2 - T_1} = \eta_{C,W} - 1. \end{aligned} \quad (4.135)$$

Ergebnisse:

- (i) Der Carnot-Wirkungsgrad ist universell, d.h. hängt nur von der Temperatur der Bäder ab.
- (ii) Wärme kann nicht vollständig in Arbeit umgewandelt werden (ohne daß weitere Änderungen auftreten). Das schließt die Existenz von *Perpetua mobila 2. Art* aus.

Eine Carnot-Prozess kann man explizit aus isothermen und adiabatischen Prozessen zusammensetzen, d.h. es existieren idealisierte Kreisprozesse mit $\eta = \eta_C$.

Übungsaufgaben (s. auch Übungsblatt 12):

Skizzieren Sie die zugehörigen T - S - und p - V -Diagramme und rechnen Sie anhand der vier Phasen den Wirkungsgrad nach. Nehmen Sie dabei Bezug auf

$$\begin{aligned} -W &= \oint p \, dV \\ &= Q = \oint T \, dS. \end{aligned} \tag{4.136}$$

4.3.7 Der zweite Hauptsatz

Formulierungen des 2. Hauptsatzes

- (i) Im Gleichgewicht maximieren die intensiven Variablen die Entropie (s. auch Gleichgewichtsbedingungen in Abschnitt 4.2.3.2), d.h. die Entropie kann insbesondere in isolierten thermodynamische Systeme nicht abnehmen.
- (ii) Wärme kann nicht vollständig in Arbeit umgewandelt werden, ohne dass irgendwo weitere Änderungen auftreten (Thomson; Planck), d.h. $\eta < 1$.
- (iii) Wärme kann nicht von einem kälteren zu einem wärmeren System übergehen, ohne dass weitere Änderungen auftreten (Clausius 1850). Sonst könnte man Wärme vom T_1 -Reservoir zum T_2 -Reservoir transferieren und so $\eta < 1$ verletzen (Skizze in VL).
- (iv) Alle zwischen den Reservoiren mit Temperaturen T_1 und T_2 reversibel arbeitenden Carno-Kreisprozesse haben denselben Wirkungsgrad η_C ; dieser ist der maximal mögliche Wirkungsgrad für alle Vorwärtsprozesse, irreversible Prozesse eingeschlossen. Sonst könnte man die unterschiedlichen Carnot-Maschinen „verketteten“, einen als Wärmepumpe und einen als Wärmekraftmaschine und so den 2. Hauptsatz verletzen (s. VL).
- (v) Der Carnot-Wirkungsgrad ist der höchstmögliche, d.h. $\eta \leq \eta_C$ für alle Kreisprozesse (sonst könnte man den Prozess mit einem Carnot-Prozess verketteten und ein Perpetuum mobile erhalten).

Irreversibel arbeitende Maschinen haben einen Vorwärts-Wirkungsgrad η_+ (Wärmekraftmaschine) und einen Rückwärtswirkungsgrad η_- (Wärmepumpe) mit

$$\eta_+ \leq \eta_- . \tag{4.137}$$

Wie oben führt $\eta_+ > \eta_-$ über Verkettung dieser Prozesse zu einem Widerspruch.

- (vi) Es existiert eine Zustandsfunktion S mit $dS = \delta Q_{\text{rev}}/T$, die bei reversiblen adiabatischen Prozessen konstant ist. Allgemein gilt für abgeschlossene Systeme $dS \geq 0$.

Wir haben gezeigt, dass (i) \Rightarrow (ii).

Übungsaufgabe: Zeigen Sie, dass (ii) \Leftrightarrow (iii) und (ii) \Rightarrow (iv) (in der 22. VL vorgeführt). In Abschnitt 4.3.9 werden wir sehen, dass (iv) \Rightarrow (vi). Da (vi) im Wesentlichen eine Umformulierung von (i) ist, sind die unterschiedlichen Formulierungen des 2. Hauptsatzes äquivalent (bis auf die statistischen Interpretation von (i)).

4.3.8 Absolute Temperatur (Zusatzinfo)

Wodurch ist unsere Temperaturskala ausgezeichnet? Könnten wir z.B. auch eine Skala verwenden, die statt die Klevin-Skala eine Kelvin-Quadrat-Skala verwendet (z.B. $10\text{ K} \rightarrow 100\text{ K}^2$)?

In diesem Abschnitt werden wir zeigen, dass der Carnot-Prozess die Kelvin-Skala bis auf einen konstanten Faktor auszeichnet. Dabei gehen wir von einer willkürlichen Temperaturquantifizierung ϑ aus, wobei ϑ monoton bei Wärmezufuhr steigt. Sei nun $\vartheta_1 < \vartheta_2$. Vom Carnot-Prozess wissen wir, dass

$$f(\vartheta_1, \vartheta_2) := 1 - \eta_C(\vartheta_1, \vartheta_2) = -\frac{Q_1}{Q_2} \quad (4.138)$$

eine universelle Funktion ist (hängt nur von ϑ_1 und ϑ_2 ab). Nun betrachten wir zwei Carnot-Prozesse, einer für $\vartheta_1 < \vartheta_2$ und einer für $\vartheta_2 < \vartheta_3$. Da wir die Carnot-Maschinen aneinanderhängen können, gilt

$$f(\vartheta_1, \vartheta_3) = -\frac{Q_1}{Q'_3} = \left(-\frac{Q_1}{Q_2}\right) \left(-\frac{Q'_2}{Q'_3}\right) = f(\vartheta_1, \vartheta_2)f(\vartheta_2, \vartheta_3), \quad (4.139)$$

wobei für die Carnot-Maschine zwischen $\vartheta_2 < \vartheta_3$ die Wärme Q'_3 die Rolle von Q_2 und Q'_2 die Rolle von Q_1 spielt (Skizze). Nun wähle wir ϑ_2 als beliebigen Referenzwert. Daher ist

$$f(\vartheta_1, \vartheta_3) = a(\vartheta_1)b(\vartheta_3) \quad (4.140)$$

mit $a(\vartheta_1) = f(\vartheta_1, \vartheta_2)$ und $b(\vartheta_3) = f(\vartheta_2, \vartheta_3)$. Mit $f(\vartheta, \vartheta) = 1$ für alle ϑ kann man zeigen, dass

$$b(\vartheta_3) = \frac{1}{a(\vartheta_3)} =: \frac{1}{T(\vartheta_3)} \quad (4.141)$$

(ohne die Annahme $\vartheta_1 < \vartheta_2 < \vartheta_3$) und so

$$f(\vartheta_1, \vartheta_3) = \frac{T(\vartheta_1)}{T(\vartheta_3)}. \quad (4.142)$$

Wir bezeichnen $T(\vartheta)$ als *absolute Temperatur*. Mit $T_i := T(\vartheta_i)$ ist

$$\eta = 1 - f(\vartheta_1, \vartheta_2) = 1 - \frac{T_1}{T_2}. \quad (4.143)$$

Die Temperaturskala lässt noch einen Faktor $\alpha > 0$ offen, d.h. auch die Wahl $T'(\vartheta) = \alpha T(\vartheta)$ wäre möglich. Dies entspricht einer anderen Wahl des Referenzwertes ϑ_2 . Üblicherweise wählen wir ihn so, dass die Schmelztemperatur von Wasser den Wert 273.15 ergibt.

23. VL

Eine Datei pro E-Abgabe

4.3.9 Phänomenologische Entropie

Wir nehmen an, dass die Formulierung (vi) des 2. Hauptsatzes gilt und folgern in diesem Abschnitt die Existenz und Grundlegende Eigenschaft der Entropie.

Für alle reversiblen Kreisprozesse gilt, dass $\eta \stackrel{!}{=} \eta_C$, d.h.

$$\eta = \frac{-W}{Q_2} = \frac{Q_1 + Q_2}{Q_2} \stackrel{!}{=} 1 - \frac{T_1}{T_2} = \eta_C \quad (4.144)$$

und damit

$$\frac{Q_1}{T_1} + \frac{Q_2}{T_2} = 0. \quad (4.145)$$

Allgemeiner gilt für eine quasistatische Folge von Gleichgewichtszuständen mit Temperaturen $T_0, \dots, T_n = T_0$, dass

$$\sum_{i=1}^n \frac{\Delta Q_i}{T_i} = 0 \quad (4.146)$$

mit $\Delta Q_i = Q_i - Q_{i-1}$, wobei Anfangszustand mit T_0 und Endzustand gleich sind. Daher gilt

$$\oint \frac{\delta Q_{\text{rev}}}{T} = 0 \quad \forall \text{ reversiblen Kreisprozesse,} \quad (4.147)$$

denn diese Integrale können jeweils beliebig genau durch eine Diskretisierung (4.146) approximiert werden; der Index „rev“ betont, dass es sich um einen reversiblen Prozess handelt. Daraus folgt, dass eine Zustandsfunktion S existiert, so dass

$$dS = \frac{1}{T} \delta Q_{\text{rev}}, \quad (4.148)$$

d.h. $1/T$ ist ein integrierender Faktor für δQ_{rev} . Wir nennen S die *phänomenologische Entropie*, die bis auf eine konstante eindeutig ist und auch unserer zuvor definierten Entropie entspricht.

Die Formulierung (vi) des 2. Hauptsatzes besagt außerdem, dass für beliebige (nicht-reversible) Kreisprozesse

$$\eta = \frac{-W}{Q_2} = \frac{Q_1 + Q_2}{Q_2} \leq 1 - \frac{T_1}{T_2} = \eta_C \quad (4.149)$$

gilt und damit

$$\frac{Q_1}{T_1} + \frac{Q_2}{T_2} \leq 0. \quad (4.150)$$

Dies impliziert, dass

$$\oint \frac{\delta Q}{T} \leq 0 \quad \forall \text{ Kreisprozesse.} \quad (4.151)$$

Wenn wir nun einen beliebigen Prozess $A \rightarrow B$ haben, können wir ihn reversibel zu einem Kreisprozess $A \rightarrow B \xrightarrow{\text{rev}} A$ ergänzen. Dann ist

$$0 \geq \oint \frac{\delta Q}{T} = \int_A^B \frac{\delta Q}{T} + \underbrace{\int_B^A \frac{\delta Q_{\text{rev}}}{T}}_{S_B - S_A}, \quad (4.152)$$

d.h.

$$S_B - S_A \geq \int_A^B \frac{\delta Q}{T}. \quad (4.153)$$

Da dies für alle Prozesse $A \rightarrow B$ gilt, haben wir das folgende Resultat.

$$dS \geq \frac{\delta Q}{T} \quad (4.154)$$

Außerdem haben wir gesehen, dass eine Entropiefunktion existiert, die in isolierten Systemen nicht abnehmen kann, was weitgehend der Formulierung (i) entspricht.

4.3.10 Der 3. Hauptsatz der Thermodynamik

Im Kern besagt der 3. HS, dass es *nicht möglich ist, ein System bis zum absoluten Nullpunkt abzukühlen*.

Satz 4.4 (3. HS von Nernst 1905 und Präzisierung von Planck 1911):

Wir betrachten ein Quantensystem mit einem Hamiltonoperator \hat{H} . Sei g_0 die Entartung des Grundzustands von \hat{H} . Dann ist die Entropie des verallgemeinerten kanonischen Dichteoperators (3.66)

$$\lim_{T \rightarrow 0} S(M) = k \ln(g_0), \quad (4.155)$$

wobei $kT = 1/\beta = 1/\lambda_U$ ist.

Beweis für den kanonischen Zustand. Wir schreiben den Hamiltonoperator \hat{H} in seiner Eigenzerlegung als

$$\hat{H} = \sum_j E_j \hat{P}_j, \quad (4.156)$$

wobei $E_0 < E_1 < \dots$ die Energieeigenwerte und \hat{P}_j die Eigenprojektoren sind. Mit $g_j := \text{Tr}[\hat{P}_j]$ bezeichnen wir die Energieentartungen. Die Zustandssumme ist

$$Z = \text{Tr}[e^{-\beta \hat{H}}] = \sum_j e^{-\beta E_j} \text{Tr}[\hat{P}_j] = \sum_j g_j e^{-\beta E_j} \quad (4.157)$$

und damit der kanonische Zustand

$$\begin{aligned} \hat{\rho} &= \sum_j \frac{e^{-\beta E_j}}{\sum_k g_k e^{-\beta E_k}} \hat{P}_j \\ &= \sum_j \frac{e^{\beta E_0} e^{-\beta(E_j - E_0)}}{e^{\beta E_0} \sum_k g_k e^{-\beta(E_k - E_0)}} \hat{P}_j \\ &= \sum_j \frac{e^{-\beta(E_j - E_0)}}{\sum_k g_k e^{-\beta(E_k - E_0)}} \hat{P}_j \\ &= \frac{1}{\sum_k g_k e^{-\beta(E_k - E_0)}} \hat{P}_0 + \frac{1}{\sum_k g_k e^{-\beta(E_k - E_0)}} \sum_{j>0} e^{-\beta(E_j - E_0)} \hat{P}_j. \end{aligned} \quad (4.158)$$

Wegen

$$\lim_{\beta \rightarrow \infty} e^{-\beta(E_k - E_0)} = \delta_{k,0} \quad (4.159)$$

ist der Zustand bei $T = 0$

$$\lim_{\beta \rightarrow \infty} \hat{\rho} = \frac{1}{g_0} \hat{P}_0, \quad (4.160)$$

also der maximal gemischte Zustand auf dem Grundzustandseigenraum. Da die Entropie (3.32) stetig ist, haben wir damit

$$\lim_{T \rightarrow 0} S(\hat{\rho}) = \lim_{\beta \rightarrow \infty} S(\hat{\rho}) = k \ln(g_0). \quad (4.161)$$

Der allgemeine Beweis geht analog oder man kann mit der Äquivalenz der Ensemble argumentieren. \square

Bemerkungen und Konsequenzen:

- Typischerweise ist die Entartung g_0 klein, so dass $\ln(g_0)/N \approx 0$ für makroskopische Systeme ist. Genauer: die Spektren von Vielteilchenhamiltonoperatoren

sehen meist wie eine passend skalierte Normalverteilung aus, was man z.B. für bestimmte Fälle mithilfe eines nicht-kommutativen zentralen Grenzwertsatzes rigoros zeigen [23, 24] kann.

Damit ist die molare Entropie dann auch praktisch gleich Null.

- Mithilfe der Wärmekapazität $C_X = T \left(\frac{\partial S}{\partial T} \right)_X$ kann die Entropie als

$$S(T) - S(T=0) = \int_0^T dT' \frac{C_X(T')}{T'} \quad (4.162)$$

geschrieben werden. Daraus folgt, dass

$$\lim_{T \rightarrow 0} C_X(T) = 0, \quad (4.163)$$

da das Integral sonst keinen endlichen Wert ergeben würde (Übungsaufgabe 12.3).

- Es ist nicht möglich, ein System bis zum absoluten Nullpunkt abzukühlen:

Dazu müsste man dem System sämtliche Entropie entziehen. Für die bekannten Prozesse und realistische Hamiltonoperatoren kann man jedoch einfach argumentieren, dass dies nicht geht. Optimalerweise würde man hier wieder zwischen isothermen und adiabatischen Teilprozessen wechseln. Im S - T -Diagramm würde man sich dann auf einem „Kegel“ Schrittweise in Richtung $T = 0$ bewegen, wobei jeder Schritt zwar die Temperatur verringert, jedoch nicht auf $T = 0$ führt (Übungsaufgabe 12.3).

5 Modellsysteme und Phasenübergänge

In diesem Kapitel wollen wir einige spezielle Modellsysteme genauer diskutieren.

5.1 Ideales Molekül-Gas

Das ideale Gas haben wir als paradigmatisches Beispiel wiederholt diskutiert. Es beschreibt gleichartige Mikroteilchen ohne gegenseitige Wechselwirkung. Daher werden verdünnte Gase, wo der mittlere Teilchenabstand viel größer als die Reichweite des Wechselwirkungspotentials ist, korrekt beschrieben.

Die kanonische Zustandssumme (2.88) ist

$$Z_N = \frac{(V/\lambda^3)^N}{N!} \quad (5.1)$$

und die Großkanonische Zustandssumme (2.117)

$$\mathcal{Z} = \exp(zV/\lambda^3) \quad (5.2)$$

mit thermischer Wellenlänge $\lambda = \frac{h}{\sqrt{2\pi m k T}}$ und Fugazität $z = e^{\beta\mu}$. Weiterhin ergeben

sich die folgenden Zustandsgleichungen (Abschnitte 2.3.1.1, 2.3.3.2 und 4.1.1)

$$\frac{S}{kN} = \ln \left[\frac{V}{N\lambda^3} \right] + \frac{5}{2} \quad (\text{Sackur-Tetrode}) \quad (5.3)$$

$$U = \frac{3}{2} NkT \quad (\text{kalorische Zustandsgl.}), \quad (5.4)$$

$$N = \frac{Vz}{\lambda^3}, \quad (5.5)$$

$$pV = NkT \quad (\text{thermische Zustandsgl.}). \quad (5.6)$$

Die in der Herleitung auftauchenden Faktoren von $3/2$ müssen für ideale Molekülgase durch $f/2$ ersetzt werden, wobei f die Zahl der nicht-eingefrorenen Freiheitsgrade des Moleküls ist, die als quadratische Terme in der Einteilchenhamiltonfunktion auftauchen (Äquipartitionsprinzip, Übungsaufgaben 6.1 und 8.1). Beispielsweise hat ein 2-Atomiges Molekül (hantelförmig) 3 Translationsfreiheitsgrade, 2 Rotationsfreiheitsgrade und 2 Schwingungsfreiheitsgrade (kinetische + potentielle Energie). Unter Normalbedingungen ($T = 0^\circ\text{C}$, $p = 1\text{atm}$) sind bei Gasen wie Sauerstoff, Stickstoff, Wasserstoff die Schwingungsfreiheitsgrade eingefroren, was zu $f \approx 5$ führt und experimentell über die Messung der Wärmekapazitäten bestätigt werden kann. Genauer gibt es Temperaturen Θ_{rot} und Θ_{vib} , bei denen das Einfrieren der Freiheitsgrade stattfindet, was durch

$$C_V = \left(\frac{\partial U}{\partial T} \right)_{V,N} = \frac{f}{2} Nk \quad (5.7)$$

erfasst wird; genaueres in der VL oder [2, Abb. 5.3 und Erklärungen].

24. VL

Weil mein kleiner Sohn krank geworden ist und ich mich daher um ihn kümmern muss, fällt diese Vorlesung leider kurzfristig aus. Bitte lesen Sie bis zur nächsten Vorlesung am 17.1. den Abschnitt 5.2 bis einschließlich Abschnitt 5.2.2. Fragen zum Inhalt werden zu Beginn der nächsten Vorlesung besprochen (umgedrehter Unterricht). Zur Hilfe und für Skizzen kann z.B. [2, Kapitel 5.3] herangezogen werden.

*Viele Grüße
Martin Kliesch*

5.2 Reale Gase

Nun betrachten wir die Situation, wo die Wechselwirkungen zwischen den Teilchen nicht mehr vernachlässigt werden kann. Weiterhin nehmen wir an, dass die Temperatur so groß ist, dass Quanteneffekte vernachlässigt werden können. Entsprechend schreibe wir die Hamiltonfunktion von N Teilchen als

$$H(p, q) = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m} + \sum_{i < j} \varphi(\mathbf{q}_i - \mathbf{q}_j) + V_{\text{Wand}}(q), \quad (5.8)$$

wobei m die Teilchenmasse und φ das Wechselwirkungspotential der Teilchen ist. Mit (s. z.B. Übungsaufgabe 4.3)

$$\frac{1}{h} \int d\mathbf{p}_k \exp\left(-\frac{\beta}{2m} p_k^2\right) = \frac{1}{\lambda} \quad (5.9)$$

ist die kanonische Zustandssumme

$$Z = \int d\xi e^{-\beta H(\xi)} = \frac{1}{h^{3N} N!} \int_{V^N} d^3 \mathbf{p}_1 \dots d^3 \mathbf{p}_N d^3 \mathbf{q}_1 \dots d^3 \mathbf{q}_N e^{-\beta H(\mathbf{p}, \mathbf{q})} \quad (5.10)$$

$$= \underbrace{\frac{V^N}{N! \lambda^{3N}}}_{=: Z_{\text{ideal}}} \underbrace{\frac{1}{V^N} \int_{V^N} d^3 \mathbf{q}_1 \dots d^3 \mathbf{q}_N \exp \left(-\beta \sum_{i < j} \varphi(\mathbf{q}_i - \mathbf{q}_j) \right)}_{=: Z_{\text{pot}}}, \quad (5.11)$$

d.h. sie faktorisiert in einen kinetische Anteil Z_{ideal} , der identisch zur kanonischen Zustandssumme (2.88) eines idealen Gases ist, und einen Anteil für die potentielle Wechselwirkungsenergie Z_{pot} .

Um die Zustandssumme weiter auszurechnen werden wir sie für kleine Abweichungen vom idealen Gas nähern, indem wir die sogenannte Virialentwicklung von Z durchführen. Dazu führen wir

$$f_{i,j} := \exp(-\beta \varphi(\mathbf{q}_i - \mathbf{q}_j)) - 1 \quad (5.12)$$

ein, so dass $f_{i,j} \rightarrow 0$ für ein verschwindendes Wechselwirkungspotential φ . Nun entwickeln wir die Zustandssumme (5.11) nach Potenzen von $f_{i,j}$, was die *Virialentwicklung* liefert:

$$\begin{aligned} Z &= \frac{1}{N! \lambda^{3N}} \int_{V^N} d^3 \mathbf{q}_1 \dots d^3 \mathbf{q}_N \prod_{i < j} (1 + f_{i,j}) \\ &= \frac{1}{N! \lambda^{3N}} \int_{V^N} d^3 \mathbf{q}_1 \dots d^3 \mathbf{q}_N \left(1 + \sum_{i < j} f_{i,j} + \sum_{\substack{i < j \\ k < \ell}} f_{i,j} f_{k,\ell} + \dots \right). \end{aligned} \quad (5.13)$$

Hier betrachten wir die niedrigste Ordnung, die vom idealen Gas abweicht:

$$Z = \frac{1}{N! \lambda^{3N}} \left[V^N + V^{N-2} \sum_{i < j} \int_{V^2} d^3 \mathbf{q}_i d^3 \mathbf{q}_j f_{i,j}(\mathbf{q}_i - \mathbf{q}_j) + O(V^{N-2}) \right]. \quad (5.14)$$

Mit der Relativkoordinate $\mathbf{r} := \mathbf{q}_1 - \mathbf{q}_2$, der Ununterscheidbarkeit der Teilchen, der Annahme, dass $\varphi(\mathbf{r} + \mathbf{a}) = \varphi(\mathbf{r}) = \varphi(-\mathbf{r})$ für alle \mathbf{r} und \mathbf{a} ist und der Zustandssumme des idealen Gases (5.2) ergibt das

$$\begin{aligned} Z &= \frac{1}{N! \lambda^{3N}} \left[V^N + V^{N-2} \frac{N(N-1)}{2} \int d^3 \mathbf{r} f_{i,j}(\mathbf{r}) + O(V^{N-2}) \right] \\ &= Z_{\text{ideal}} \underbrace{\left[1 + \frac{N(N-1)}{V} B(T) + O(1/V^2) \right]}_{=: Z_{\text{pot}}} \end{aligned} \quad (5.15)$$

mit dem 2. *Virialkoeffizienten*

$$B(T) := \frac{1}{2} \int d^3 \mathbf{r} f_{1,2}(\mathbf{r}) \geq 0, \quad (5.16)$$

welcher die intermolekulare 2-Teilchenwechselwirkung berücksichtigt.

5.2.1 Thermische Zustandsgleichung

Es gilt (s. z.B. die Tabelle in Abschnitt 4.2)

$$p = -\frac{\partial F}{\partial V} = kT \frac{\partial \ln(Z)}{\partial V}. \quad (5.17)$$

Mit (5.15) ist die freie Energie¹

$$\begin{aligned} F &= -kT \ln(Z) = -kT (\ln(Z_{\text{ideal}}) + \ln(Z_{\text{pot}})) \\ &\approx F_{\text{ideal}} - kT \ln\left(1 + \frac{N(N-1)}{V} B(T)\right) \\ &\approx F_{\text{ideal}} - kT \frac{N(N-1)}{V} B(T). \end{aligned} \quad (5.18)$$

wobei F_{ideal} die freie Energie eines idealen Gases bezeichnet und wir $\ln(1+x) = x + O(x^2)$ verwendet haben. Mit $-\frac{\partial F_{\text{ideal}}}{\partial V} = p_{\text{ideal}} = NkT/V$ und $N(N-1) \approx N^2$ erhalten wir schließlich

$$\begin{aligned} p &\approx \frac{NkT}{V} - kT B(T) \frac{N^2}{V^2} \\ &= \frac{NkT}{V} \left(1 - \frac{N}{V} B(T)\right). \end{aligned} \quad (5.19)$$

Dies sind die ersten zwei Summanden einer Reihenentwicklung in der Teilchendichte N/V mit dem 1. Virialkoeffizienten 1 und 2. Virialkoeffizienten $B(T)$. Den Koeffizienten $B(T)$ kann man experimentell bestimmen oder mit ausgefeilteren Herleitungen ausrechnen. Dazu muss man ein konkretes Potential φ annehmen. Ein gutes Modell ist durch das Lennard-Jones-Potential

$$\varphi(|\mathbf{r}|) = \frac{c_1}{|\mathbf{r}|^{12}} - \frac{c_2}{|\mathbf{r}|^6} \quad (5.20)$$

gegeben. Der r^6 -Term kommt aus einer Dipol-Dipol-Näherung anziehender Wechselwirkungen und der r^{12} -Term liefert eine starke kurzreichweitige Abstoßung (die 12 im Exponenten ist kein Tippfehler!).

Ein einfacheres Modell ist eines mit harten Kugeln festen Eigenvolumens, außerhalb dessen ein anziehendes Potential schnell abfällt, d.h.

$$\varphi(r) \begin{cases} = \infty & \text{für } r \leq d \\ < 0 & \text{für } r > d \end{cases}$$

in Kugelkoordinaten. Für (5.12) ergibt ein solches Potential

$$f_{1,2}(r) \approx \begin{cases} -1 & \text{für } r < d, \\ -\beta\varphi(r) & \text{für } r > d. \end{cases} \quad (5.21)$$

Bezogen auf ein Mol (d.h. $N \mapsto N_A$) führt dies für den 2. Virialkoeffizienten (5.16) zu

$$\begin{aligned} B(T) &= -\frac{1}{2} \frac{4\pi}{3} d^3 - \frac{1}{2} \int_d^\infty \beta\varphi(r) 4\pi r^2 dr \\ &= -\frac{b}{N_A} + \frac{a}{kTN_A^2} \end{aligned} \quad (5.22)$$

mit molarem *Eigenvolumen*

$$b := 4N_A \frac{4\pi}{3} \left(\frac{d}{2}\right)^3 \geq 0, \quad (5.23)$$

¹Wir haben Z nur approximativ ausgerechnet. Im allgemeinen impliziert $f(x) \approx g(x)$ jedoch nicht, dass auch $\frac{\partial f}{\partial x}(x) \approx \frac{\partial g}{\partial x}(x)$ sein muss. Zur der hier gemachten und üblichen Approximation von F geht noch die Hypothese ein, dass sich die Natur auf makroskopischen Skalen ausreichen „gutartig“ verhält.

wobei d der Durchmesser der Kugeln ist (Skizze), und dem *Binnendruck*

$$a := \frac{1}{2} N_A^2 \int_d^\infty |\varphi(r)| 4\pi r^2 dr \geq 0 \quad (\text{Kugelkoordinaten}). \quad (5.24)$$

Mit (5.19) und Gaskonstante $R := kN_A$ ergibt das

$$p = \frac{RT}{v} \left(1 + \frac{b}{v} \right) - \frac{a}{v^2}. \quad (5.25)$$

Für *verdünnte Gase*, d.h. $b \ll v$ stimmt diese Gleichung überein mit der empirischen Van-der-Waals-Gleichung

$$p = \frac{RT}{v-b} - \frac{a}{v^2} \quad (5.26)$$

(s. auch [2, Abb. 5.11] oder die VL für eine Skizze), da dann $\frac{1}{1-b/v} \approx 1 + b/v$ ist. Häufig wird sie auch wie folgt geschrieben.

Van-der-Waals-Gleichung

$$RT = \left(p + \frac{a}{v^2} \right) (v - b) \quad (5.27)$$

Hier können a und b durch weitere Rechnungen genauer abgeschätzt und mit dem führenden Fehler aus (5.18) ein Gültigkeitsbereich der Van-der-Waals-Gleichung beschrieben werden; genauer: $v \gg b$ und $kT \gg \langle \varphi \mathbf{1}_{[d,\infty)} \rangle := \int_d^\infty r^2 \varphi$. Man kann a und b aber auch empirisch bestimmen und damit den Gültigkeitsbereich der Van-der-Waals-Gleichung vergrößern.

Für eine ausreichend kleine konstante Temperatur T hat $p(v)$ aus (5.26) ein lokales Maximum, ein lokales Minimum und einen Wendepunkt. Wir werden sehen, dass das Van-der-Waals-Gas bei diesen Temperaturen Phasenübergänge durchläuft, wenn man das Volumen geeignet variiert. Für eine ausreichend große konstante Temperatur T ist $p(v)$ monoton fallend. Daraus folgt, dass ein Punkt (T_c, p_c, v_c) existiert, bei dem $p(v)$ einen Sattelpunkt hat, der auch *kritischer Punkt* genannt wird. Für (T_c, v_c) muss

$$\frac{\partial p}{\partial v} = 0 = \frac{\partial^2 p}{\partial v^2} \quad (5.28)$$

sein. Daraus ergeben sich die Werte (Übungsaufgabe 13.1a)

$$v_c = 3b, \quad RT_c = \frac{8}{27} \frac{a}{b}, \quad p_c = \frac{a}{27b^2} \quad (5.29)$$

für den kritischen Punkt.

5.2.2 Kalorische Zustandsgleichung

Mit (2.20) und (5.15) haben wir

$$\begin{aligned} U &= -\frac{\partial \ln(Z)}{\partial \beta} \\ &\approx -\frac{\partial \ln(Z_{\text{ideal}})}{\partial \beta} - \frac{\partial}{\partial \beta} \ln \left(1 + \frac{N^2}{V} B(T) \right) \\ &\approx \frac{3}{2} NkT + \frac{N^2}{V} \frac{\partial B(T)}{\partial T} \frac{\partial T}{\partial \beta}, \end{aligned} \quad (5.30)$$

wobei wir wieder $\ln(1+x) = x + O(x^2)$ verwendet haben. Mit $\frac{\partial T}{\partial \beta} = -kT^2$ ist

$$U = \frac{3}{2}NkT + \frac{N^2kT^2}{V} \frac{\partial B(T)}{\partial T} \quad (5.31)$$

die kalorische Zustandsgleichung. Für den Van-der-Waals-Fall (5.22) haben wir

$$\frac{\partial B(T)}{\partial T} = -\frac{a}{kT^2N_A^2} \quad (5.32)$$

und damit (molarer Schreibweise):

Van-der-Waals-Gas: Kalorische Zustandsgleichung

$$u = \frac{3}{2}RT - \frac{a}{v}. \quad (5.33)$$

5.2.3 Der Joule-Thomson-Prozess

(In der 26. VL behandelt.)

Dieser Prozess beschreibt die gedrosselte Expansion eines Gases (Skizze in VL).

- Anfangszustand: V_1, U_1
- Endzustand: V_2, U_2

Dabei wird das Gas mittels eines Stempels aus dem Volumen V_1 durch eine Drossel in ein Volumen V_2 gepresst, welches wiederum an einen Stempel gekoppelt ist. Durch die beiden Stempel werden die Drücke jeweils auf p_1 und p_2 gehalten. Der gesamte Prozess erfolgt adiabatisch.

Die Gesamtänderung der inneren Energie ist daher

$$U_2 - U_1 = \int_{V_1}^0 (-p_1) dV + \int_0^{V_2} (p_2) dV = p_1 V_1 - p_2 V_2. \quad (5.34)$$

Daraus folgt, dass die Enthalpie $H = U + pV$ bei dem Vorgang unverändert ist. Die Temperaturänderung ist daher durch den *Joule-Thomson-Koeffizienten*

$$\mu_{JT} := \left(\frac{\partial T}{\partial p} \right)_H \quad (5.35)$$

gegeben. Man kann zeigen (Übungsaufgabe 13.3b), dass

$$\mu_{JT} = \frac{V}{C_p} (T\alpha - 1), \quad (5.36)$$

wobei C_p die Wärmekapazität (1.14) und α der thermische Ausdehnungskoeffizient (4.50) ist.

Für ein ideales Gas gilt nach (4.70) $\alpha = 1/T$ und es kommt zu keiner Temperaturänderung. Für reale Gase kann α jedoch größer oder auch kleiner als $1/T$ sein, d.h., sowohl Erwärmungen (negativer Joule-Thomson-Effekt) als auch Abkühlungen (positiver Joule-Thomson-Effekt) des Gases sind möglich. Die Grenze zwischen diesen beiden Effekten wird durch die *Inversionskurve* $p_{\text{inv}}(V)$ gegeben, bei der

$$\alpha = \frac{1}{T} \quad (5.37)$$

gilt.

Beispiel: Für ein Van-der-Waals-Gas gilt nach (5.26)

$$p = \frac{RT}{v-b} - \frac{a}{v^2} \quad (5.38)$$

und Differenzieren nach T liefert

$$0 = \frac{R}{v-b} - \frac{RT}{(v-b)^2} \left(\frac{\partial v}{\partial T} \right)_p + \frac{2a}{v^3} \left(\frac{\partial v}{\partial T} \right)_p. \quad (5.39)$$

Einsetzen der Inversionsbedingung $\alpha := \frac{1}{v} \left(\frac{\partial v}{\partial T} \right)_p = \frac{1}{T}$ ergibt

$$\begin{aligned} 0 &= \frac{R}{v-b} - \frac{Rv}{(v-b)^2} + \frac{2a}{Tv^2} \\ \Leftrightarrow 0 &= \frac{RT}{v} - \frac{RT}{v-b} + \frac{2a(v-b)}{v^3} = -\frac{b}{v} \frac{RT}{v-b} + \frac{2a(v-b)}{v^3} \end{aligned} \quad (5.40)$$

und mit (5.38)

$$\begin{aligned} 0 &= -\frac{b}{v} \left(p_{\text{inv}} + \frac{a}{v^2} \right) + \frac{2a(v-b)}{v^3} \\ \Leftrightarrow p_{\text{inv}} &= \frac{2a}{bv} - \frac{3a}{v^2}. \end{aligned} \quad (5.41)$$

Außerdem kann man eine Inversionskurve $p_{\text{inv}}(T)$ herleiten, s. [Wikipedia](#) für die Herleitung und [2, Abb. 3.14a] oder die VL für eine Skizze. Oberhalb der Kurve ist der Joule-Thomson-Koeffizient immer negativ, unterhalb immer positiv.

25. VL

5.2.4 Phasenübergänge

Die Van-der-Waals-Gleichung (5.27) lässt sich equivalent als

$$pv^3 - (RT + pb)v^2 + av - ab = 0 \quad (5.42)$$

schreiben. Damit gibt es T, p , so dass drei Lösungen für v existieren, was insbesondere durch eine Skizze von $p(v)$ für verschiedenen Werte von T deutlich wird; s. [2, Abbildung 5.11], die VL und Übungsaufgabe 13.1b.

Wir erinnern uns dazu an den kritischen Punkt (5.29) (vgl. Übungsaufgabe 13.1a). Danach gibt es ein Temperaturintervall mit $T < T_c$, für den

$$\left(\frac{\partial p}{\partial v} \right)_T = -\frac{RT}{(v-b)^2} + \frac{2a}{v^3} > 0 \quad (5.43)$$

ist. Insbesondere ist die isotherme Kompressibilität negativ,

$$\kappa_T = -\frac{1}{v} \left(\frac{\partial v}{\partial p} \right)_T < 0. \quad (5.44)$$

Damit wird die Stabilitätsbedingung (4.52) verletzt!

Die Stabilitätsgrenze ist durch $\left(\frac{\partial p}{\partial v} \right)_T = 0$ definiert, wobei für $\left(\frac{\partial p}{\partial v} \right)_T > 0$ mechanisch Stabilität gegeben und für $\left(\frac{\partial p}{\partial v} \right)_T < 0$ verletzt ist. Daraus folgt, dass wenn wir uns der Stabilitätsgrenze nähern die Kompressibilität divergiert,

$$\kappa_T \propto -\frac{1}{\left(\frac{\partial p}{\partial v} \right)_T} \rightarrow \infty. \quad (5.45)$$

Mithilfe von (4.40) folgt auch, dass der thermische Ausdehnungskoeffizient und die Wärmekapazitäten divergieren,

$$\alpha = \frac{1}{v} \left(\frac{\partial v}{\partial T} \right)_p = - \frac{\left(\frac{\partial p}{\partial T} \right)_V}{v \left(\frac{\partial p}{\partial v} \right)_T} \rightarrow \infty \quad (5.46)$$

$$c_p = c_v + \frac{T v \alpha^2}{\kappa_T} = c_v - T \frac{\left(\frac{\partial v}{\partial T} \right)_p^2}{\left(\frac{\partial v}{\partial p} \right)_T} = c_v + T \left(\frac{\partial v}{\partial T} \right)_p \left(\frac{\partial p}{\partial T} \right)_v \rightarrow \infty, \quad (5.47)$$

wobei für c_p und c_v zusätzlich (4.67) verwendet wurde. In der Übungsaufgabe 13.1d wird gezeigt, dass das singuläre Verhalten durch sogenannte *kritische Exponenten* beschrieben werden kann, d.h. z.B.

$$\kappa_T \propto \left(\frac{T - T_c}{T_c} \right)^{-\gamma} \quad (5.48)$$

mit kritischem Exponenten $\gamma = -1$.

Nach dem Fluktuations-Dissipations-Satz (2.43) sind die Volumenfluktuationen bzw. Dichtefluktuationen durch

$$\text{Var}(v) \equiv \langle (\Delta v)^2 \rangle = -kT \left(\frac{\partial v}{\partial p} \right)_T = kT v \kappa_T \rightarrow \infty \quad (5.49)$$

gegeben und divergieren an der Stabilitätsgrenze ebenfalls.

Allgemein können wir festhalten, dass an der Stabilitätsgrenze die Suszeptibilitäten, die als 2. Ableitungen Gibbs-Energie $G(T, p)$ gegeben sind (s. Abschnitt 4.2.2), divergieren. Solche ein singuläres Verhalten wird allgemeiner als *Phasenübergang* definiert. Physikalisch wird an der Stabilitätsgrenze eine Koexistenz verschiedener Phasen beobachtet. Im p - V -Diagramm für $T < T_c$ ist der Bereich zwischen den zwei lokalen Extrema instabil und „trennt“ die flüssige und gasförmige Phase (s. VL für eine Skizze).

5.2.4.1 Maxwell-Konstruktion für Phasenkoexistenz

Nach der Gleichgewichtsbedingung (4.99) für die Gibbs-Energie gilt

$$g_1(T, P(T)) = g_2(T, P(T)), \quad (5.50)$$

wobei g_1 und g_2 jeweils die molaren Gibbs-Energien in der flüssigen und gasförmigen Phase bezeichnen und $P(T)$ den Gleichgewichtsdampfdruck. In Bezug auf die molare freie Energie $f = g - pv$ bedeutet das

$$f_2 - f_1 + (v_2 - v_1)P = 0 \quad (5.51)$$

ist. Mit $\left(\frac{\partial f}{\partial v} \right)_T = -p$ folgt

$$\int_{v_1}^{v_2} \left(\frac{\partial f}{\partial v} \right)_T dv + (v_2 - v_1)P = 0, \quad (5.52)$$

was die folgende *Flächengleichheitsregel* impliziert.

Maxwell-Konstruktion

Das Integral

$$(v_2 - v_1)P = \int_{v_1}^{v_2} p dv \quad (5.53)$$

bestimmt die *Maxwell-Gerade* $P(T)$ (im p - V -Diagramm) sowie die molaren Volumina v_2 und v_1 .

Siehe [2, Abbildung 5.14] oder die VL für eine Skizze. Daraus folgt, dass $v(p)$ bei konstantem $T < T_c$ unstetig ist. Wir erinnern uns auch an die Dampfdruck-Kurve (4.110) eines idealen Gases, $P(T) \propto e^{-q/(RT)}$, was die Temperaturabhängigkeit approximativ beschreibt.

Unter *Metastabilität* versteht man Phänomäne wie den *Siedeverzug* oder den übersättigten Dampf, die mit der Van-der-Waals-Beschreibung kompatibel sind. Solche metastabilen Zustände sind allgemein typisch für Phasenübergänge.

26. VL

Von Raphael Brieger gehalten; der Joule-Thomson-Prozess aus Abschnitt 5.2.3 wurde ebenfalls in dieser VL behandelt.

5.2.4.2 Klassifikation von Phasenübergängen (Ehrenfest)

Wir haben gesehen, dass im Van-der-Waals-Gas bei dem Phasenübergang zwischen flüssiger und gasförmiger Phase die zweiten Ableitungen der Gibbsschen Energie divergieren und $V = \left(\frac{\partial G}{\partial p}\right)_T$ unstetig ist. Das motiviert die folgende Definition.

Definition 5.1 (Phasenübergang n -ter Ordnung (Ehrenfest)):

Bei einem Phasenübergang n -ter Ordnung ist mindestens eine der n -ten Ableitungen der Gibbs-Energie G unstetig.

Bemerkungen zum Van-der-Waals-Gas:

- Ein *Phasenübergang 1. Ordnung* liegt vor, wenn $\left(\frac{\partial G}{\partial p}\right)_T = V$ unstetig ist. Hier beobachten wir (i) eine Koexistenz von Phasen und (ii) dass beim Phasenübergang *latente Wärme* (oben Verdampfungswärme $q = (s_2 - s_1)T$, vgl. Clausius-Glaueyron-Gleichung in Abschnitt 4.2.3.3) auftritt. Das liegt daran, dass

$$S = -\left(\frac{\partial G}{\partial T}\right)_p \quad (5.54)$$

dann auch unstetig ist.

- Ein *Phasenübergang 2. Ordnung* liegt vor, wenn $\left(\frac{\partial G}{\partial p}\right)_T = V$ stetig ist, aber $\left(\frac{\partial^2 G}{\partial p^2}\right)_T = \left(\frac{\partial V}{\partial p}\right)_T$ unstetig ist. Ein solcher Phasenübergang liegt vor, wenn im p - V -Diagramm der kritische Punkt vertikal und senkrecht zum Phasenkoexistenzgebiet durchlaufen wird.

Dann gibt es keine Phasenkoexistenz, es tritt keine latente Wärme auf, da $S = -\left(\frac{\partial G}{\partial T}\right)_p$ stetig ist und muß durch den kritische Punkt führen. Dies kann durch universelle kritische Exponenten beschrieben werden, geht mit kritischen Fluktuationen einher und es tritt eine kritische Verlangsamung der Relaxation ins Gleichgewicht auf.

- In einer alternativen Formulierung der Ehrenfest-Klassifikation wird die freie Energie anstelle der Gibbs-Energie verwendet.
- Durch die Klassifikation ist nicht abgedeckt, dass Ableitungen der Potentiale bei Phasenübergängen divergieren. Es gibt eine modernere Klassifikation in Phaseübergänge 1. und 2. Ordnung. Diese Klassifikation ist und z.B. auch *Supraleitung* mit einschließen.

5.3 Ideale Quantengase II

In Abschnitt 3.7 haben wir die großkanonische Zustandssumme

$$\mathcal{Z} = \sum_{\{N_{\mathbf{p}}\}} \exp\left(-\beta \sum_{\mathbf{p}} (E_{\mathbf{p}} - \mu) N_{\mathbf{p}}\right) \quad (5.55)$$

für ideale Quantengase kennengelernt, wobei $\sum_{\mathbf{p}} N_{\mathbf{p}} = N$; hier haben wir die Summen über die Impulseigenwerte $\mathbf{p} = (p_x, p_y, p_z)$ anstelle über einen Zählindex j parametrisieren, da wir ja ein ideales Gas mit Hamiltonoperator der Form $\hat{H} = \frac{\hat{\mathbf{p}}^2}{2m}$ betrachten. Aus dem Kontext heraus sollte klar sein, wenn p den Impuls und wann den Druck bezeichnet.

Das ergibt

$$\mathcal{Z} = \prod_{\mathbf{p}} (1 + t_{\mathbf{p}}) \quad \text{für Fermionen,} \quad (5.56)$$

$$\mathcal{Z} = \prod_{\mathbf{p}} \frac{1}{1 - t_{\mathbf{p}}} \quad \text{für Bosonen} \quad (5.57)$$

mit

$$t_{\mathbf{p}} := e^{-\beta(E_{\mathbf{p}} - \mu)}, \quad (5.58)$$

wobei für Bosonen $E_{\mathbf{p}} > \mu$ für alle \mathbf{p} angenommen werden muss.

Damit folgt für ideales Quantengas das *grosskanonische Potential*

$$\Phi(T, V, \mu) = -kT \ln(\mathcal{Z}) = \frac{\xi}{\beta} \sum_{\mathbf{p}} \ln(1 - \xi e^{-\beta(E_{\mathbf{p}} - \mu)}) , \quad (5.59)$$

mit $\xi = 1$ für Bosonen und $\xi = -1$ für Fermionen.

Daraus folgen alle thermodynamischen Größen:

- Mittlere Teilchenzahl bei vorgegebenem μ :

$$N = - \left(\frac{\partial \Phi}{\partial \mu} \right)_{T, V} = \sum_{\mathbf{p}} \frac{e^{-\beta(E_{\mathbf{p}} - \mu)}}{1 - \xi e^{-\beta(E_{\mathbf{p}} - \mu)}} = \sum_{\mathbf{p}} N_{\mathbf{p}} \quad (5.60)$$

mit der *Bose-Einstein* bzw. *Fermi-Dirac-Verteilungsfunktion*

$$N_{\mathbf{p}} = N(E_{\mathbf{p}}) = \frac{1}{e^{\beta(E_{\mathbf{p}} - \mu)} - \xi} \quad (5.61)$$

Wir prüfen, dass $N_{\mathbf{p}}$ die mittlere Besetzungszahl des Impulseigenzustandes $|\mathbf{p}\rangle$ ist (diskret da im Volumen V eingeschränkt),

$$N_{\mathbf{p}} = \langle \hat{N}_{\mathbf{p}} \rangle = \text{Tr}[\hat{\rho} \hat{N}_{\mathbf{p}}] = \dots = N(E_{\mathbf{p}}), \quad (5.62)$$

wobei $\hat{\rho}$ der großkanonische Dichteoperator ist. Die Energie spaltete also in Einteilchenenergien auf, da es sich um ein ideales Gas handelt.

- Innere Energie

$$U = - \left(\frac{\partial \ln(\mathcal{Z})}{\partial \beta} \right)_{\beta \mu} = - \frac{1}{\mathcal{Z}} \left(\frac{\partial \mathcal{Z}}{\partial \beta} \right) \quad (5.63)$$

da $\lambda_U = \beta$ und $\lambda_N = -\beta\mu$; vgl. (3.77). Mit (5.55) folgt durch entsprechendes Ableiten

$$U = \frac{1}{\mathcal{Z}} \sum_{\{N_{\mathbf{p}}\}} E_{\mathbf{p}} \exp\left(-\beta \sum_{\mathbf{p}} (E_{\mathbf{p}} - \mu) N_{\mathbf{p}}\right) = \langle \hat{H} \rangle = \sum_{\mathbf{p}} E_{\mathbf{p}} N_{\mathbf{p}} \quad (5.64)$$

Bemerkungen:

- Damit $N(E_{\mathbf{p}}) \geq 0$ muss für Bosonen

$$\mu \leq \min(E_{\mathbf{p}}) \quad (5.65)$$

erfüllt sein. Bei Fermionen ist $0 \leq N(E_{\mathbf{p}}) \leq 1$ wegen des Pauli-Prinzips automatisch erfüllt und μ somit beliebig.

- Klassischer Grenzfall: $e^{-\beta(E_{\mathbf{p}}-\mu)} \ll 1$, also hohe Energien $(E_{\mathbf{p}} - \mu) \gg kT$ (nicht einfach hohe Temperatur!). Dann folgen Maxwell-Boltzmann Verteilung,

$$N_{\mathbf{p}}^{\text{MB}} = e^{-\beta(E_{\mathbf{p}}-\mu)} \quad (5.66)$$

und (durch Faktorisieren von (5.55)) das klassische Potential

$$\Phi = -kT \sum_{\mathbf{p}} e^{-\beta(E_{\mathbf{p}}-\mu)}. \quad (5.67)$$

Mit Fugazität $z = \exp(\beta\mu)$ folgt (Kontinuumslimit: $\sum_{\mathbf{p}} \rightarrow (V/h^3) \int d^3\mathbf{p}$, vgl. (3.119))

$$\Phi = -kTz \frac{V}{(2\pi\hbar)^3} \int d^3\mathbf{p} e^{-\beta p^2/2m} = -kTzV/\lambda^3 \quad (5.68)$$

mit $\lambda = h/\sqrt{2\pi mkT}$. Dies ist konsistent mit der bereits berechneten Zustandssumme (2.117).

5.3.1 Zustandsgleichung

Wir werden nun eine Zustandsgleichung der Variablen p, V , und U herleiten. Doch wir starten mit der Frage: Was ist die mittlere Teilchenzahl eines idealen Quantengases? Da $E_{\mathbf{p}}$ unabhängig vom Spin s ist, ist der Kontinuumslimit durch

$$\sum_{\mathbf{p}} \dots \rightarrow g \frac{V}{h^3} \int d^3\mathbf{p} \dots \quad (5.69)$$

gegeben, wobei hier $\int d^3\mathbf{p} \dots \rightarrow 4\pi \int_0^\infty dp p^2 \dots$, und mit Entartungsfaktor $g = 2s + 1$ verwendet wurden. Die Gesamtteilchenzahl ist damit

$$N = \frac{gV}{(2\pi\hbar)^3} \int d^3\mathbf{p} N(E_{\mathbf{p}}) = \frac{gV}{2\pi^2\hbar^3} \int_0^\infty dp p^2 N(E_{|p|}). \quad (5.70)$$

Nun verwenden wir (5.61) und gehen zur Integrationsvariable $\epsilon = p^2/2m$ über, wobei

$$p^2 dp \mapsto m\sqrt{2m\epsilon} d\epsilon \quad (5.71)$$

(dies sieht man aus $d\epsilon = p dp/m$, also $p^2 dp = pm d\epsilon$) und damit

$$N = \frac{gV m^{3/2}}{\sqrt{2} \pi^2 \hbar^3} \int_0^\infty d\epsilon \frac{\epsilon^{1/2}}{e^{\beta(\epsilon-\mu)} - \xi} \quad (5.72)$$

Mit $x = \beta\epsilon$ und $v = V/N$ (inverse Dichte) folgt ($z = e^{\beta\mu}$ ist die Fugazität, und $\lambda = h/\sqrt{2\pi mkT}$ die thermische Wellenlänge)

$$\frac{1}{v} = \frac{g}{\lambda^3} \begin{cases} g_{3/2}(z) & \text{für Bosonen} \\ f_{3/2}(z) & \text{für Fermionen} \end{cases} \quad (5.73)$$

mit verallgemeinerten Riemannschen ζ -Funktionen², oder auch *polylogarithmischen*

²Die übliche Riemannsche ζ -Funktion ist $\nu \mapsto g_\nu(1)$.

Funktionen³ ($x = \beta\epsilon$)

$$g_\nu(z) = \frac{1}{\Gamma(\nu)} \int_0^\infty dx \frac{x^{\nu-1}}{z^{-1}e^x - 1}, \quad (5.74)$$

$$f_\nu(z) = \frac{1}{\Gamma(\nu)} \int_0^\infty dx \frac{x^{\nu-1}}{z^{-1}e^x + 1} \quad (5.75)$$

und Gamma-Funktion mit z.B. $\Gamma(3/2) = \sqrt{\pi}/2$, $\Gamma(1) = 1$ und der Funktionalgleichung

$$\Gamma(z+1) = z\Gamma(z). \quad (5.76)$$

Potenzreihenentwicklung der verallgemeinerten Zetafunktionen ist (ohne Beweis):

$$\left\{ \begin{array}{c} g_\nu(z) \\ f_\nu(z) \end{array} \right\} = \sum_{k=1}^{\infty} \xi^{k+1} \frac{z^k}{k^\nu} \quad (5.77)$$

wobei $\xi = +1$ für Bosonen (g_ν), $\xi = -1$ für Fermionen (f_ν).

Genauso folgt

$$\Phi = -\beta^{-1} \ln(\mathcal{Z}) = -pV = \frac{\xi}{\beta} \sum_p \ln(1 - \xi e^{-\beta(E_p - \mu)}) \quad (5.78)$$

Im Kontinuumslimites ergibt das

$$\Phi = \xi \frac{gV m^{3/2}}{\sqrt{2} \pi^2 \hbar^3 \beta} \int_0^\infty d\epsilon \sqrt{\epsilon} \ln[1 - \xi \exp(-\beta(\epsilon - \mu))] \quad (5.79)$$

$$\begin{aligned} &= -\frac{2}{3} \frac{gV m^{3/2}}{\sqrt{2} \pi^2 \hbar^3} \int_0^\infty d\epsilon \frac{\epsilon^{3/2}}{e^{\beta(\epsilon - \mu)} - \xi} \\ &= \frac{-gV kT}{\lambda^3} \begin{cases} g_{5/2}(z) & \text{(Bosonen),} \\ f_{5/2}(z) & \text{(Fermionen),} \end{cases} \end{aligned} \quad (5.80)$$

wobei im vorletzten Schritt partielle Integration verwendet wurde, mit Hilfsrelationen

$$\sqrt{\epsilon} = \frac{2}{3} \partial_\epsilon \epsilon^{3/2} \quad (5.81)$$

$$\partial_\epsilon \ln(1 - \xi e^{-\beta(\epsilon - \mu)}) = \frac{\xi \beta}{e^{\beta(\epsilon - \mu)} - \xi} \quad (5.82)$$

d.h. da Randterm verschwindet, ergibt partielle Integration

$$\int_0^\infty d\epsilon \sqrt{\epsilon} \ln(1 - \xi e^{-\beta(\epsilon - \mu)}) = -\frac{2}{3} \int_0^\infty d\epsilon \epsilon^{3/2} \underbrace{\partial_\epsilon \ln(\cdots)}_{\frac{\xi \beta}{e^{\beta(\epsilon - \mu)} - \xi}}; \quad (5.83)$$

beachte auch $\Gamma(5/2) = (3/2)\Gamma(3/2) = (3/4)\sqrt{\pi}$.

Ebenso folgt die innere Energie

$$U = \frac{gV}{(2\pi\hbar)^3} \int d^3p E_p N(E_p) = \frac{gV m^{3/2}}{\sqrt{2} \pi^2 \hbar^3} \int_0^\infty d\epsilon \frac{\epsilon^{3/2}}{e^{\beta(\epsilon - \mu)} - \xi} = -\frac{3}{2} \Phi = +\frac{3}{2} pV \quad (5.84)$$

Insgesamt folgt wie beim klassischen idealen Gas die die *Zustandsgleichung eines idealen Quantengases*

$$pV = \frac{2}{3} U \quad (5.85)$$

unabhängig von der Statistik. Beim klassischen idealgas ist auch $pV = NkT = 2U/3$

³Man schreibt auch $g_\nu(z) \equiv \text{Li}_\nu(z)$ für die polylogarithmische Funktion Li.

wegen $U = (3/2)NkT$. Vorsicht: $U = (3/2)NkT$ gilt nur bei klassischem Idealgas, $pV = 2U/3$ aber allgemein bei quadratischer Dispersionsrelation (also massiven nicht-relativistischen Teilchen). Bei Photonen ist die Zustandsgleichung eine andere. Im folgenden Abschnitt werden wir sehen, dass die Zustandsgleichung (5.73) für Bosonen einen zusätzlichen Term erhält, was die folgende Gleichung liefert.

Zustandsgleichung eines idealen Bose-Gases

$$\frac{N}{V} = \frac{1}{\lambda^3} g_{3/2}(z) + \frac{1}{V} \frac{z}{1-z} \quad (5.86)$$

27. VL

In dieser VL wurden die idealen Quantengase ausführlich wiederholt.

5.3.2 Klassischer Grenzfall

Im klassischen Grenzfall ist $E - \mu \gg kT$. Bei niedrigen Energien E folgt $\mu \ll -kT$, also ist $z = \exp(\mu/kT) \ll 1$. Damit folgt aus (5.73) mit (5.77), dass

$$\frac{\lambda^3}{gv} = \sum_{k=1}^{\infty} \xi^{k+1} z^k k^{-3/2} = z + \xi z^2 / \sqrt{8} + O(z^3). \quad (5.87)$$

Lösen der quadratischen Gleichung und Entwickeln 2. Ordnung in $\lambda^3/gv \ll 1$ ergibt

$$z \approx \frac{\lambda^3}{vg} - \xi \frac{1}{\sqrt{8}} \left(\frac{\lambda^3}{vg} \right)^2. \quad (5.88)$$

Damit folgt mit (5.77) aus (5.80), dass

$$\begin{aligned} \Phi &= \frac{-gVkT}{\lambda^3} (z + \xi z^2 / \sqrt{32} + \dots) \\ &= -\frac{gVkT}{\lambda^3} \left(\frac{\lambda^3}{vg} - \xi \frac{1}{\sqrt{8}} \left(\frac{\lambda^3}{vg} \right)^2 + \xi \frac{1}{\sqrt{32}} \left(\frac{\lambda^3}{vg} \right)^2 + \dots \right), \end{aligned} \quad (5.89)$$

also die Zustandsgleichung

$$pV = -\Phi = NkT \left(1 - \xi \frac{\lambda^3}{\sqrt{32}vg} + \dots \right) = \frac{2}{3}U \quad (5.90)$$

vgl. hierzu das klassische ideale Gas. Korrekturen skalieren wie $\hbar^3/T^{3/2}$, da $\lambda = h/\sqrt{2\pi mkT}$.

Führende Quantenkorrekturen (Austauschkorrekturen):

- Sie sind proportional zu \hbar^3 , hängen vom Symmetrieverhalten ($\xi = \pm$) der Wellenfunktion ab.
- Für Bosonen (symmetrische Zustandsvektoren) ergibt sich eine Verringerung des Drucks, als ob sich Atome anziehen (vgl. van der Waals Gleichung!). Es gibt eine Tendenz zur *Clusterbildung bei Bosonen*, obwohl es sich um nichtwechselwirkende Teilchen handelt.
- Für Fermionen (antisymmetrische Zustandsvektoren) ergibt sich eine Erhöhung des Drucks, d.h. eine effektive Abstoßung. Das *Pauliprinzip wirkt wie eine abstossende Kraft*, obwohl es sich um nichtwechselwirkende Teilchen handelt.

Dies ist ein Effekt der Quantenstatistik.

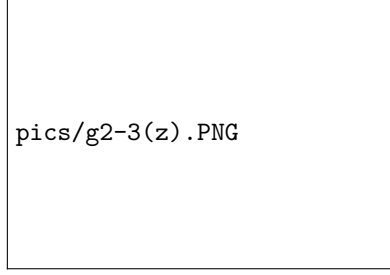


Abbildung 5.1: Graph der Funktion $g_{3/2}$; es gilt $g_{3/2}(1) = \zeta(3/2) \approx 2.612$ (mit ζ -Funktion ζ).

- Der klassische Grenzfall gilt, falls $\lambda \ll v^{1/3}$ (typischer Abstand zweier Teilchen ist $v^{1/3}$, da inverse Dichte gerade v ist), also bei starker Verdünnung und/oder hoher Temperatur.
- Mittlere Besetzungszahl folgt aus Fermi-Dirac- bzw. Bose-Einstein-Verteilung. Im klassischen Grenzfall ist

$$\frac{1}{z^{-1}e^{\beta E} \pm 1} \rightarrow e^{-\beta(E-\mu)}, \quad (5.91)$$

da dann $-\beta\mu \gg 1$ und somit $z = e^{\beta\mu} \ll 1$ ist. Dieser Grenzfall ergibt also genau die *Maxwell-Boltzmann Verteilung*

$$N^{\text{MB}}(E) = e^{-\beta(E-\mu)}. \quad (5.92)$$

Es gibt insbesondere keinen Unterschied in der Verteilung für Bosonen und Fermionen im klassischen Limes.

5.3.3 Ideale Bosegase

Im Folgenden wollen wir in Abhängigkeit der Dispersionsrelation $p \mapsto E_p$ einige physikalisch relevante Spezialfälle betrachten.

5.3.3.1 Bose-Einstein-Kondensation

Wir betrachten ein ideales Gas aus N spinlosen ($g = 2s + 1 = 2$) Bosonen mit Dispersionsrelation

$$E_{\mathbf{p}} = \frac{\mathbf{p}^2}{2m}. \quad (5.93)$$

Im Grundzustand befinden sich alle Bosonen im Einteilchenzustand mit $\mathbf{p} = 0$.

Bose-Einstein-Kondensation bezeichnet das Phänomen, dass unterhalb einer gewissen kritischen Temperatur T_c der Einteilchenzustand mit $\mathbf{p} = 0$ *makroskopisch* besetzt ist, also mit einer Teilchenzahl, die von der Größenordnung der Gesamtteilchenzahl N ist. Hier beobachtet man eine Abspaltung in Kondensat und angeregte „normale“ thermische Bosonen. Dies wurde von Einstein und Bose 1924 vorhergesagt und 1995 von Ketterle, Cornell, und Wieman beobachtet (Nobelpreis 2001).

Für Bosonen ist das inverse spezifische Volumen aus (5.73) als

$$\frac{\lambda^3}{v} = g_{3/2}(z) \quad (\text{Zustandsgleichung}) \quad (5.94)$$

gegeben. Für Bosonen muss $\mu \leq 0$ sein, also $0 < z \leq 1$. Damit ist der Maximalwert von $g_{3/2}(z)$ bei $z = 1$ gegeben:

$$g_{3/2}(1) = \zeta(3/2) \approx 2.612, \quad (5.95)$$

s. Abbildung 5.1.

Im Folgenden seien N, V und damit auch $v = V/N$ konstant. Damit liefert (5.94) das chemische Potential μ bzw. die Fugazität $z = e^{\beta\mu}$. Bei einer Verringerung von T werden sowohl λ^3/v als auch $z = \exp(-|\mu|/(kT))$ grösser (s. Abbildung 5.1). Dies kann nur bis $z = 1$ so sein, d.h. es muss eine kritische Temperatur T_c geben, die durch $\lambda^3 = v g_{3/2}(1)$ bestimmt und unterhalb derer $z = 1$ ist.

28. VL

Aus der Existenz der kritischen Temperatur T_c bei $z = 1$ folgt aus (5.94), dass

$$kT_c = \frac{2\pi\hbar^2/m}{(2.612v)^{2/3}}. \quad (5.96)$$

Je höher Dichte ist (also je kleiner spezifisches Volumen v) und je leichter die Atome sind, desto höher ist T_c . Unterhalb von T_c existiert ein Kondensat zusätzlich zum thermischen Bosegas. Im Extremfall bei $T = 0$ gäbe es nur noch das Kondensat.

Bei T_c ist z und damit auch das chemische Potential μ nicht stetig differenzierbar. Andererseits ist $\mu = \left(\frac{\partial G}{\partial N}\right)_{T,p}$, d.h. eine zweite Ableitung der Gibbs-Energie hat eine Unstetigkeit. Damit handelt es sich beim dem Kondensationseffekt um einen *Phasenübergang 2. Ordnung*.

Bei Temperaturen, die nicht deutlich größer als T_c sind, muss Grenzübergang $d^3p \mapsto d\epsilon$ sorgfältiger durchgeführt werden. Wir behandeln die $\mathbf{p} = 0$ Mode in (5.60) gesondert, wobei unterhalb T_c das chemische Potential $\mu = 0$, als $z = 1$. Die Gesamtteilchenzahl ist

$$\begin{aligned} N &= N(\mathbf{p} = 0) + \sum_{\mathbf{p} \neq 0} N(E_{\mathbf{p}}) \\ &\approx N_0 + \frac{V}{h^3} \int d^3\mathbf{p} N(E_{\mathbf{p}}) \\ &= N_0 + N \frac{v}{\lambda^3} g_{3/2}(z), \end{aligned} \quad (5.97)$$

wobei die Anzahl der kondensierten Atome ist und wir das Integral bereits in (5.73) berechnet hatten. Es gilt

$$N_0 = N(\mathbf{p} = 0) = \frac{1}{e^{\beta(0-\mu)} - 1} = \frac{1}{1/z - 1}, \quad (5.98)$$

was auch (5.86) impliziert.

Bemerkungen:

- Schon die nächstliegende Mode mit $\mathbf{p} \neq 0$ hat eine um Grössenordnungen kleinere Besetzungswahrscheinlichkeit. Bei $T = 0$ ist dieses Verhalten extremal.
- Mit der Definition der kritischen Temperatur folgt bei beliebigem T :

$$N = N_0 + N(T/T_c)^{3/2} \frac{g_{3/2}(z)}{g_{3/2}(1)} \quad (5.99)$$

Für $T > T_c$ ist $z < 1$, und N_0 ist von Ordnung 1, kann daher gegenüber dem 2. Term vernachlässigt werden. Damit folgt $g_{3/2}(z) = g_{3/2}(1)(T/T_c)^{-3/2}$, was uns $\mu(T)$ liefert.

- Für $T < T_c$ ist $z = 1$, d.h. $\mu(T < T_c) = 0$, und

$$N = N_0 + N(T/T_c)^{3/2}. \quad (5.100)$$

Der Anteil des Kondensats ist also bei grossen N

$$\nu_0 := \frac{N_0}{N} = 1 - (T/T_c)^{3/2} \quad (5.101)$$

(s. VL oder [2, Abbildung 4.13] für eine Skizze).

- Spezifische Wärme des Bosegases:
Bei tiefen T ist $C_V \propto T^{3/2}$ (längere Rechnung über $S = -\left(\frac{\partial(pV)}{\partial T}\right)_{V,\mu}$). Bei $T = T_c$ gibt es jedoch einen scharfen Peak ('cusp'), was eine Signatur des Phasenübergangs ist; s. VL oder [2, Abbildung 4.17] für eine Skizze.

Physikalische Anwendungen des idealen spinlosen Bosegases:

- Helium-4. Bei $T_\lambda = 2.18$ K (Lambda-Temperatur) Phasenübergang in suprafluiden Zustand. Wechselwirkung hier aber sehr wichtig, sonst wäre es „einfache“ Bose-Einstein-Kondensation. Dadurch starke Entleerung der makroskopisch besetzten Mode („quantum depletion“) auf wenige Prozent.
- Atomarer Wasserstoff in starkem Magnetfeld (Spinpolarisation um Rekombination zu H_2 zu verhindern). Hier wurde Bose-Einstein-Kondensation 1998 nachgewiesen ($T_c = 50\mu K$).
- Experimentell erster Nachweis von Bose-Einstein-Kondensaten in ultrakalten Atomgasen (Rb, Li) in Fallen (1995). Metastabil (Festkörper ist eigentlich stabile Form bei ultratiefen Temperaturen), aber lange Lebensdauer da sehr verdünnt, wenige Dreiteilchenstösse. T_c ist hier im Nano-Kelvin-Bereich.

5.3.3.2 Photonengas

Die Dispersionsrelation für masselose Teilchen ist

$$E_{\mathbf{p}} = c|\mathbf{p}| = c\hbar|\mathbf{k}| \quad (5.102)$$

mit Wellenvektor \mathbf{k} . Photonen haben einen Spin $s = 1$. Die Entartung der Energien ist jedoch $g = 2$ und nicht $g = (2s+1) = 3$, da die Polarisation $\sigma = \pm$ wegen Ausbreitung mit Lichtgeschwindigkeit senkrecht auf \mathbf{k} stehen muss. Der Hamiltonoperator ist daher

$$\hat{H} = \sum_{\mathbf{p},\sigma} E_{\mathbf{p}} \hat{N}(\mathbf{p},\sigma), \quad \mathbf{p} \neq 0, \quad (5.103)$$

wobei $\hat{N}(\mathbf{p},\sigma)$ der Besetzungszahloperator zum Impuls \mathbf{p} unter der Polarization σ ist. Die kanonische Zustandssumme ist

$$\begin{aligned} Z &= \text{Tr}[e^{-\beta\hat{H}}] = \sum_{\{N(\mathbf{p},\pm)\}} e^{-\beta \sum_{\mathbf{p},\sigma} E_{\mathbf{p}} N(\mathbf{p},\sigma)} \\ &= \left(\prod_{\mathbf{p} \neq 0} \frac{1}{1 - e^{-\beta c|\mathbf{p}|}} \right)^2, \end{aligned} \quad (5.104)$$

wobei der Exponent gerade $g = 2$ ist. Unter Verwendung des Kontinuumslimes (5.69) und anschließendes Integrierenlassen sich damit die folgenden thermodynamischen

Gleichungen herleiten:

$$F = -kT \ln(\ln Z) = \frac{4\sigma}{3c} VT^4 \quad (\text{freie Energie}), \quad (5.105)$$

$$S = -\left(\frac{\partial F}{\partial T}\right)_V = \frac{16\sigma}{3c} VT^3 \quad (\text{Entropie}), \quad (5.106)$$

$$C_V = T\left(\frac{\partial S}{\partial T}\right)_V = \frac{16\sigma}{c} VT^3 \quad (\text{Wärmekapazität}), \quad (5.107)$$

$$U = F + TS = \frac{4\sigma}{c} VT^4 \quad (\text{innere Energie}), \quad (5.108)$$

$$p = -\left(\frac{\partial F}{\partial V}\right)_T = \frac{4\sigma}{3c} T^4 \quad (\text{Strahlungsdruck}), \quad (5.109)$$

wobei σ die *Stefan-Boltzmann-Konstante*

$$\sigma = \frac{2\pi^5 k_B^4}{15h^3 c^2} = \frac{\pi^2 k_B^4}{60h^3 c^2} \quad (5.110)$$

ist. Für die innere Energie s. auch Übungsaufgabe 8.2 (Stefan-Boltzmann-Gesetz). Wir beobachten, dass aufgrund der relativistischen Dispersionsrelation

$$U = 3pV \quad (5.111)$$

anstelle von (5.85) gilt.

Mithilfe der Bose-Einstein-Verteilung (5.61) lässt sich auch die *spektrale Energiedichte* (Energie pro Volumen- und Frequenzeinheit)

$$u(\omega) = \frac{\hbar}{\pi^2 c^3} \frac{\omega^3}{e^{\hbar\omega/(kT)} - 1} \quad (5.112)$$

berechnen, was im klassischen Grenzfall $kT \gg \hbar\omega$ dem Rayleigh-Jeans-Gesetz und im Quantengrenzfall $kT \ll \hbar\omega$ dem Wien'schen Gesetz entspricht, s. die VL oder [2, Abbildung 4.23] für eine Skizze.

5.3.3.3 Phononengas

Beim „Phononengas“ handelt es sich um nicht wechselwirkende quantisierte Gitterschwingungen in Festkörpern. Hier ist die Dispersionsrelation ebenfalls linear. Allerdings gibt es zwei transversale plus eine longitudinale Schwingungsmoden.

5.3.4 Ideales Fermigas

5.3.4.1 Physikalische Beispiele

Reale Fermionsysteme sind wechselwirkend, oft spielen Coulomb-Wechselwirkungen geringe oder keine Rolle. Theoretische Basis: Landaus *Fermiflüssigkeitstheorie*, würde hier zu weit führen. Physikalischer Mechanismus zum „Ausschalten“ der Wechselwirkung

- Abschirmung: Elektronen bilden Polarisationswolken um Ladungen. Langreichweitiger Schwanz des Coulombpotentials wird abgeschirmt, (s. Elektrodynamik).
- Wir werden sehen, dass das Pauliprinzip impliziert, dass es nur wenige Zustände gibt, die zur Elektron-Elektron Streuung beitragen (Phasenraumargument).

Beim idealen Fermigas wird wieder von nichtwechselwirkenden Teilchen ausgegangen. Das ideale Fermigas ist in guter Näherung anwendbar auf

- Elektronengas in vielen Metallen, z.B. Alkalimetalle, Edelmetalle

- Helium-3 (dort auch Wechselwirkungseffekte im Rahmen der Fermi-Flüssigkeitstheorie wichtig)
- Kernmaterie: Neutronen und Protonen sind Fermionen. Auch in Neutronensternen.
- Weiße Zwerge (am Ende eines Sternenlebens), z.B. Sirius B, bestehen aus ionisierten Kernen und freien Elektronen (Plasma).

Das ideale Fermigas ist nicht anwendbar auf Korrelationseffekte wie Magnetismus oder Supraleitung.

5.3.4.2 Grundzustand (Fermi-See)

Wir betrachten den Grundzustand ($T = 0$) eines N -Fermionensystems mit Spin und haben dabei den Spin $s = 1/2$ im Hinterkopf (z.B. Elektronen). Wegen des Pauliprinzipes ist jeder Impulseigenzustand $|p\rangle$ entweder nicht oder nur einfach besetzt. Jeder Wert von \mathbf{p} kommt z.B. bei Elektronen genau 2-fach vor ($g = 2$).

Im Grundzustand sind alle Impulseigenzustände unterhalb von $p_F = \hbar k_F$ besetzt, so dass die Teilchendichte N/V realisiert ist. Mann nennt p_F *Fermiimpuls*,

$$E_F = \frac{\hbar^2 k_F^2}{2m} = kT_F \quad (5.113)$$

Fermienergie bzw. *Fermitemperatur*, mit *Fermi-Wellenzahl* k_F .

Eine 2D-Skizze im Impulsraum ist hier hilfreich, um zu visualisieren, welche Zustände besetzt sind (Fermikugel), s. VL oder [2, Abbildung 4.2].

Beim sogenannten *entarteten Fermigas* bedeutet *entartet* eine Abweichung vom klassischen idealen Gas.

Im Grundzustand sind also alle Einteilchenzustände mit einer Energie kleiner als die Fermienergie E_F bzw. Impuls $p := |\mathbf{p}|$ kleiner als p_F besetzt (Pauliprinzip!). Damit ist die Teilchenzahl im Kontinuumslimites (5.69)

$$\begin{aligned} N &= g \sum_{p \leq p_F} 1 = \frac{gV}{h^3} \int_0^\infty \Theta(p_F - p) 4\pi p^2 dp \quad (\text{Kugelkoordinaten}) \\ &= \frac{gV p_F^3}{6\pi^2 \hbar^3}, \end{aligned} \quad (5.114)$$

wobei wir $h = 2\pi\hbar$ verwendet haben. Damit ist

$$p_F = \left(\frac{6\pi^2}{g} \right)^{1/3} \hbar \varrho^{1/3} \quad (5.115)$$

mit Teilchendichte $\varrho := N/V$. Weiterhin ist die Fermienergie

$$E_F = \frac{p_F^2}{2m} = \left(\frac{6\pi^2}{g} \right)^{2/3} \frac{\hbar^2}{2m} \varrho^{2/3}. \quad (5.116)$$

Die Grundzustandsenergie ist

$$U = \frac{gV}{h^3} \int_0^\infty \frac{p^2}{2m} \Theta(p_F - p) 4\pi p^2 dp = \frac{gV p_F^5}{20\pi^2 \hbar^3 m}, \quad (5.117)$$

also

$$U = \frac{3}{5} E_F N \quad (5.118)$$

Die Zustandsgleichung (5.85) liefert den Druck

$$p = \frac{2}{3} \frac{U}{V} = \frac{2}{5} E_F \varrho = \frac{1}{5} \left(\frac{6\pi^2}{g} \right)^{2/3} \frac{\hbar^2}{m} \varrho^{5/3}. \quad (5.119)$$

Die Entartung des Grundzustands ist so gering, dass dessen Entropie makroskopisch verschwindet (vgl. Abschnitt 4.3.10). Die Euler-Gleichung (4.2) liefert damit

$$\mu = \frac{1}{N} (U + pV - TS) = \dots = E_F, \quad (5.120)$$

was mit der Definition des chemischen Potentials konsistent ist (es ist die Energie, die notwendig ist, um ein weiteres Teilchen in das Ensemble einzubringen). Man beachte, dass die Existenz der Fermienergie jedoch in der Quantenmechanik begründet ist.

5.3.4.3 Endliche Temperaturen

Nun betrachten wir $T \ll T_F$. Hier führt die Fermi-Dirac-Verteilung

$$N(E) = \frac{1}{e^{(E-\mu)/kT} + 1} \quad (5.121)$$

zum einem „Versmieren“ der *Fermikante* um $\pm kT$, s. Skizze in VL oder [2, Abbildung 4.3]. Für $T \rightarrow 0$ wird aus der Kante eine Stufe („entartetes Fermigas“). Man kann ausrechnen [2, Abschnitt 4.3], dass der Effekte der Versmierung durch die folgende Korrektur zur Energie

$$U = U(T=0) + C \left(\frac{k_B T}{E_F} \right)^2 + O(T^4) \quad (5.122)$$

mit einer absoluten Konstante C gegeben ist. Für kleine Temperaturen ist die Wärmekapazität also linear in T ,

$$C_V \propto T/T_F, \quad (5.123)$$

meist also klein, da T_F für Metalle sehr groß ist. Allgemeiner sind thermische Effekte in Fermigasen meist klein. Analog kann man unter Verwendung von $N(T) = N(T=0) = N(\epsilon_F)$ zeigen, dass

$$\mu(T) = E_F \left(1 - \frac{\pi^2}{12} \left(\frac{kT}{E_F} \right)^2 + O(T^4) \right) \quad (5.124)$$

ist.

29. VL

5.4 Das Ising-Modell

Das Ising-Modell ist ein Spezialfall von *Quantenspinningittersystemen*, die z.B. noch das Heisenbergmodell umfassen. Diese Modelle bilden die Grundlage z.B. für das Verständnis von Magnetismus. Eine Übersicht über die mathematische Behandlung des Ising-Modells kann in [25] gefunden werden.

Wir betrachten ein System aus N (klassischen) Spins, jeweils mit magnetischem Moment $\mu_0 = \mu_0 s_i$ mit $s_i = \pm 1$ auf einem kubischen Gitter, siehe Abbildung 5.2.

- Spins können hier nur nach oben oder unten zeigen, was also stark anisotrope Version des SU(2)-invarianten quantenmechanischen angesehen werden kann.

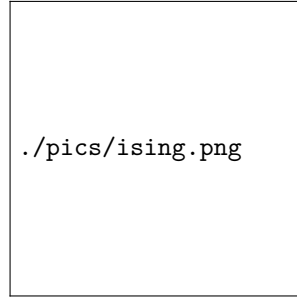


Abbildung 5.2: Ising-Modell.

- Die quantenmechanischen Heisenbergmodelle mit $SU(2)$ -Symmetrie sind deutlich komplizierter zu analysieren. Als Quantenmodell geschrieben ist das Ising-Modell diagonal in der σ_z -Basis.
- In Festkörpern sind die Wechselwirkungen zwischen Spins meist kurzreichweitig.

Das *Ising-Modell* ist ein grundlegendes Modell der Statistischen Physik, welches Magnetismus erklären kann. Es ist ein klassisches Modell mit Hamiltonfunktion

$$H(s) = -J \sum_{\langle i,j \rangle} s_i s_j - \mu_0 B \sum_{i=1}^N s_i, \quad (5.125)$$

wobei $s \in \{-1, 1\}^N$, B ist die äußere Magnetfeldstärke in z -Richtung, J ist eine Kopplungskonstante und $\langle i, j \rangle$ bezeichnet die nächsten Nachbarn auf dem vorgegebenen Wechselwirkungsgraphen, der hier durch ein kubische Gitter in D räumlichen Dimensionen ist (s. Abbildung 5.2 für $D = 2$).

- Falls $J > 0$, ist es energetisch günstig, die Spins in die gleiche Richtung auszurichten: Tendenz zum *Ferromagnetismus*.
- Falls $J < 0$, gibt es entsprechend eine Tendenz zum *Antiferromagnetismus*.
- Ein äußeres Magnetfeld $B > 0$ versucht Spins nach oben auszurichten.
- Wir betrachten $J > 0$. Als Funktion der Temperatur sollte bei der *Curietemperatur* $T = T_c$ ein Phasenübergang zwischen geordneter ferromagnetischer Phase ($T < T_c$) und ungeordneter paramagnetischer Phase ($T > T_c$) auftreten.
- Ferromagnetische Phase hat eine *spontane Magnetisierung*. Kooperative Prozesse setzen ein, die alle Spins in die selbe Richtung zeigen lassen, obwohl die Hamiltonfunktion H nur kurzreichweitig Wechselwirkung enthält. Das bedeutet, dass aus Regeln für das Verhalten nächster Nachbarn *langreichweitige Korrelationen* und *langreichweitige Ordnung* entsteht!

Die *Magnetisierung* ist die Observable

$$M(s) := \mu_0 \sum_{i=1}^N s_i. \quad (5.126)$$

Eine *spontane Magnetisierung* liegt vor, falls

$$\langle M \rangle = \mu_0 \sum_{i=1}^N \langle s_i \rangle \quad (5.127)$$

für $B = 0$ einen makroskopisch nichtverschwindenden Wert hat, siehe Abbildung 5.3 für die spontane Magnetisierung im 2D-Ising-Modell. Häufig schreiben wir auch

$$\langle M \rangle = \mu_0(N_+ - N_-), \quad (5.128)$$

wobei N_{\pm} die erwartete Anzahl der Spins s_i mit Wert ± 1 ist.

Beide Richtungen (Vorzeichen) von M sind bei $B = 0$ möglich, eine davon wird realisiert. Dies nennt man *Spontane Symmetriebrechung*:

- Hamiltonfunktion besitzt volle Symmetrie unter $s_i \rightarrow -s_i$ (für alle Spins gleichzeitig). Dies gilt jedoch nicht für den Grundzustand und die Magnetisierung M .
- Spontane Symmetriebrechung ist ein sehr wichtiges Konzept, z.B. bei der Beschreibung der schwachen Wechselwirkung in der Teilchenphysik, dem Meissnereffekt in der Supraleitung u.v.m.
- Das realisierte Vorzeichen von M hängt von der Vorgeschichte ab.

5.4.1 Zustandssumme

Implizit haben wir U als Makronebenbedingung und N und B als Mikronebenbedingung gewählt. Die Zustandssumme ist

$$\begin{aligned} Z(\beta; B, N) &= \sum_{s \in \{-1, 1\}^N} e^{-\beta H(s)} \\ &= \sum_{s_1 = \pm} \dots \sum_{s_N = \pm} e^{\beta \mu_0 B \sum_{i=1}^N s_i} e^{\beta J \sum_{\langle ij \rangle} s_i s_j}. \end{aligned} \quad (5.129)$$

Das Magnetfeld entspricht dem chemischen Potential, welches die Magnetisierung $\langle M \rangle = N_+ - N_-$ steuert.

5.4.2 Das 1D-Ising-Modell

In einer Dimension (Ising-Kette) lässt sich die Zustandssumme in elementarer Weise exakt ausrechnen. Für die Rechnung ist es hilfreich $\mu_0 = 1$ zu setzen. Die Wechselwirkung ist die Summe über einen Index, d.h.

$$H(s) = - \sum_{i=1}^N (J s_i s_{i+1} + B s_i), \quad (5.130)$$

wobei wir periodische Randbedingungen $s_{N+1} = s_1$ wählen. Dann lässt sich die *Transfermatrixmethode* anwenden: Wir werden gleich sehen, dass die Zustandssumme die

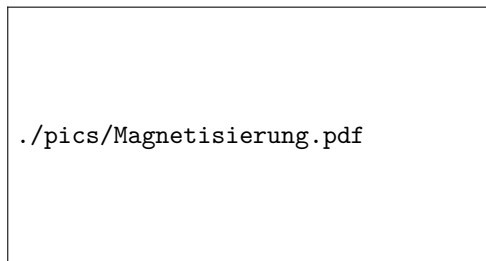


Abbildung 5.3: Spontanmagnetisierung

Spur eines Produktes von N identischen 2×2 Matrizen

$$A = \begin{pmatrix} e^{\beta(J+B)} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J-B)} \end{pmatrix} \quad (5.131)$$

mit Einträgen $A(s_i, s_{i+1})$ ist. Die Matrix A heißt *Transfermatrix*, da sie einen Platz weiter entlang der Isingkette propagiert. Die Zustandssumme ist damit

$$\begin{aligned} Z &= \sum_{s_1=\pm} \dots \sum_{s_N=\pm} \prod_{i=1}^N e^{\beta(Js_i s_{i+1} + Bs_i)} \\ &= \sum_{s_1=\pm} \dots \sum_{s_N=\pm} \prod_{i=1}^N e^{\beta(\frac{B}{2}s_i + Js_i s_{i+1} + \frac{B}{2}s_{i+1})} \\ &= \text{Tr} \left[\begin{pmatrix} e^{\beta(J+B)} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J-B)} \end{pmatrix}^N \right] \\ &= \text{Tr}[A^N] \\ &= \lambda_+^N + \lambda_-^N \end{aligned} \quad (5.132)$$

mit den Eigenwerten λ_{\pm} der Matrix A . Jeder Eigenwert λ von A muss $\det(A - \lambda \mathbb{1}) = 0$ und damit

$$\lambda^2 - 2 \cosh(\beta B) e^{\beta J} \lambda + e^{2\beta J} - e^{-2\beta J} = 0 \quad (5.133)$$

erfüllen. Damit ist

$$\lambda_{\pm} = e^{\beta J} \left(\cosh(\beta B) \pm \sqrt{\cosh^2(\beta B) - 1 + e^{-4\beta J}} \right) \quad (5.134)$$

Für $N \rightarrow \infty$ dominiert der grössere Eigenwert die Zustandssumme (5.132). Damit folgt für große N

$$\ln Z(T, B, N) \approx N \ln(\lambda_+) \quad (5.135)$$

bzw. $F \approx -kTN \ln(\lambda_+)$. Für $B = 0$ folgt

$$Z = [2 \cosh(\beta J)]^N. \quad (5.136)$$

Die Magnetisierung erhalten wir als Ableitung nach B (was wieder aus der Definition der Zustandssumme folgt),

$$\langle M \rangle =: M = kT \frac{1}{Z} \frac{\partial}{\partial B} Z(T, B, N)|_{B=0} = kT \left. \frac{\partial}{\partial B} \ln(Z(T, B, N)) \right|_{B=0}. \quad (5.137)$$

Für $B \rightarrow 0, T > 0$ ist Z bzw. $\ln(Z)$ eine analytische Funktion von B^2 (Taylorentwicklung in B^2 existiert). Damit folgt sofort

$$M(B = 0) = 0 \quad (5.138)$$

In einer Dimension gibt es bei endlicher Temperatur keine spontane Magnetisierung.

Dies ist ein Spezialfall des *Mermin-Wagner Theorems*. In diskreten Systemen (wie Spinsysteme) zerstören Fluktuationen jegliche Ordnung für $T > 0$ in einer Raumdimension (völlig allgemein, nicht auf diese Art von Modell beschränkt).

Physikalische Erklärung: Wir betrachten die Spinkonfiguration wo die ersten $N/2$ Spins den Wert $s_i = +$ und die anderen Spins den Wert $s_j = -$ haben. Voll ausgerichteter Ferromagnet hätte Energie $-NJ$. Dieser Zustand hat jedoch Energie $H(s) = (-N + 4)J$, und Magnetisierung 0 („ungeordnet“). Hier haben wir zwei

Domänenwände (dort ändern sich die Spinausrichtungen), mit jeweils Energieverlust von $2J$. Für grosse N ist die relative Energieänderung also nur von Ordnung $1/N$, d.h. kleine Energiefluktuationen zerstören bereits die magnetische Ordnung.

5.4.3 2D Isingmodell

In zwei Dimensionen sind Domänenwände Linien, und kosten relative Energie $\propto N^{-1/2}$. Hier gibt es eine kritische Temperatur T_c , so dass für $T < T_c$ eine ferromagnetische Phase gegeben ist (ebenso in 3D). Dies ist aber höchst nichttrivial, und wurde erst durch die exakte Lösung des 2D Isingmodells durch Lars Onsager (1944) gezeigt. Die Herleitung würde zu weit führen, das Resultat ist mit $\beta = 1/(kT)$ für $B = 0$

$$Z = [2 \cosh(\beta J) e^I]^N = e^{-\beta F} \quad (5.139)$$

mit

$$I = \int_0^\pi \frac{d\phi}{2\pi} \ln \left(\frac{1}{2} [1 + \sqrt{1 - \kappa^2 \sin^2 \phi}] \right) \quad (5.140)$$

und

$$\kappa = 2 \frac{\sinh(2\beta J)}{\cosh^2(2\beta J)}. \quad (5.141)$$

Die freie Energie F hat für $\kappa > 1$ eine *Verzweigungsschnitt* (Uneindeutigkeit des Logarithmus), und damit auch tatsächlich eine spontane Magnetisierung. Die kritische Temperatur T_c , auch *Curie-Temperatur* genannt, erfüllt

$$\kappa_c = \frac{2 \sinh(2J/kT_c)}{1 + \sinh^2(2J/kT_c)} = 1, \quad (5.142)$$

was impliziert, dass

$$\sinh(2J/kT_c) = 1 \rightarrow \frac{kT_c}{J} = \frac{2}{\ln(1 + \sqrt{2})} \approx 2.269. \quad (5.143)$$

Eine spontane Magnetisierung $M(T)$ nahe $T = T_c$ (aber unterhalb) folgt dann in der Form

$$\frac{M}{N} \propto (T_c - T)^\eta \quad (5.144)$$

mit *kritischem Exponenten* $\eta = 1/8$. Die Wärmekapazität C_V folgt aus der exakten Lösung in der Form (nahe T_c)

$$\frac{C_V}{N} = \frac{k_B \beta^2}{N} \frac{\partial^2 \ln Z(\beta)}{\partial \beta^2} \approx \frac{8k}{\pi} (\beta J)^2 \ln |1/(T - T_c)|. \quad (5.145)$$

d.h. die Wärmekapazität divergiert logarithmisch.

Die kanonische Verteilung ist hier explizit durch den Wahrscheinlichkeitsvektor p mit Komponenten

$$p_s = \frac{e^{-\beta H(s)}}{Z} \quad (5.146)$$

gegeben. Die *Korrelation* zwischen zwei Spins $i, j \in [N]$ ist

$$\Gamma(i, j) := \langle (S_i - \langle S_i \rangle) (S_j - \langle S_j \rangle) \rangle \quad (5.147)$$

$$= \langle S_i S_j \rangle - \langle S_i \rangle \langle S_j \rangle, \quad (5.148)$$

wobei alle Erwartungswerte bzgl. p genommen werden und S_i die Zufallsvariable mit Werten in $\{-1, +1\}$ ist, die den i -ten Spin beschreibt und hier thermisch verteilt ist, d.h. der Zufallsvektor S ist gemäss des Wahrscheinlichkeitsvektors p ver-

teilt. Onsager hat auch gezeigt, dass die Korrelationen wie

$$\Gamma_p(i, j) \leq c_\beta e^{-\beta|i-j|/\xi(\beta)} \quad (5.149)$$

abfallen mit Korrelationslänge

$$\xi \propto \frac{1}{|T - T_c|^\nu} \quad (5.150)$$

und kritischem Exponenten $\nu = 1$.

Das 3D Isingmodell ist bisher nicht analytisch gelöst worden. Numerische Arbeiten zeigen $T_c^{3D} \approx 2T_c^{2D}$ mit kritischem Exponenten

$$\eta \approx 0.313. \quad (5.151)$$

Literaturverzeichnis

- [1] F. Schlögl, *Probability and heat: fundamentals of thermostatics* (Springer-Verlag, 2013).
- [2] F. Schwabl, *Statistische Mechanik* (Springer Verlag, 2006).
- [3] W. Nolting, *Statistische Physik*, 5 ed., Grundkurs Theoretische Physik, Vol. 6 (Springer Verlag, 2005).
- [4] F. Scheck, *Theoretische Physik 5 – Statistische Theorie der Wärme*, 1st ed. (Springer-Verlag Berlin Heidelberg, 2008).
- [5] T. Fließbach, *Statistische Physik* (Springer, 1993).
- [6] M. Kliesch, *Sciebo-Verzeichnis zur Veranstaltung ‘Statistische Mechanik’*, <https://uni-duesseldorf.sciebo.de/s/hV5DgfmUEEvBrRp> (2019), Passwort: “Z4fD.jM5q”.
- [7] J. Elstrodt, *Maß- und Integrationstheorie*, Vol. 7 (Springer, 1996).
- [8] R. T. Rockafellar, *Convex analysis*, 2nd ed. (Princeton university press, 1970).
- [9] T. M. Cover and J. A. Thomas, *Elements of information theory* (John Wiley and Sons, New York, 2012).
- [10] M. Kliesch, *Sciebo-File-Drop zur Abgabe von Übungen zur ‘Statistische Mechanik’*, <https://uni-duesseldorf.sciebo.de/s/zGMI6G66sidwIcv> (2019), Passwort: “Z4fD.jM5q”.
- [11] R. K. P. Zia, E. F. Redish, and S. R. McKay, *Making sense of the Legendre transform*, *American Journal of Physics* **77**, 614 (2009), [arXiv:0806.1147](https://arxiv.org/abs/0806.1147) [physics.ed-ph].
- [12] E. Jaynes, *Information Theory and Statistical Mechanics*, *Phys. Rev.* **106**, 620 (1957).
- [13] A. Caticha and A. Giffin, Updating Probabilities, in *Bayesian Inference and Maximum Entropy Methods In Science and Engineering*, American Institute of Physics Conference Series, Vol. 872, edited by A. Mohammad-Djafari (2006) pp. 31–42, [arXiv:physics/0608185](https://arxiv.org/abs/physics/0608185) [physics.data-an].

- [14] A. Giffin and A. Caticha, Updating Probabilities with Data and Moments, in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, American Institute of Physics Conference Series, Vol. 954, edited by K. H. Knuth, A. Caticha, A. Giffin, and C. C. Rodríguez (2007) pp. 74–84, [arXiv:0708.1593 \[physics.data-an\]](#).
- [15] C. Gogolin and J. Eisert, *Equilibration, thermalisation, and the emergence of statistical mechanics in closed quantum systems*, *Rep. Prog. Phys.* **79**, 056001 (2016), [arXiv:1503.07538 \[quant-ph\]](#).
- [16] J. Goold, M. Huber, A. Riera, L. del Rio, and P. Skrzypczyk, *The role of quantum information in thermodynamics — a topical review*, *J. Phys. A: Math. Theor.* **49**, 143001 (2016), [arXiv:1505.07835 \[quant-ph\]](#).
- [17] F. Binder, L. A. Correa, C. Gogolin, J. Anders, and G. Adesso, eds., *Thermodynamics in the Quantum Regime*, Fundamental Theories of Physics, Vol. 195 (Springer International Publishing, 2019).
- [18] M. Kliesch, C. Gogolin, M. J. Kastoryano, A. Riera, and J. Eisert, *Locality of Temperature*, *Phys. Rev. X* **4**, 031019 (2014), [arXiv:1309.0816 \[quant-ph\]](#).
- [19] J. Watrous, *The Theory of Quantum Information* (Cambridge University Press, 2018).
- [20] E. Jaynes, *Information Theory and Statistical Mechanics. II*, *Phys. Rev.* **108**, 171 (1957).
- [21] Y. Guryanova, S. Popescu, A. J. Short, R. Silva, and P. Skrzypczyk, *Thermodynamics of quantum systems with multiple conserved quantities*, *Nat. Commun.* **7**, 12049 (2016), [arXiv:1512.01190 \[quant-ph\]](#).
- [22] N. Linden, S. Popescu, A. J. Short, and A. Winter, *Quantum mechanical evolution towards thermal equilibrium*, *Phys. Rev. E* **79**, 061103 (2009).
- [23] J. P. Keating, N. Linden, and H. J. Wells, *Spectra and Eigenstates of Spin Chain Hamiltonians*, *Commun. Math. Phys.* **338**, 81 (2015), [arXiv:1403.1121 \[math-ph\]](#).
- [24] F. G. S. L. Brandao and M. Cramer, *Equivalence of Statistical Mechanical Ensembles for Non-Critical Quantum Systems*, [arXiv:1502.03263 \[quant-ph\]](#).
- [25] S. M. Bhattacharjee and A. Khare, *Fifty Years of the Exact Solution of the Two-Dimensional Ising Model by Onsager*, *Curr. Sci.* **69**, 816 (1996), [arXiv:cond-mat/9511003 \[cond-mat\]](#).