

GITTERGASE UND DAS CURIE-WEISS MODELL

SEMINARARBEIT

Westfälische Wilhelms-Universität Münster
Fachbereich Mathematik und Informatik
Institut für Mathematische Statistik

Betreuung:

Prof. Dr. Matthias Löwe

Dr. Andrea Winkler

Eingereicht von:

Sven Upgang

Münster, 11. Februar 2013

Inhaltsverzeichnis

1. Einleitung	1
2. Gittergase und das Curie-Weiss Modell	3
2.1. Gibbsmaß und Hamiltonfunktion	3
2.2. Modelle für Ferromagnetismus	4
2.2.1. Das Curie-Weiss Modell	6
2.3. Kurz zusammengefasst	7
3. Das Verhalten der Magnetisierung im Curie-Weiss Modell	9
3.1. Das Prinzip der Großen Abweichung	9
3.2. Anwendung auf das Curie-Weiss Modell	10
A. Anhang	21
A.1. Weitere Erläuterungen	21

1. Einleitung

Diese Seminararbeit soll einen kleinen Einblick in das Curie-Weiss Modell liefern, ein Beispiel für ein Spin-System aus der statistischen Mechanik.

Die Statistische Mechanik beschäftigt sich mit Vielteilchenmodellen (z.B. Gittergasen oder Spinsystemen) und verwendet wahrscheinlichkeitstheoretische Aussagen, um beispielsweise physikalische Phänomene zu erklären. Zum besseren Verständnis dieser Modelle möchte ich kurz erläutern, was man unter einem Makro- bzw. Mikrozustand versteht.

Auf mikroskopischer Ebene beobachten wir einzelne Teilchen, z.B. einzelne Atome oder Moleküle, und weisen diesen Eigenschaften zu, wie z.B. deren Ort, Geschwindigkeit oder Spin. Diese Daten bei einem Experiment zu messen und zu verfolgen ist unmöglich, da wir uns bei der Teilchenzahl in einem Bereich von 10^{23} Teilchen bewegen (siehe Einheit 1 Mol = $6,022 \cdot 10^{23}$ Teilchen). In der statistischen Mechanik werden daher die Bewegungen, Orte oder andere Eigenschaften mit Wahrscheinlichkeiten belegt, die aus physikalischen Gesetzmäßigkeiten gezogen werden.

Wenn man für jedes Teilchen in dem System die Eigenschaften auf einen Wert festlegt, erhält man einen Mikrozustand. Aber der Aufenthaltsort eines Teilchens ist oft gar nicht entscheidend, sondern uns interessiert, was das gesamte System für Eigenschaften besitzt. Die Eigenschaften des gesamten System sind die Makrozustände, wie z.B. Energie, Druck oder Magnetisierung. Diese resultieren aus den Mikrozuständen, wichtig ist aber, dass dies keine 1 zu 1 Beziehung ist.

Zum Verständnis ein kurzes Beispiel ohne physikalischen Hintergrund:

Bei einem Gesellschaftsspiel darf man 3 Würfel mit jeweils 6 Seiten werfen. Ein Mikrozustand wäre dann der Vektor (n_1, n_2, n_3) wobei n_i die Zahl für den Würfel $i = 1, 2, 3$ angibt. Ein Makrozustand wäre die Summe der Augenzahlen, also $n_1 + n_2 + n_3$ bzw. auch $\frac{1}{3}(n_1 + n_2 + n_3)$. Für das Spiel ist es hilfreich eine hohe Augenzahl zu würfeln, aber ob wir $(1, 3, 6)$ oder $(2, 3, 5)$ werfen, macht für uns keinen Unterschied. Physika-

lich entspricht hierbei die Zahl eines Würfels den Eigenschaften eines Teilchens und die Gesamtaugenanzahl ist beispielsweise die resultierende Gesamtenergie. Wir haben demnach in diesem Beispiel $6^3 = 216$ Mikrozustände und 16 Makrozustände (Zahlen von 3 bis 18). Die statistische Mechanik würde nun untersuchen, wie die Verteilung der Makrozustände bei einer sehr großen Anzahl von Würfeln ist. In diesem Kontext würde die Anzahl der „Würfel“ im Bereich von 10^{23} liegen und deshalb betrachten wir den Grenzwert für die Anzahl der Würfel gegen unendlich, ähnlich wie im Gesetz der Großen Zahlen. (Die absolute Augenzahl würde in diesem Fall natürlich nicht konvergieren, daher betrachtet man die Augenzahl pro Würfel, also $\frac{1}{k} \sum_{i=1}^k n_i$. Dies ist bei unabhängigen Würfeln uninteressant und nicht schwer zu lösen, deshalb untersucht man Würfel, die sich gegenseitig beeinflussen.)

Der physikalische Hintergrund des Curie-Weiss Modells:

Die Eigenschaft der einzelnen Teilchen, die uns im Curie-Weiss Modell interessiert, ist der Spin. Dieser Spin kann in zwei Richtungen wirken, die wir als „Spin up“ (bzw. Spin nach oben) oder als „Spin down“ (bzw. Spin nach unten) bezeichnen werden. Dadurch dass sich in einem Stoff die Spins ausrichten können, hat er die Möglichkeit, magnetisch zu reagieren, das heißt zum einen von einem Magnet angezogen zu werden und zum anderen eventuell selbst zu einem Magneten zu werden. Dies geschieht, indem viele Spins den gleichen Wert annehmen, sich also gleich ausrichten. Hierbei addieren sich die magnetischen Momente der einzelnen Teilchen und ergeben ein spürbares magnetisches Moment für den gesamten Stoff.

Interessant ist nun, wann und wie sich diese Spins ausrichten, d.h. ob sie sich gegenseitig beeinflussen, eher die gleiche oder gegensätzliche Richtung bevorzugen und ob beispielsweise ein externes Magnetfeld benötigt wird, damit diese Ausrichtung erfolgt. Im Curie-Weiss Modell soll der Ferromagnetismus untersucht werden und das bedeutet, dass sich die Spins durch ein externes Magnetfeld gleich ausrichten, aber selbst nach dem Entfernen dieses Magnetfeldes noch eine Magnetisierung erhalten bleibt. Dies nennt man spontane Magnetisierung.

2. Gittergase und das Curie-Weiss Modell

2.1. Gibbsmaß und Hamiltonfunktion

Wir betrachten ein Gitter, d.h. für ein $d \in \mathbb{N}$ sei $\Lambda \subset \mathbb{Z}^d$ eine endliche Teilmenge. Für den 2-dimensionalen Fall könnte Λ beispielsweise wie in Abbildung 2.1 dargestellt aussehen. Die Elemente aus Λ (schwarze Punkte) stellen unsere Moleküle, Atome oder in

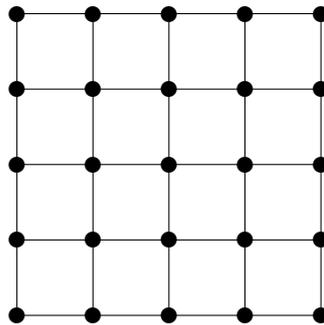


Abbildung 2.1.: Beispiel für ein Gitter im \mathbb{Z}^2

sozialen Modellen einzelne Personen dar. Über die Eigenschaften, die die einzelnen Elemente annehmen können, erhalten wir die Menge der Mikrozustände Ω . Des Weiteren benötigen wir für das System eine Hamiltonfunktion.

Definition 2.1.

Sei Ω die Menge aller Mikrozustände. Eine Hamiltonfunktion $H_\Lambda : \Omega \rightarrow \mathbb{R}$ ordnet jedem Mikrozustand einen Energiewert zu.

Welche σ aus Ω auftreten und welche Zustände bevorzugt werden, wird über das Gibbsmaß $\mu_{\Lambda,\beta}$ angegeben, welches jedem Mikrozustand σ seine Wahrscheinlichkeit zuordnet.

Definition 2.2.

Sei Ω die Menge aller Mikrozustände und H_Λ eine Hamiltonfunktion. Für $\sigma \in \Omega$ ist das Gibbsmaß definiert durch

$$\mu_{\Lambda,\beta}(\sigma) := \frac{e^{-\beta H_\Lambda(\sigma)}}{\sum_{\sigma \in \{-1,1\}^N} e^{-\beta H_\Lambda(\sigma)}}$$

für ein $\beta \in \mathbb{R}$ mit $\beta > 0$.

β stellt hier einen Parameter für das ganze System dar. In den physikalischen Modellen ist β beispielsweise die inverse Temperatur, also $\beta = \frac{1}{T} > 0$. Dabei gilt $T > 0$ durch die Betrachtung der Kelvinskala anstelle der Celsiusskala. Bei einem Gibbsmaß lässt sich folgendes Verhalten beobachten:

- Falls β sehr groß ist, d.h. bei tiefen Temperaturen, konzentriert sich $\mu_{\Lambda,\beta}$ auf die σ mit minimaler Energie $H_\Lambda(\sigma)$.
- Falls β sehr klein, d.h. bei hohen Temperaturen, herrscht nahezu eine Gleichverteilung. Dies spiegelt die Eigenschaft wieder, dass ein System bei höheren Temperaturen leichter Zustände annehmen kann, die nicht der minimalen Energie entsprechen.

Durch diesen Einfluss von β kann eine sprunghafte Änderung des Verhaltens eines Systems auftreten. Denn bei niedrigem β ist die Entropie, d.h. die Anzahl der Mikrozustände für einen Makrozustand, ausschlaggebend und bei großem β erhält der Makrozustand mit der geringsten Energie die größte Wahrscheinlichkeit.

2.2. Modelle für Ferromagnetismus

Wir möchten uns näher mit den Modellen für Ferromagnetismus beschäftigen. Dabei besitzen die Teilchen einen sogenannten elektromagnetischen Spin. Ein einzelnes Teilchen kann den Zustand „spin-up“ (1) oder „spin-down“ (−1) annehmen. Der Spin ist die einzige Eigenschaft, die wir für ein einzelnes Teilchen betrachten, und daher ist die Menge der Mikrozustände gegeben durch $\{-1, 1\}^\Lambda$. Ein Element dieser Menge nennen wir auch Spinkonfiguration.

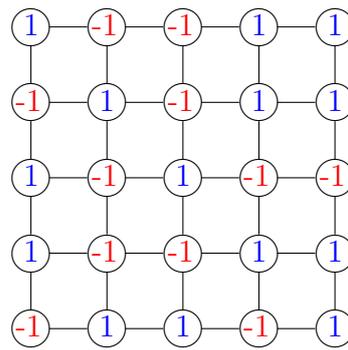


Abbildung 2.2.: Beispiel einer Spinkonfiguration

Der für uns relevante Makrozustand ist die Magnetisierung, welcher anzeigt ob die Spins eher gleich ausgerichtet sind oder nicht.

Definition 2.3.

Die Magnetisierung für eine Spinkonfiguration $\sigma \in \{-1, 1\}^\Lambda$ ist definiert als

$$m_\Lambda(\sigma) := \frac{1}{|\Lambda|} \sum_{i \in \Lambda} \sigma_i.$$

Die Normierung durch die Anzahl der Elemente von Λ gewährleistet eine einfache Interpretation von m_N ohne die genaue Anzahl der Teilchen zu kennen. Wie zuvor schon bemerkt wurde, ist bei vielen gleich ausgerichteten Spins die Magnetisierung hoch. Zusätzlich gibt das Vorzeichen an in welche Richtung die Magnetisierung wirkt. Wie in der Einleitung erwähnt, ist es nicht wichtig, wie der Spin an einer festen Stelle ist, weil die Magnetisierung invariant gegenüber jeder Permutation ist.

Sehr allgemein wird als Hamiltonfunktion

$$H_\Lambda(\sigma) = - \sum_{A \cap \Lambda \neq \emptyset} \varphi_A(\sigma_A)$$

verwendet, bei der jeder Teilmenge A von Λ über eine Funktion φ_A ein Interaktionspotential zugeordnet werden kann. Es ist $\sigma_A = (\sigma_i)_{i \in A}$. Bei Magnetismus geht man davon aus, dass die Wechselwirkungen zwischen den Spins vom Ort abhängig sind. Dabei nimmt der Einfluss den ein Spin auf den Anderen hat, bei größerer Entfernung ab. Im Ising Modell schränkt man den Einfluss ein, sodass nur jeweils zwei benachbarte Spins physikalisch aufeinander einwirken. Man verwendet in dem Ising Modell bei gegebenen

$h \in \mathbb{R}$ die Hamiltonfunktion

$$H_{\Lambda,h}(\sigma) = - \sum_{\substack{i,j \in \Lambda, \\ \langle i,j \rangle}} \sigma_i \sigma_j - h \sum_{i \in \Lambda} \sigma_i.$$

Hierbei bedeutet $\langle i, j \rangle$, dass die Spins σ_i und σ_j auf dem Gitter Nachbarn sind. Dieser kurze Erläuterung zum Ising Modell soll aber nur verdeutlichen, wie die Gitterstruktur in einem Modell genutzt werden kann. In Abbildung 2.2 wird dies deutlich gemacht durch die Verbindungslinien zwischen den Spins. Genauere Erläuterung zu dieser Art von Hamiltonfunktion folgt nun im Curie-Weiss Modell.

2.2.1. Das Curie-Weiss Modell

Der Einfluss von den nächsten Nachbarn ist schwierig zu handhaben und daher nehmen wir im Curie-Weiss Modell an, dass für jedes Spinpaar die gleiche Wechselwirkung auftritt. Dadurch hat der Ort des Spins keinen Einfluss mehr und wir vereinfachen unsere Spinkonfiguration zu $\sigma = (\sigma_1, \dots, \sigma_N)$ für $N = |\Lambda|$. Die Magnetisierung lässt sich nun schreiben als

$$m_N(\sigma) := \frac{1}{N} \sum_{i=1}^N \sigma_i.$$

Für das Curie-Weiss Modell verwenden wir die Hamilton- oder Energiefunktion

$$\begin{aligned} H_{N,h}(\sigma) &= - \underbrace{\frac{1}{2N} \sum_{1 \leq i,j \leq N} \sigma_i \sigma_j}_A - h \underbrace{\sum_{1 \leq i \leq N} \sigma_i}_B \\ &= - \underbrace{\frac{1}{2} \sum_{1 \leq i \leq N} \sigma_i \cdot \left(\frac{1}{N} \sum_{1 \leq j \leq N} \sigma_j \right)}_A - h \underbrace{\sum_{1 \leq i \leq N} \sigma_i}_B \end{aligned}$$

Erinnern wir uns an das Gibbsmaß und die Tatsache, dass Zustände mit niedriger Energie bevorzugt werden, so lässt sich diese Funktion leicht interpretieren:

- Für Teil A gilt, dass wir eine Wechselwirkung zwischen σ_i und m_N (bzw. σ_i und σ_j) in der Art haben, dass wenn beide die selbe Ausrichtung besitzen der Summand, durch das Minuszeichen am Anfang des Terms, die Gesamtenergie verringert.

- Die Variable h in Teil B stellt ein externes Magnetfeld dar. Hier greifen wir auf das selbe Instrument zurück wie in A, denn durch das negative Vorzeichen und das Produkt von σ_i mit h , wird durch Spins mit der selben Ausrichtung wie das externe Magnetfeld eine geringere Energie erreicht.

Eine Besonderheit im Curie-Weiss Modell ist, dass die Hamiltonfunktion sich umschreiben lässt zu

$$\begin{aligned}
 H_{N,h}(\sigma) &= -\frac{1}{2} \sum_{1 \leq i \leq N} \sigma_i \cdot \left(\frac{1}{N} \sum_{1 \leq j \leq N} \sigma_j \right) - h \sum_{i=1}^N \sigma_i \\
 &= -\frac{N}{2} \left(\frac{1}{N^2} \sum_{1 \leq i, j \leq N} \sigma_i \sigma_j \right) - hN \frac{1}{N} \sum_{i=1}^N \sigma_i \\
 &= -\frac{N}{2} \left(\frac{1}{N} \sum_{i=1}^N \sigma_i \right)^2 - hN \frac{1}{N} \sum_{i=1}^N \sigma_i \\
 &= -\frac{N}{2} \cdot [m_N(\sigma)]^2 - hN \cdot m_N(\sigma).
 \end{aligned}$$

Damit ist das Gibbsmaß im Curie-Weiss Modell für $\sigma = (\sigma_1, \dots, \sigma_N)$ von der Form

$$\mu_{N,\beta,h}(\sigma) = \frac{e^{\beta \frac{N}{2} \cdot [m_N(\sigma)]^2 + N\beta h \cdot m_N(\sigma)}}{\sum_{\sigma \in \{-1,1\}^N} e^{\beta \frac{N}{2} \cdot [m_N(\sigma)]^2 + N\beta h \cdot m_N(\sigma)}}.$$

Der Nenner, der das Gibbsmaß zu einem Wahrscheinlichkeitsmaß normiert, wird als Zustandssumme bezeichnet.

$$Z_{N,\beta,h} := \sum_{\sigma \in \{-1,1\}^N} e^{\beta \frac{N}{2} \cdot [m_N(\sigma)]^2 + N\beta h \cdot m_N(\sigma)}$$

Wir möchten nun untersuchen, auf welche Werte sich das Wahrscheinlichkeitsmaß konzentriert, wenn wir sehr viele Teilchen in unserem Gitter haben, also betrachten wir den Limes $N \rightarrow \infty$.

2.3. Kurz zusammengefasst

Wir betrachten nun folgendes Szenario

- N Teilchen, wobei Teilchen i den Spin $\sigma_i \in \{-1, 1\}$ annehmen kann.

- Einer Spinkonfiguration $\sigma = (\sigma_1, \dots, \sigma_N)$ wird die Energie $H_{N,h}(\sigma)$ zugewiesen.
- Die Hamiltonfunktion fördert die Ausrichtung der Spins in eine Richtung, vorzugsweise in die des externen Magnetfeldes.
- Das Gibbsmaß bestimmt die Wahrscheinlichkeit einer Konfiguration σ und hängt außerdem von der inversen Temperatur β ab.
- Durch β ändert sich der Einfluss von Energie und Entropie auf die Wahrscheinlichkeiten für m_N .
- Frage: Was geschieht mit der Magnetisierung m_N für den Grenzwert $N \rightarrow \infty$ in Abhängigkeit von β , also wodurch wird die Magnetisierung stärker beeinflusst, durch die Energie oder die Entropie?

3. Das Verhalten der Magnetisierung im Curie-Weiss Modell

3.1. Das Prinzip der Großen Abweichung

Zunächst benötigen wir ein Prinzip der Großen Abweichung.

Definition 3.1 (Prinzip der Großen Abweichung).

Eine Folge S_1, S_2, \dots von reellwertigen Zufallsvariablen genügt einem Prinzip der großen Abweichungen mit Geschwindigkeit n und Ratenfunktion I , falls gilt

i.) I ist von unten halbstetig und $0 \leq I \leq +\infty$, $I \not\equiv 0$.

ii.) Die Niveaumengen $N_L = \{x : I(x) \leq L\}$ von I sind für alle $L \geq 0$ kompakt.

iii.) Für alle $A \in \mathcal{B}$ gilt:

$$-\inf_{x \in \overset{\circ}{A}} I(x) \leq \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(S_n \in \overset{\circ}{A}) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(S_n \in \bar{A}) \leq -\inf_{x \in \bar{A}} I(x)$$

Hierbei bezeichnet $\overset{\circ}{A}$ das topologische Innere von A und \bar{A} den topologischen Abschluss von A .

Bemerkung 3.2.

Halbstetigkeit von unten bedeutet anschaulich, dass in den Stellen wo die Funktion springt, immer der untere Punkt den Funktionswert liefert. Also die Funktion nicht nach unten springen kann. Formal wird in einem metrischen Raum in einem Punkt x die Halbstetigkeit von unten ausgedrückt als Bedingung, dass

$$\liminf_{y \rightarrow x} f(y) \leq f(x) = f(\lim_{y \rightarrow x} y)$$

Die Idee einer Großen Abweichung wird deutlicher, wenn man die Ungleichungen umformt. Zum Beispiel gilt für eine abgeschlossene Menge \bar{A} und hinreichend großes n ,

dass für ein $\epsilon \geq 0$.

$$\begin{aligned} \frac{1}{n} \log \mathbb{P}(S_n \in \bar{A}) &\leq - \inf_{x \in \bar{A}} I(x) + \epsilon \\ \Leftrightarrow \log \mathbb{P}(S_n \in \bar{A}) &\leq -n \left(\inf_{x \in \bar{A}} I(x) + \epsilon \right) \\ \Leftrightarrow \mathbb{P}(S_n \in \bar{A}) &\leq e^{-n \left(\inf_{x \in \bar{A}} I(x) + \epsilon \right)} \end{aligned}$$

Über die Ratenfunktion auf der Menge \bar{A} kann also ausgedrückt werden, wodurch die Wahrscheinlichkeit in dieser Menge zu landen beschränkt ist. Je größer das Infimum auf \bar{A} ist, desto kleiner können die Schranken gewählt werden.

3.2. Anwendung auf das Curie-Weiss Modell

Um die Sätze und Aussagen, die für ein Prinzip der Großen Abweichung gelten, zu nutzen, wählen wir den Wahrscheinlichkeitsraum

$$(\Omega, \mathcal{F}, \mathbb{P}) \text{ mit } \Omega = \{-1, 1\}^\infty, \mathcal{F} = \otimes_{i=1}^\infty \mathcal{P}(\{-1, 1\}).$$

Wir betrachten unabhängige, identisch verteilte Zufallsvariablen X_1, X_2, \dots . Über

$$\mathbb{P}_i(\{-1\}) = \frac{1}{2} = \mathbb{P}_i(\{1\})$$

erhalten wir auf (Ω, \mathcal{F}) das Produktmaß

$$\mathbb{P} = \otimes_{i=1}^\infty \mathbb{P}_i.$$

Wir werden immer die ersten N bzw. n Stellen betrachten. Dort ist die Wahrscheinlichkeit gegeben durch

$$\begin{aligned} \mathbb{P}((\sigma_1, \dots, \sigma_n) \times \{-1, 1\} \times \dots) &= \prod_{i=1}^n \mathbb{P}(\{\sigma_i\}) \\ &= \left(\frac{1}{2}\right)^n \end{aligned}$$

wobei $\sigma_i \in \{-1, 1\}$ für alle $i \leq n$.

Wir betrachten diesen Wahrscheinlichkeitsraum, weil so eine Konvergenz für das Bild-

maß der Magnetisierung des Gibbsmaßes hergeleitet werden kann. Das Bildmaß der Magnetisierung unter dem Gibbsmaß lässt sich durch eine Dichte auf dem Raum $(\Omega, \mathcal{F}, \mathbb{P})$ definieren. Über das Prinzip der Großen Abweichung und eine Folgerung für das über die Dichte definierte Maß, kann dann das Verhalten der Magnetisierung im Curie-Weiss Modell im Limes $N \rightarrow \infty$ abgeleitet werden.

Lemma 3.3.

Seien X_1, X_2, \dots unabhängige Zufallsvariablen mit einer Laplaceverteilung auf $\{-1, 1\}$. Dann genügt die Folge $(S_n)_{n \in \mathbb{N}}$ mit $S_n = \frac{1}{n} \sum_{i=1}^n X_i$ einem Prinzip der großen Abweichung mit Geschwindigkeit n und Ratenfunktion

$$I(x) = \frac{1+x}{2} \cdot \log(1+x) + \frac{1-x}{2} \cdot \log(1-x).$$

Dies folgt aus einem Prinzip der Großen Abweichung für die Partialsummenfolge von Bernoulli verteilten Zufallsvariablen. Die Aussage von Lemma 3.3 kann dann über eine Transformation auf die Menge $\{-1, 1\}$ bewiesen werden.

Aus Lemma 3.3 erhalten wir die Aussage, dass $(m_N)_{N \in \mathbb{N}}$ unter dem Produktmaß \mathbb{P} einem Prinzip der großen Abweichung genügt mit der oben genannten Ratenfunktion I .

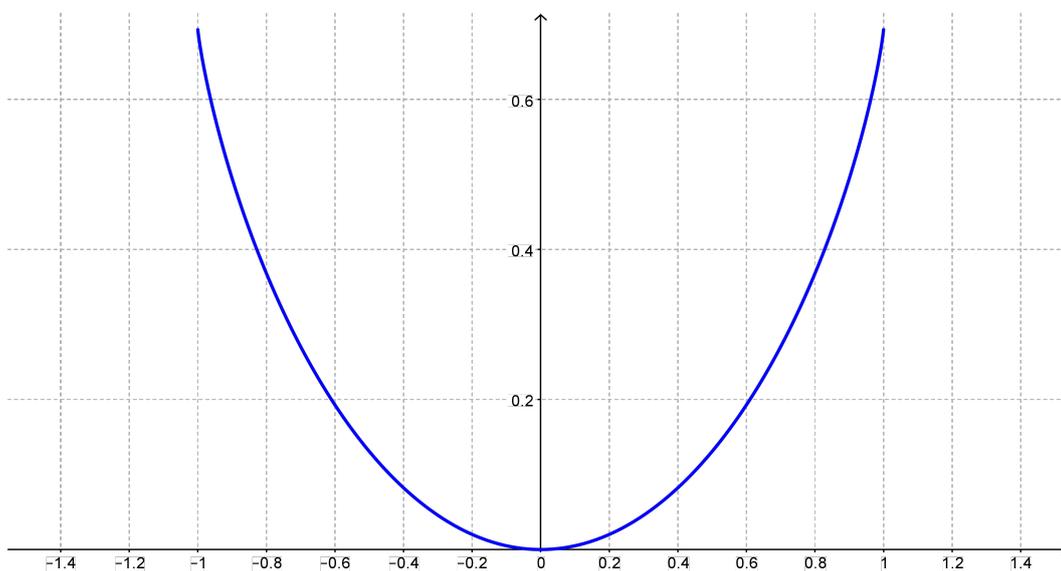


Abbildung 3.1.: Ratenfunktion aus Lemma 3.3

In der Abbildung 3.1 ist noch mal zu erkennen, dass die Funktion $I(x)$ stetig, insbesondere halbstetig von unten, positiv, aber nicht konstant 0 ist. Die Niveaumengen entsprechen immer abgeschlossen Intervallen und sind somit auch kompakt.

m_N hat unter \mathbb{P} eine andere Verteilung als unter dem Gibbsmaß. Trotzdem hilft uns die Aussage des Prinzips der großen Abweichung von m_N unter \mathbb{P} indem wir den nachfolgenden Satz verwenden.

Satz 3.4.

Sei S_1, S_2, \dots eine Folge die einem Prinzip der Großen Abweichungen mit Geschwindigkeit n und Ratenfunktion I genügt und $f : \mathbb{R} \rightarrow \mathbb{R}$ eine stetige und beschränkte Funktion. Sei \mathbb{P}_n^f ein Wahrscheinlichkeitsmaß, welches mit

$$J_n(A) = \int_A e^{n \cdot f(x)} d\mathbb{P}^{S_n}(x), \text{ für eine Menge } A \in \mathcal{B}$$

definiert ist durch

$$\mathbb{P}_n^f(A) = \frac{J_n(A)}{J_n(\mathbb{R})} = \int_A \frac{e^{n \cdot f(x)}}{\int_{\mathbb{R}} e^{n \cdot f(y)} d\mathbb{P}^{S_n}(y)} d\mathbb{P}^{S_n}(x).$$

Dann genügt die Folge $\mathbb{P}_1^f, \mathbb{P}_2^f, \dots$ von Wahrscheinlichkeitsmaßen einem Prinzip der Großen Abweichung mit Geschwindigkeit n und Ratenfunktion

$$I^f(x) = I(x) - f(x) + \sup_{y \in \mathbb{R}} [f(y) - I(y)].$$

Wenden wir nun Satz 3.4 an, so können wir das Prinzip der Großen Abweichung von m_N unter \mathbb{P} auf das Gibbsmaß überführen.

Satz 3.5.

Die Folge der Magnetisierungen m_1, m_2, \dots genügt unter dem Gibbsmaß $\mu_{N,\beta,h}$ einem Prinzip der Großen Abweichungen mit Geschwindigkeit N und Ratenfunktion J , wobei J definiert ist als

$$J(x) = I(x) - \frac{\beta x^2}{2} - \beta h x + \sup_{y \in [-1,1]} \left[\frac{\beta y^2}{2} + \beta h y - I(x) \right]$$

und I bezeichnet wieder die Ratenfunktion

$$I(x) = \frac{1+x}{2} \log(1+x) + \frac{1-x}{2} \log(1-x)$$

Beweis. Wir haben folgendes Setting für diesen Beweis

$$(\Omega, \mathcal{F}, \mathbb{P}) \text{ mit } \Omega = \{-1, 1\}^\infty, \mathcal{F} = \otimes_{i=1}^\infty \mathcal{P}(\{-1, 1\}) \text{ und } \mathbb{P} = \otimes_{i=1}^\infty \mathbb{P}_i.$$

Dabei gilt

$$\mathbb{P}_i(\{-1\}) = \frac{1}{2} = \mathbb{P}_i(\{1\})$$

und somit

$$\begin{aligned} \mathbb{P}((\sigma_1, \dots, \sigma_n) \times \{-1, 1\} \times \dots) &= \prod_{i=1}^n \mathbb{P}(\{\sigma_i\}) \\ &= \left(\frac{1}{2}\right)^n \end{aligned}$$

Wir betrachten die Magnetisierung als Folge, die unter \mathbb{P} einem Prinzip der Großen Abweichung genügt und setzen

$$\begin{aligned} m_N &= \frac{1}{N} \sum_{i=1}^N \sigma_i \\ f(x) &= \frac{\beta}{2} x^2 + \beta h x \end{aligned}$$

Jetzt ist es eine einfache Rechnung zu zeigen, dass die Folge \mathbb{P}_n^f aus dem Satz die Folge der Bildmaße von m_N unter dem Gibbsmaß ist, und nach Satz 3.4 einem Prinzip der Großen Abweichung genügt. Sei $S \in \mathcal{B}$, dann gilt

$$\begin{aligned} \mathbb{P}_N^f(S) &= \frac{\int_S e^{\frac{\beta N}{2} x^2 + \beta h N x} d\mathbb{P}^{m_N}(x)}{\int_{\mathbb{R}} e^{\frac{\beta N}{2} x^2 + \beta h N x} d\mathbb{P}^{m_N}(x)} \\ &= \frac{\int_{m_N^{-1}(S) \times \{-1, 1\}^\infty} e^{\frac{\beta N}{2} m_N(\sigma_1, \dots, \sigma_N)^2 + \beta h N \cdot m_N(\sigma_1, \dots, \sigma_N)} d\mathbb{P}(\sigma)}{\int_{m_N^{-1}(\mathbb{R}) \times \{-1, 1\}^\infty} e^{\frac{\beta N}{2} m_N(\sigma_1, \dots, \sigma_N)^2 + \beta h N \cdot m_N(\sigma_1, \dots, \sigma_N)} d\mathbb{P}(\sigma)} \\ &= \frac{\int_{m_N^{-1}(S)} e^{\frac{\beta N}{2} m_N(\sigma_1, \dots, \sigma_N)^2 + \beta h N \cdot m_N(\sigma_1, \dots, \sigma_N)} d\mathbb{P}_1(\sigma_1) \dots d\mathbb{P}_N(\sigma_N)}{\int_{m_N^{-1}(\mathbb{R})} e^{\frac{\beta N}{2} m_N(\sigma_1, \dots, \sigma_N)^2 + \beta h N \cdot m_N(\sigma_1, \dots, \sigma_N)} d\mathbb{P}_1(\sigma_1) \dots d\mathbb{P}_N(\sigma_N)} \\ &= \frac{\sum_{x \in m_N^{-1}(S)} e^{\frac{\beta N}{2} m_N(x)^2 + \beta h N \cdot m_N(x)} \cdot \frac{1}{2}^N}{\sum_{y \in m_N^{-1}(\mathbb{R})} e^{\frac{\beta N}{2} m_N(y)^2 + \beta h N \cdot m_N(y)} \cdot \frac{1}{2}^N} \\ &= \frac{\sum_{x \in m_N^{-1}(S)} e^{\frac{\beta N}{2} m_N(x)^2 + \beta h N \cdot m_N(x)}}{\sum_{\sigma \in \{-1, 1\}^N} e^{\frac{\beta N}{2} m_N(\sigma)^2 + \beta h N \cdot m_N(\sigma)}} \end{aligned}$$

$$\begin{aligned}
&= \frac{\sum_{x \in m_N^{-1}(S)} e^{\frac{\beta N}{2} m_N(x)^2 + \beta h N \cdot m_N(x)}}{\sum_{\sigma \in \{-1,1\}^N} e^{\frac{\beta N}{2} m_N(\sigma)^2 + \beta h N \cdot m_N(\sigma)}} \\
&= \mu_{N,\beta,h}(m_N \in S)
\end{aligned}$$

Zur Erinnerung:

$$\mu_{N,\beta,h}(\sigma) = \frac{e^{\beta \frac{N}{2} \cdot [m_N(\sigma)]^2 + N\beta h \cdot m_N(\sigma)}}{\sum_{\sigma \in \{-1,1\}^N} e^{\beta \frac{N}{2} \cdot [m_N(\sigma)]^2 + N\beta h \cdot m_N(\sigma)}}.$$

Damit ist die Gleichheit der Maße gezeigt und die Berechnung der Ratenfunktion folgt durch einfaches Einsetzen von $f(x)$:

$$J(x) = I(x) - \frac{\beta x^2}{2} - \beta h x + \sup_{y \in [-1,1]} \left[\frac{\beta y^2}{2} + \beta h y - I(x) \right]$$

□

Wir haben nun bewiesen, dass das Maß $\mu_{N,\beta,h}(m_N \in \cdot)$ einem Prinzip der Großen Abweichung genügt. Das bedeutet, dass m_N unter $\mu_{N,\beta,h}$ einem Prinzip der Großen Abweichung genügt. Wir können somit durch die Ratenfunktion

$$J(x) = I(x) - \frac{\beta x^2}{2} - \beta h x + \sup_{y \in [-1,1]} \left[\frac{\beta y^2}{2} + \beta h y - I(x) \right]$$

Grenzen für die Wahrscheinlichkeiten angeben. Nun folgt aus dem Prinzip der Großen Abweichung direkt eine Konvergenz der Wahrscheinlichkeit, mit der wir letztendlich konkret das Ergebnis für das Curie-Weiss Modell berechnen können.

Satz 3.6.

Sei S_1, S_2, \dots eine Folge, die einem Prinzip der Großen Abweichung mit Geschwindigkeit n und Ratenfunktion I genügt. Sei

$$\mathcal{N} = \{x \in \mathbb{R} : I(x) = 0\}$$

die Menge aller Nullstellen der Ratenfunktion I und für jedes $\epsilon > 0$ sei

$$\mathcal{N}_\epsilon = \{x : \|x - \mathcal{N}\| < \epsilon\}$$

die Menge aller Punkte, die weniger als ϵ von einer Nullstelle entfernt sind. Dann gilt

$$\mathbb{P}(S_n \in \mathcal{N}_\epsilon^c) \rightarrow 0 \text{ für } n \rightarrow \infty.$$

Beweis. Die Funktion I ist von unten Halbstetig und \mathcal{N} war die Menge der globalen Minima, deshalb ist

$$a := \inf_{x \in \mathcal{N}_\epsilon^c} I(x) > 0.$$

Weiter ist N_ϵ die Vereinigung der offenen ϵ -Umgebungen der Nullstellen von I . Daraus folgt, dass N_ϵ^c eine abgeschlossene Menge sein muss und wir somit für N_ϵ^c die obere Abschätzung aus dem Prinzip der Großen Abweichung nutzen können. Daher gilt für hinreichend großes n

$$\begin{aligned} 0 \leq \mathbb{P}(S_n \in N_\epsilon^c) &\leq e^{-n \inf_{x \in N_\epsilon^c} I(x)} \\ &= e^{-na} \xrightarrow{n \rightarrow \infty} 0 \end{aligned}$$

und daraus folgt direkt

$$\mathbb{P}(S_n \in \mathcal{N}_\epsilon^c) \rightarrow 0 \text{ für } n \rightarrow \infty$$

□

Nur die Nullstellen der Ratenfunktionen können im Grenzwert noch Masse erhalten und in unserem Modell bedeutet dies, dass für eine sehr große Anzahl von Teilchen nur diese Magnetisierungen vorkommen und die restlichen exponentiell unwahrscheinlich sind. Da die Ratenfunktion nicht-negativ ist, können wir diese einfach auf Minimalstellen untersuchen. Wir untersuchen nun

$$J(x) = I(x) - \frac{\beta x^2}{2} - \beta h x + \sup_{y \in [-1,1]} \left[\frac{\beta y^2}{2} + \beta h y - I(y) \right]$$

auf Minimalstellen. I bezeichnet wieder die Ratenfunktion

$$I(x) = \frac{1+x}{2} \log(1+x) + \frac{1-x}{2} \log(1-x)$$

Durch ableiten erhalten wir

$$\begin{aligned} J'(x) &= -\beta x - \beta h + I'(x) \\ &= -\beta(x + h) + I'(x) \end{aligned}$$

Die Ableitung von I berechnet sich über die Produktregel zu

$$\begin{aligned} I'(x) &= \frac{1+x}{2} \cdot \frac{1}{1+x} + \frac{1}{2} \log(1+x) - \frac{1-x}{2} \cdot \frac{1}{1-x} - \frac{1}{2} \log(1-x) \\ &= \frac{1}{2} + \frac{1}{2} \log(1+x) - \frac{1}{2} - \frac{1}{2} \log(1-x) \\ &= \frac{1}{2} \log\left(\frac{1+x}{1-x}\right) \end{aligned}$$

Wir müssen eine Extremstelle $x \in [-1, 1]$ finden, die

$$\begin{aligned} -\beta(x+h) + \frac{1}{2} \log\left(\frac{1+x}{1-x}\right) &= 0 \\ \Leftrightarrow \frac{1}{2} \log\left(\frac{1+x}{1-x}\right) &= \beta(x+h) \\ \Leftrightarrow \frac{1+x}{1-x} &= e^{2\beta(x+h)} \\ \Leftrightarrow 1+x &= e^{2\beta(x+h)} \cdot (1-x) \\ \Leftrightarrow 1+x &= e^{2\beta(x+h)} - x \cdot e^{2\beta(x+h)} \\ \Leftrightarrow x + x \cdot e^{2\beta(x+h)} &= e^{2\beta(x+h)} - 1 \\ \Leftrightarrow x(e^{2\beta(x+h)} + 1) &= e^{2\beta(x+h)} - 1 \\ \Leftrightarrow x &= \frac{e^{2\beta(x+h)} - 1}{e^{2\beta(x+h)} + 1} \end{aligned}$$

Als Nullstellen kommen nur die Lösungen der Gleichung

$$x = \frac{e^{2\beta(x+h)} - 1}{e^{2\beta(x+h)} + 1} = \tanh(\beta(x+h)).$$

in Frage. Aber nicht alle Lösungen führen zu einer Nullstelle. Insgesamt ergibt sich somit folgendes Bild:

Satz 3.7.

Unter dem Gibbsmaß $\mu_{N,\beta,h}$ konvergiert die Magnetisierung m_N exponentiell, im Sinne von Lemma 3.6, gegen folgende Grenzwerte:

1. Für $h > 0$ konvergiert m_N gegen die positive Lösung von

$$x = \tanh(\beta(x + h)).$$

2. Für $h < 0$ konvergiert m_N gegen die negative Lösung von

$$x = \tanh(\beta(x + h)).$$

3. Für $h = 0$ und $\beta \leq 1$ konvergiert m_N gegen 0.

4. Für $h = 0$ und $\beta > 1$ konvergiert m_N gegen die beiden Lösungen von

$$x = \tanh(\beta x)$$

welche beide von 0 verschieden sind.

5. Insbesondere erhält man bei $\beta > 1$

$$\lim_{h \downarrow 0} \lim_{N \rightarrow \infty} \mu_{N,\beta,h} \circ m_N^{-1} \Rightarrow \delta_{m^*(\beta)}$$

wobei $m^*(\beta)$ die größte Lösung von

$$x = \tanh(\beta x)$$

ist.

Die spontane Magnetisierung ist im fünften Fall zu erkennen, wo durch den Limes $h \downarrow 0$ das Entfernen eines externen Magnetfeldes dargestellt wird. Hier konvergiert die Magnetisierung also gegen eine Positive Lösung von $x = \tanh(\beta x)$, was bedeutet, dass wir auch wenn $h = 0$ wird, eine Magnetisierung beibehalten.

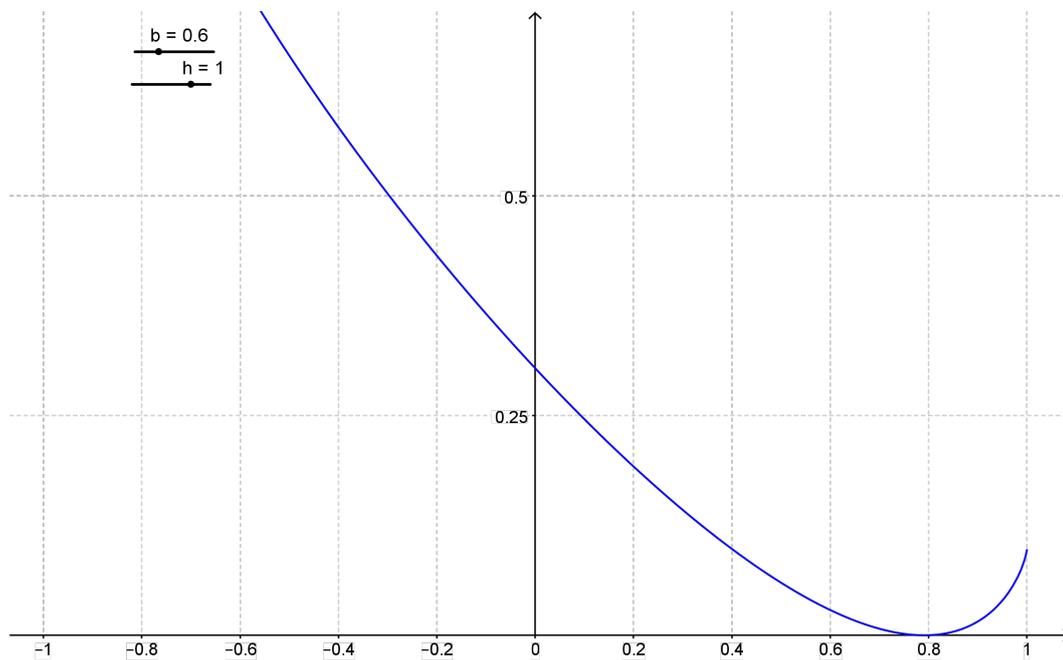


Abbildung 3.2.: Ratenfunktion für den Fall 1.

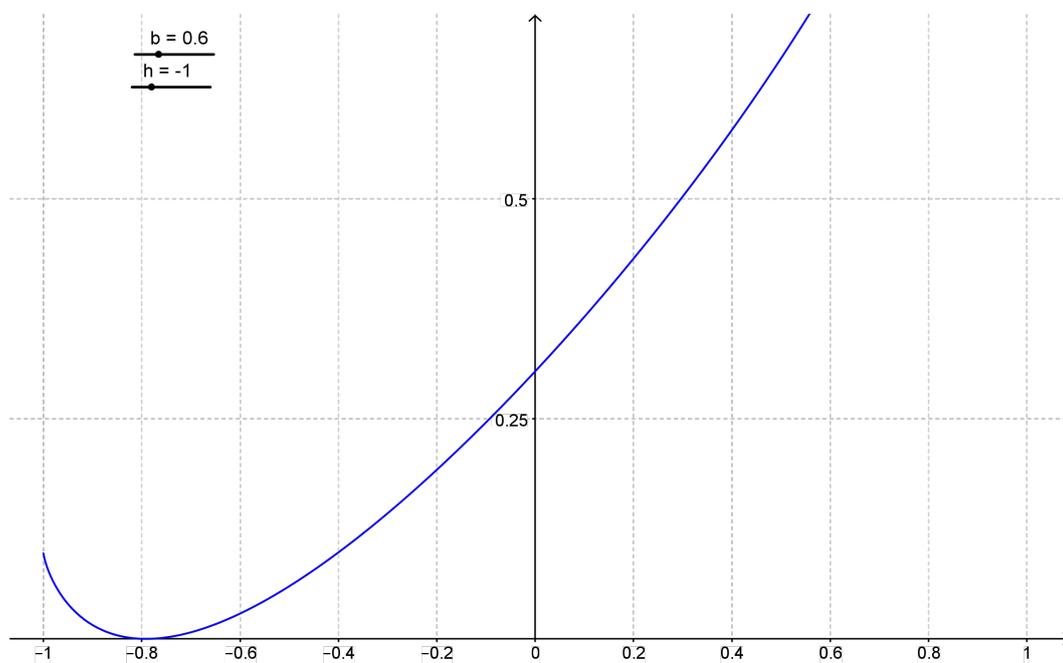


Abbildung 3.3.: Ratenfunktion für den Fall 2.

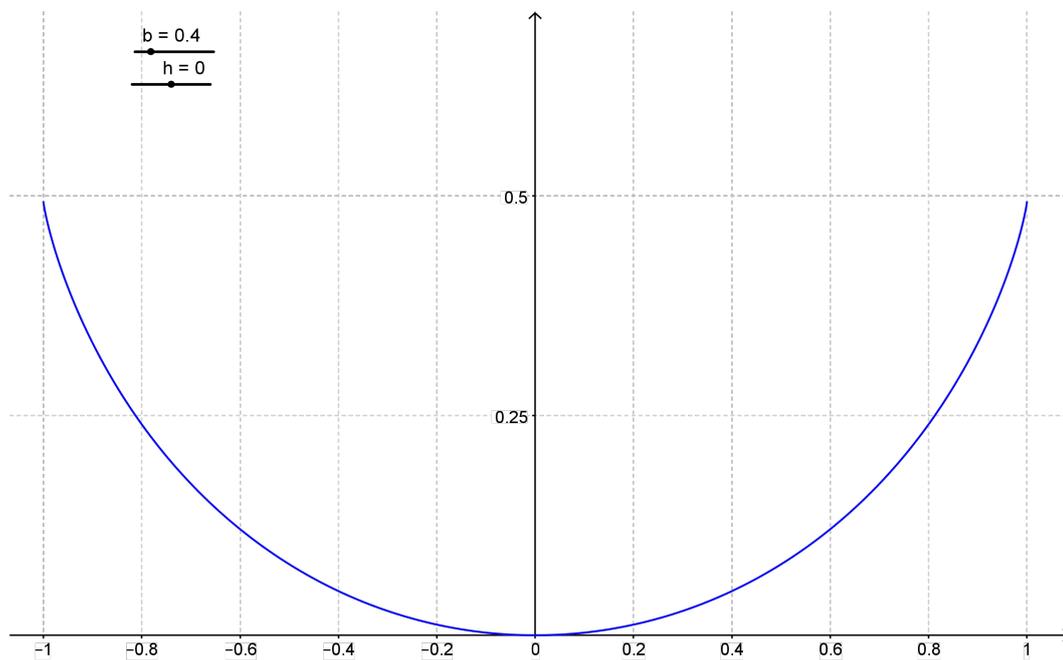
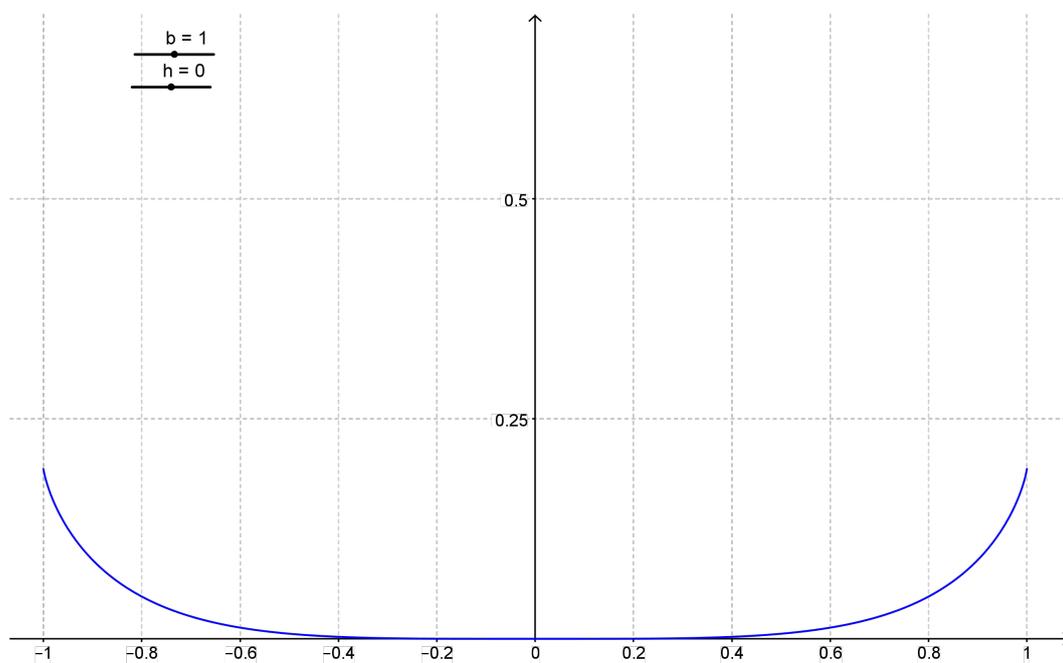
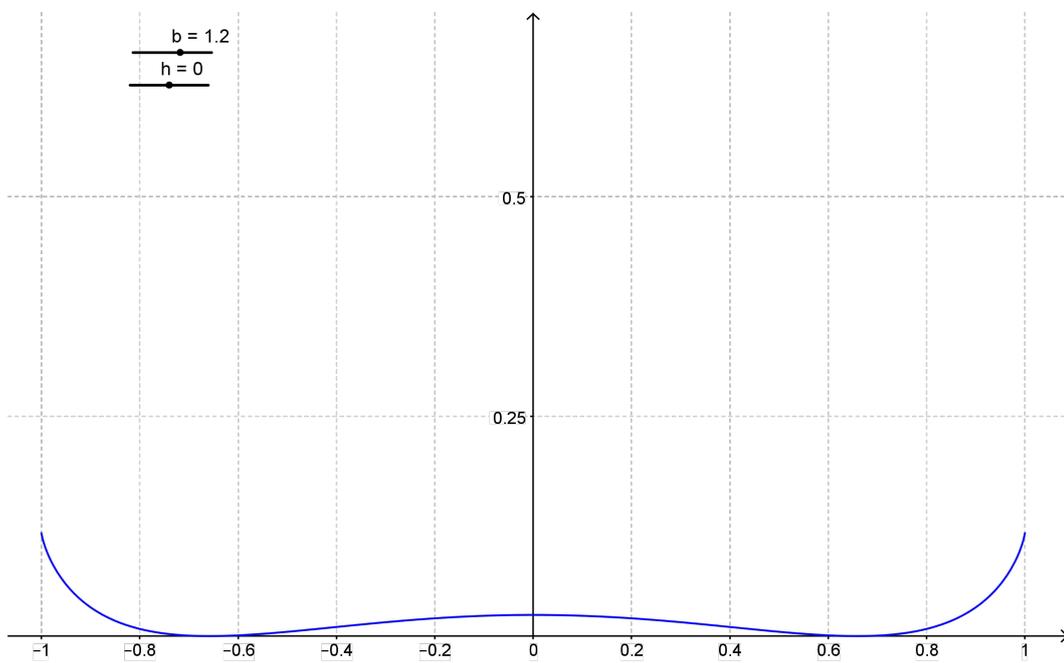


Abbildung 3.4.: Ratenfunktion für den Fall 3.

Abbildung 3.5.: Ratenfunktion im Grenzfall: $\beta = 1$

Abbildung 3.6.: Ratenfunktion im Grenzfall: $\beta > 1$

A. Anhang

A.1. Weitere Erläuterungen

Über das Prinzip der großen Abweichung von m_N , kann Grenzwert der freien Energie pro Teilchen berechnet werden. Mit der freien Energie können viele Eigenschaften des Systems ermittelt werden. Beispielsweise entspricht die Ableitung der freien Energie pro Teilchen nach h genau die erwartete Magnetisierung. Die freie Energie im Curie-Weiss Modell ist dabei definiert als

$$F_{N,\beta,h} := -\frac{1}{\beta} \log(Z_{N,\beta,h}).$$

Wir betrachten den Grenzwert der freien Energie pro Teilchen

$$\begin{aligned} f_{\beta,h} &:= \lim_{N \rightarrow \infty} \frac{1}{N} \cdot F_{N,\beta,h} \\ &= \lim_{N \rightarrow \infty} -\frac{1}{\beta N} \cdot \log(Z_{N,\beta,h}) \\ &= \lim_{N \rightarrow \infty} -\frac{1}{\beta N} \cdot \log \left(\sum_{\sigma \in \{-1,1\}^N} e^{\beta \frac{N}{2} \cdot [m_N(\sigma)]^2 + N\beta h \cdot m_N(\sigma)} \right). \end{aligned}$$

Zunächst benötigen wir einen Satz für Folgen von Zufallsvariablen, die einem Prinzip der Großen Abweichung genügt.

Satz A.1 (Lemma von Varadhan).

Genügt die Folge X_1, X_2, \dots einem Prinzip der Großen Abweichung mit Geschwindigkeit n und Ratenfunktion I , so gilt für jede stetige und beschränkte Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E} [e^{n \cdot f(X_n)}] = \sup_{x \in \mathbb{R}} [f(x) - I(x)].$$

Beweis.

Die \mathbb{R} -wertige Funktion f ist beschränkt und stetig. Durch die Stetigkeit (Vergleiche ϵ - δ -Kriterium) finden wir für jedes $\delta > 0$ eine Überdeckung von \mathbb{R} mit endlich vielen abgeschlossenen Mengen A_1, \dots, A_M mit der Eigenschaft, dass

$$\begin{aligned} \sup_{y \in A_i} f(y) - \inf_{y \in A_i} f(y) &\leq \delta \text{ für alle } i = 1, \dots, M \\ \Leftrightarrow \sup_{y \in A_i} f(y) &\leq \inf_{y \in A_i} f(y) + \delta \text{ für alle } i = 1, \dots, M \end{aligned}$$

Nun kann folgende Abschätzung durch die Überdeckung (keine disjunkte Partition) gemacht werden

$$\begin{aligned} \mathbb{E}[e^{n \cdot f(X_n)}] &= \int_{\Omega} e^{n \cdot f(X_n)} d\mathbb{P}(\omega) \\ &= \int_{\mathbb{R}} e^{n \cdot f(y)} d\mathbb{P}^{X_n}(y) \\ &\leq \int_{\mathbb{R}} \sum_{i=1}^M e^{n \cdot f(y)} \cdot \mathbb{1}_{A_i}(y) d\mathbb{P}^{X_n}(y) \\ &= \sum_{i=1}^M \int_{\mathbb{R}} e^{n \cdot f(y)} \cdot \mathbb{1}_{A_i}(y) d\mathbb{P}^{X_n}(y) \\ &\leq \sum_{i=1}^M \int_{\mathbb{R}} e^{n \cdot \sup_{y \in A_i} f(y)} \cdot \mathbb{1}_{A_i}(y) d\mathbb{P}^{X_n}(y) \\ &\leq \sum_{i=1}^M \int_{\mathbb{R}} e^{n \cdot (\inf_{y \in A_i} f(y) + \delta)} \cdot \mathbb{1}_{A_i}(y) d\mathbb{P}^{X_n}(y) \\ &= \sum_{i=1}^M e^{n \cdot (\inf_{y \in A_i} f(y) + \delta)} \cdot \mathbb{P}(X_n \in A_i) \end{aligned}$$

Also wurde gezeigt, dass

$$\mathbb{E}[e^{n \cdot f(X_n)}] \leq \sum_{i=1}^M e^{n \cdot (\inf_{y \in A_i} f(y) + \delta)} \cdot \mathbb{P}(X_n \in A_i)$$

Weil der Logarithmus monoton wachsend ist gilt außerdem

$$\frac{1}{n} \log(\mathbb{E}[e^{n \cdot f(X_n)}]) \leq \frac{1}{n} \log\left(\sum_{i=1}^M e^{n \cdot (\inf_{y \in A_i} f(y) + \delta)} \cdot \mathbb{P}(X_n \in A_i)\right)$$

Des weiteren gilt für zwei Folgen mit $a_n \leq b_n$, dass auch $\limsup_{n \rightarrow \infty} a_n \leq \limsup_{n \rightarrow \infty} b_n$ ist.

Dadurch gelangen wir zu:

$$\begin{aligned}
& \limsup_{n \rightarrow \infty} \frac{1}{n} \log (\mathbb{E} [e^{n \cdot f(X_n)}]) \\
& \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \left(\sum_{j=1}^M e^{n \cdot (\inf_{y \in A_j} f(y) + \delta)} \cdot \mathbb{P}(X_n \in A_j) \right) \\
& \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \left(M \cdot \sup_{1 \leq j \leq M} \left(e^{n \cdot (\inf_{y \in A_j} f(y) + \delta)} \cdot \mathbb{P}(X_n \in A_j) \right) \right) \\
& \leq \underbrace{\limsup_{n \rightarrow \infty} \frac{1}{n} \log(M)}_{=0} + \limsup_{n \rightarrow \infty} \frac{1}{n} \log \left(\sup_{1 \leq j \leq M} \left(e^{n \cdot (\inf_{y \in A_j} f(y) + \delta)} \cdot \mathbb{P}(X_n \in A_j) \right) \right) \\
& \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \sup_{1 \leq j \leq M} \left(\log \left(e^{n \cdot (\inf_{y \in A_j} f(y) + \delta)} \cdot \mathbb{P}(X_n \in A_j) \right) \right) \\
& \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \sup_{1 \leq j \leq M} \left(n \cdot \left(\inf_{y \in A_j} f(y) + \delta \right) + \log (\mathbb{P}(X_n \in A_j)) \right) \\
& = \limsup_{n \rightarrow \infty} \sup_{1 \leq j \leq M} \left(\inf_{y \in A_j} f(y) + \delta + \frac{1}{n} \log (\mathbb{P}(X_n \in A_j)) \right) \\
& \leq \sup_{1 \leq j \leq M} \limsup_{n \rightarrow \infty} \left(\inf_{y \in A_j} f(y) + \delta + \frac{1}{n} \log (\mathbb{P}(X_n \in A_j)) \right) \\
& = \sup_{1 \leq j \leq M} \left(\inf_{y \in A_j} f(y) + \delta + \limsup_{n \rightarrow \infty} \frac{1}{n} \log (\mathbb{P}(X_n \in A_j)) \right)
\end{aligned}$$

Durch das Prinzip der Großen Abweichung gilt für die abgeschlossenen Mengen A_j

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log (\mathbb{P}(X_n \in A_j)) \leq - \inf_{y \in A_j} I(y)$$

Setzen wir dies in der obigen Ungleichung ein, erhalten wir

$$\begin{aligned}
\limsup_{n \rightarrow \infty} \frac{1}{n} \log (\mathbb{E} [e^{n \cdot f(X_n)}]) & \leq \sup_{1 \leq j \leq M} \left(\inf_{y \in A_j} f(y) + \delta - \inf_{y \in A_j} I(y) \right) \\
& \leq \sup_{1 \leq j \leq M} \left(\sup_{y \in A_j} (f(y) - I(y)) + \delta \right) \\
& = \sup_{y \in \mathbb{R}} (f(y) - I(y)) + \delta
\end{aligned}$$

Für δ gegen 0 folgt nun die eine Seite der Abschätzung:

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log (\mathbb{E} [e^{n \cdot f(X_n)}]) \leq \sup_{y \in \mathbb{R}} (f(y) - I(y))$$

Bemerkung A.2.

Das ϵ - δ -Kriterium für Stetigkeit lautet

$$\begin{aligned} f : D \rightarrow \mathbb{R} \text{ ist stetig in } x \in D \\ \Leftrightarrow \forall \epsilon > 0 \exists \delta > 0 \text{ sodass } \forall y \in D \text{ mit } |x - y| \leq \delta \Rightarrow |f(x) - f(y)| \leq \epsilon \end{aligned}$$

Da f stetig ist, können wir für jeden Punkt $y_0 \in \mathbb{R}$ und für jedes $\epsilon > 0$ eine δ -Umgebung $U_\delta^\epsilon(y_0)$ finden, sodass für $y \in U_\delta^\epsilon(y_0)$ gilt

$$|f(y_0) - f(y)| < \epsilon \Rightarrow f(y) \geq f(y_0) - \epsilon$$

Daraus können wir nun abschätzen

$$\mathbb{E} [e^{n \cdot f(X_n)}] \geq \int_{U_\epsilon^\delta(y_0)} e^{n \cdot f(X_n)} d\mathbb{P} \geq e^{n \cdot (f(y_0) - \epsilon)} \cdot \mathbb{P}(X_n \in U_\epsilon^\delta(y_0))$$

Nutzen wir hier die Ungleichung aus dem LDP:

$$-\inf_{x \in \mathring{A}} I(x) \leq \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(X_n \in \mathring{A})$$

so erhalten wir

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{1}{n} \log (\mathbb{E} [e^{n \cdot f(X_n)}]) &\geq \liminf_{n \rightarrow \infty} \frac{1}{n} [\log (e^{n \cdot (f(y_0) - \epsilon)} \cdot \mathbb{P}(X_n \in U_\epsilon^\delta(y_0)))] \\ &= \liminf_{n \rightarrow \infty} \frac{1}{n} [(n \cdot (f(y_0) - \epsilon) + \log (\mathbb{P}(X_n \in U_\epsilon^\delta(y_0)))] \\ &= \liminf_{n \rightarrow \infty} \left[(f(y_0) - \epsilon) + \frac{1}{n} \log (\mathbb{P}(X_n \in U_\epsilon^\delta(y_0))) \right] \\ &= (f(y_0) - \epsilon) + \liminf_{n \rightarrow \infty} \frac{1}{n} \log (\mathbb{P}(X_n \in U_\epsilon^\delta(y_0))) \\ &\geq (f(y_0) - \epsilon) - \inf_{x \in U_\epsilon^\delta(y_0)} I(x) \\ &= f(y_0) - \epsilon - \inf_{x \in U_\epsilon^\delta(y_0)} I(x) \\ &\geq f(y_0) - I(y_0) - \epsilon \end{aligned}$$

Diese Ungleichung gilt für alle $y \in \mathbb{R}$ und alle $\epsilon > 0$ damit ist auch die andere Abschätzung gezeigt, d.h.

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log (\mathbb{E} [e^{n \cdot f(X_n)}]) \geq \sup_{y \in \mathbb{R}} (f(y) - I(y))$$

$$\begin{aligned} \sup_{y \in \mathbb{R}} (f(y) - I(y)) &\geq \limsup_{n \rightarrow \infty} \frac{1}{n} \log (\mathbb{E} [e^{n \cdot f(X_n)}]) \\ &\geq \liminf_{n \rightarrow \infty} \frac{1}{n} \log (\mathbb{E} [e^{n \cdot f(X_n)}]) \geq \sup_{y \in \mathbb{R}} (f(y) - I(y)) \end{aligned}$$

□

$f_{\beta,h}$ besitzt keine stochastischen Elemente, da wir die Summe über alle $\sigma \in \{-1, 1\}^N$ betrachten. Deshalb können wir $f_{\beta,h}$ mit dem Satz A.1 berechnen, indem wir die Funktion f passend wählen und m_N als Folge verwenden. Hierbei besteht der einzige Trick darin, dass wir den Erwartungswert aus dem Satz umformen und dadurch die gewünschte Summe aus $f_{\beta,h}$ erhalten und zusätzlich ein $\frac{1}{2}^N$ durch die Wahrscheinlichkeit von \mathbb{P} bekommen.

Satz A.3.

Für jede inverse Temperatur $\beta > 0$ und jedes magnetische Feld h ist

$$\begin{aligned} f_{\beta,h} &= \lim_{N \rightarrow \infty} -\frac{1}{\beta N} \cdot \log (Z_{N,\beta,h}) \\ &= \inf_{m \in [-1,1]} \left(-\frac{m^2}{2} - hm - \frac{1}{\beta} (\log 2 - I(m)) \right) \\ &= - \sup_{m \in [-1,1]} \left(\frac{m^2}{2} + hm + \frac{1}{\beta} (\log 2 - I(m)) \right) \end{aligned}$$

Hierbei ist I die Ratenfunktion für die Partialsummen der transformierten Bernoulli-verteilten Zufallsvariablen, d.h.

$$I(x) = \frac{1+x}{2} \cdot \log(1+x) + \frac{1-x}{2} \cdot \log(1-x)$$

Beweis.

Wir möchten Satz A.1 verwenden mit folgenden Elementen:

$$\begin{aligned} m_N(\sigma) &= \frac{1}{N} \sum_{i=1}^N \sigma_i \\ f(x) &= \frac{\beta}{2} x^2 + \beta h x \end{aligned}$$

Nach Bemerkung 3.3 genügt die Folge m_1, m_2, \dots dem Prinzip der Großen Abweichung mit Geschwindigkeit N und Ratenfunktion I . Die Funktion f ist stetig und auf dem Intervall $[-1, 1]$ beschränkt. Durch folgende Umformung wird schnell deutlich, wie der

Satz A.1 angewendet werden kann:

$$\begin{aligned}
\mathbb{E} [e^{N \cdot f(m_N)}] &= \sum_{\sigma \in \{\pm 1\}^n} e^{N \cdot f(m_N(\sigma))} \cdot \underbrace{\mathbb{P}(X_1 = \sigma_1, \dots, X_N = \sigma_N)}_{=2^{-N}} \\
&= 2^{-N} \cdot \sum_{\sigma \in \{\pm 1\}^N} e^{N \cdot f(m_N(\sigma))} \\
&= 2^{-N} \cdot \sum_{\sigma \in \{\pm 1\}^N} e^{\beta \frac{N}{2} \cdot [m_N(\sigma)]^2 + N\beta h \cdot m_N(\sigma)}
\end{aligned}$$

Erinnern wir uns an

$$Z_{N,\beta,h} = \sum_{\sigma \in \{\pm 1\}^N} e^{\beta \frac{N}{2} \cdot [m_N(\sigma)]^2 + N\beta h \cdot m_N(\sigma)},$$

so erhalten wir

$$\mathbb{E} [e^{N \cdot f(m_N)}] = 2^{-N} \cdot Z_{N,\beta,h}$$

Daher gilt nun:

$$\begin{aligned}
\lim_{N \rightarrow \infty} -\frac{1}{\beta N} \cdot \log(Z_{N,\beta,h}) &= -\frac{1}{\beta} \lim_{N \rightarrow \infty} \frac{1}{N} \cdot \log(2^N \cdot 2^{-N} \cdot Z_{N,\beta,h}) \\
&= -\frac{1}{\beta} \lim_{N \rightarrow \infty} \frac{1}{N} \cdot \log(2^N \cdot \mathbb{E} [e^{N \cdot f(m_N)}]) \\
&= -\frac{1}{\beta} \left[\lim_{N \rightarrow \infty} \frac{1}{N} (N \cdot \log(2) + \log(\mathbb{E} [e^{N \cdot f(m_N)}])) \right] \\
&= -\frac{1}{\beta} \left[\log(2) + \lim_{N \rightarrow \infty} \frac{1}{N} \cdot \log(\mathbb{E} [e^{N \cdot f(m_N)}]) \right] \\
&= -\frac{1}{\beta} \left[\log(2) + \sup_{m \in [-1,1]} \left(\frac{\beta m^2}{2} + \beta h m - I(m) \right) \right] \\
&= -\frac{1}{\beta} \log(2) - \sup_{m \in [-1,1]} \left(\frac{m^2}{2} + h m - \frac{1}{\beta} I(m) \right) \\
&= - \sup_{m \in [-1,1]} \left(\frac{m^2}{2} + h m - \frac{1}{\beta} I(m) + \frac{1}{\beta} \log(2) \right) \\
&= - \sup_{m \in [-1,1]} \left(\frac{m^2}{2} + h m + \frac{1}{\beta} (\log(2) - I(m)) \right) \\
&= \inf_{m \in [-1,1]} \left(-\frac{m^2}{2} - h m - \frac{1}{\beta} (\log(2) - I(m)) \right)
\end{aligned}$$

□

Mit Satz A.1 können wir auch den nächsten Satz beweisen, der uns die Möglichkeit bietet, ein Prinzip der Großen Abweichungen für das Bildmaß von m_N unter dem Gibbsmaß zu beweisen.

Satz A.4.

Sei X_1, X_2, \dots eine Folge die einem Prinzip der Großen Abweichungen mit Geschwindigkeit n und Ratenfunktion I genügt und $f : \mathbb{R} \rightarrow \mathbb{R}$ eine stetige und beschränkte Funktion. Sei \mathbb{P}_n^f ein Wahrscheinlichkeitsmaß, welches mit

$$J_n(A) = \int_A e^{n \cdot f(x)} d\mathbb{P}^{X_n}(x), \text{ für eine Menge } A \in \mathcal{B}$$

definiert ist durch

$$\mathbb{P}_n^f(A) = \frac{J_n(A)}{J_n(\mathbb{R})} = \int_A \frac{e^{n \cdot f(x)}}{\int_{\mathbb{R}} e^{n \cdot f(y)} d\mathbb{P}^{X_n}(y)} d\mathbb{P}^{S_n}(x).$$

Dann genügt die Folge $\mathbb{P}_1^f, \mathbb{P}_2^f, \dots$ von Wahrscheinlichkeitsmaßen einem Prinzip der Großen Abweichung mit Geschwindigkeit n und Ratenfunktion

$$I^f(x) = I(x) - f(x) + \sup_{y \in \mathbb{R}} [f(y) - I(y)].$$

Beweis.

Nach Definition erhalten wir

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n^f(A) &= \limsup_{n \rightarrow \infty} \frac{1}{n} \log \left(\frac{J_n(A)}{J_n(\mathbb{R})} \right) \\ &= \limsup_{n \rightarrow \infty} \left[\frac{1}{n} \log (J_n(A)) - \frac{1}{n} \log (J_n(\mathbb{R})) \right] \\ &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log (J_n(A)) + \limsup_{n \rightarrow \infty} -\frac{1}{n} \log (J_n(\mathbb{R})) \\ &= \limsup_{n \rightarrow \infty} \frac{1}{n} \log J_n(A) - \liminf_{n \rightarrow \infty} \frac{1}{n} \log J_n(\mathbb{R}), \end{aligned}$$

sowie analog für den Limes Inferior:

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n^f(A) &= \liminf_{n \rightarrow \infty} \frac{1}{n} \log \left(\frac{J_n(A)}{J_n(\mathbb{R})} \right) \\ &= \liminf_{n \rightarrow \infty} \left[\frac{1}{n} \log (J_n(A)) - \frac{1}{n} \log (J_n(\mathbb{R})) \right] \\ &\geq \liminf_{n \rightarrow \infty} \frac{1}{n} \log J_n(A) - \limsup_{n \rightarrow \infty} \frac{1}{n} \log J_n(\mathbb{R}). \end{aligned}$$

Aus dem vorherigen Satz folgt die Gleichheit für den Limes Superior und den Limes Inferior für den zweiten Summanden mit

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \log J_n(\mathbb{R}) &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E} [e^{n \cdot f(X_n)}] \\ &= \sup_{y \in \mathbb{R}} [f(y) - I(y)] \end{aligned}$$

Daher haben wir

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n^f(A) &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log J_n(A) - \limsup_{n \rightarrow \infty} \frac{1}{n} \log J_n(\mathbb{R}) \\ &= \limsup_{n \rightarrow \infty} \frac{1}{n} \log J_n(A) - \sup_{y \in \mathbb{R}} [f(y) - I(y)] \\ \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n^f(A) &\geq \liminf_{n \rightarrow \infty} \frac{1}{n} \log J_n(A) - \liminf_{n \rightarrow \infty} \frac{1}{n} \log J_n(\mathbb{R}) \\ &= \liminf_{n \rightarrow \infty} \frac{1}{n} \log J_n(A) - \sup_{y \in \mathbb{R}} [f(y) - I(y)] \end{aligned}$$

Nun erhält man die Aussage des Satzes unter Verwendung der beiden Ungleichungen für ein $A \in \mathcal{B}$:

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n} \log J_n(\bar{A}) &\leq - \inf_{x \in \bar{A}} [I(x) - f(x)] \\ \liminf_{n \rightarrow \infty} \frac{1}{n} \log J_n(\dot{A}) &\geq - \inf_{x \in \dot{A}} [I(x) - f(x)] \end{aligned}$$

Diese Ungleichungen werden nicht bewiesen. Somit erhalten wir

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n^f(\bar{A}) &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log J_n(\bar{A}) - \sup_{y \in \mathbb{R}} [f(y) - I(y)] \\ &\leq - \inf_{x \in \bar{A}} [I(x) - f(x)] - \sup_{y \in \mathbb{R}} [f(y) - I(y)] \\ &= - \left[\inf_{x \in \bar{A}} [I(x) - f(x)] + \sup_{y \in \mathbb{R}} [f(y) - I(y)] \right] \end{aligned}$$

$$\begin{aligned}
&= - \inf_{x \in \bar{A}} \left[I(x) - f(x) + \sup_{y \in \mathbb{R}} [f(y) - I(y)] \right] \\
\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n^f(\dot{A}) &\geq \liminf_{n \rightarrow \infty} \frac{1}{n} \log J_n(\dot{A}) - \sup_{y \in \mathbb{R}} [f(y) - I(y)] \\
&\geq - \inf_{x \in \dot{A}} [I(x) - f(x)] - \sup_{y \in \mathbb{R}} [f(y) - I(y)] \\
&= - \inf_{x \in \dot{A}} \left[I(x) - f(x) + \sup_{y \in \mathbb{R}} [f(y) - I(y)] \right]
\end{aligned}$$

Somit ergibt sich als Ratenfunktion:

$$I^f(x) = I(x) - f(x) + \sup_{y \in \mathbb{R}} [f(y) - I(y)]$$

□