

Zahl der
metastabilen Zustände
bei einem Neuronalen Netzwerk
mit Projektorkopplungen

Diplomarbeit

von
Peter Kuhlmann
aus
Ehringshausen

Institut für Theoretische Physik III
der
Justus-Liebig-Universität Gießen

1990

Inhaltsverzeichnis

1	Einleitung: Spingläser und Neuronale Netzwerke	4
1.1	Vom Ising-Modell über Spin-Gläser zum Hopfield-Modell . . .	5
1.2	Problemstellung	11
1.3	Inhaltsangabe	11
2	Berechnung der Zahl der metastabilen Zustände beim SK-Modell und beim Hopfield-Modell	12
2.1	Dynamik in Spinsystemen; metastabile Zustände	13
2.1.1	Einzelspinflip-Dynamik nach Glauber	15
2.1.2	Parallelfip-Dynamik nach Glauber	17
2.2	Allgemeines Berechnungsverfahren für die Zahl der metastabilen Zustände bei $\tau = 0$	18
2.3	Metastabile Zustände beim SK-Modell	19
2.3.1	Das Ergebnis von Tanaka und Edwards für das SK-Modell	19
2.3.2	Frustration	20
2.3.3	Selbstmittelungseigenschaft von $h = \frac{1}{N} \ln \mathcal{N}$	22
2.3.4	Ergebnisse von Bray und Moore und von Roberts für das SK-Modell	24
2.3.5	Zweierzyklen beim SK-Spinglas	26
2.4	Metastabile Zustände beim Hopfield-Modell	27
3	Ein Neuronales Netzwerk mit Projektorkopplungen	32
3.1	Die Projektormatrix	33
3.2	Eigenschaften des Modells	36
3.3	Selbstmittelungseigenschaft der Determinanten der Korrelationsmatrix \mathbf{C}	39

3.3.1	Der direkte Mittelwert $\langle\langle(\det \mathbf{C})^{-\frac{1}{2}}\rangle\rangle$	40
3.3.2	Die Replica-Rechnung für $\frac{1}{N}\langle\langle\ln(\det \mathbf{C})^{-\frac{1}{2}}\rangle\rangle$	42
3.4	Berechnung der Wahrscheinlichkeitsverteilung für J_{ii}	45
4	Berechnung der Zahl der metastabilen Zustände bei einem Neuronalen Netzwerk mit Projektorkopplungen	49
4.1	Der Rechenweg bis zum allgemeinen Ergebnis $\langle\langle\mathcal{N}(\alpha, g, \gamma, \varepsilon)\rangle\rangle$	51
4.1.1	Rechenweg bis zur Mittelung über die Muster	52
4.1.2	Entkopplung der Variablen in i und μ ; mehrdimensionale Sattelpunktintegration	57
4.1.3	Das allgemeine Endergebnis $\langle\langle\mathcal{N}(\alpha, g, \gamma, \varepsilon)\rangle\rangle$	61
4.2	Das Ergebnis $\langle\langle\mathcal{N}(\alpha, g, \gamma, \varepsilon_{max})\rangle\rangle$	62
4.2.1	Rechenweg	62
4.2.2	Das Endergebnis $\langle\langle\mathcal{N}(\alpha, g, \gamma, \varepsilon_{max})\rangle\rangle$, Sattelpunktgleichungen und Liste aller Hilfsvariablen	63
4.2.3	Grenzwerte und Näherungen im Fall $g \rightarrow 0$	64
4.3	Das Ergebnis $\langle\langle\mathcal{N}(\alpha, g = \frac{1}{2}, \gamma, \varepsilon_{max})\rangle\rangle$	67
4.3.1	Das Ergebnis $\langle\langle\mathcal{N}(\alpha, g = \frac{1}{2}, \gamma, \varepsilon_{max})\rangle\rangle$; Sattelpunktgleichung und Liste aller Hilfsvariablen	67
4.3.2	Grenzwerte im Fall $\alpha \rightarrow 0$	68
4.3.3	Grenzwerte im Fall $\alpha \rightarrow 1$	69
4.4	Das Ergebnis $\langle\langle\mathcal{N}(\alpha, g = \frac{1}{2}, \gamma, \varepsilon)\rangle\rangle$	71
5	Zahl der metastabilen Zustände bei einem Netzwerk mit Projektorkopplungen; Auswertung der Ergebnisse	73
5.1	Ergebnisse für $\langle\langle\mathcal{N}(\alpha, g = \frac{1}{2}, \gamma, \varepsilon_{max})\rangle\rangle$	75
5.1.1	Das Ergebnis für $f(\alpha, g = \frac{1}{2}, \gamma, \varepsilon_{max})$ in den Fällen $\gamma = 0$ und $\gamma = \alpha$	76
5.1.2	Der Effekt der Stabilität γ	82
5.2	Ergebnisse für $\langle\langle\mathcal{N}(\alpha, g, \gamma, \varepsilon_{max})\rangle\rangle$	84
5.2.1	Ergebnisse für $\gamma = 0$ und $\gamma = \alpha$	85
5.2.2	Der Grenzwert von $f(\alpha, g, \gamma, \varepsilon_{max})$ für $g \rightarrow 0$	93
5.3	Ergebnisse für $\langle\langle\mathcal{N}(\alpha, g = \frac{1}{2}, \gamma = \alpha, \varepsilon)\rangle\rangle$	96
5.4	Vergleich des Projektor-Netzwerks mit dem Hopfield-Modell im Fall $\alpha \rightarrow 0$	103

6	Zusammenfassung und Ausblick	106
6.1	Zusammenfassung	106
6.2	Ausblick	109
A	Gauß-Integrale und die Φ-Funktion	111
A.1	Gauß-Integrale	111
A.2	Die Φ -Funktion	112
B	Zahl der metastabilen Zustände beim Hopfield-Modell	113
C	Nebenrechnungen zu Kapitel 4	119
C.1	Die Mittelung über die σ_j^μ	119
C.2	Berechnung des Gauß-Integrals zur Potenz $p - 1$	120
C.3	Berechnung des Gauß-Integrals zur Potenz Ng	120
C.4	Vereinfachung des Exponenten Nf am Sattelpunkt durch Ableiten	121
C.5	Das Ergebnis $\langle\langle\mathcal{N}(\alpha, g, \gamma, \varepsilon_{max})\rangle\rangle$	122
C.6	Herleitung der Sattelpunktgleichungen zur Berechnung von $\langle\langle\mathcal{N}(\alpha, g, \gamma, \varepsilon_{max})\rangle\rangle$	124
C.7	Näherungen für $\langle\langle\mathcal{N}(\alpha, g = \frac{1}{2}, \gamma, \varepsilon_{max})\rangle\rangle$	128
C.8	Berechnung von $\langle\langle\mathcal{N}(\alpha, g = \frac{1}{2}, \gamma, \varepsilon)\rangle\rangle$	129
D	Ein Berechnungsverfahren für die Zahl der Zweierzyklen	133

Kapitel 1

Einleitung: Spingläser und Neuronale Netzwerke

Um eine allgemeine Definition des Neuronalen Netzwerks zu umgehen, befasse ich mich hier nur mit solchen Neuronalen Netzwerken, die auf dem Ising-Modell beruhen. Sie werden auch „Neuronale Netzwerke vom Spinglas-Typ“ genannt. Sie haben durch die bahnbrechende Arbeit von Hopfield [Ho82] einen enormen Aufschwung erfahren. Immer mehr Physiker und Ingenieure sind im Laufe der Zeit mit ihnen in Berührung gekommen und arbeiten nun auf dem Gebiet der Neuronalen Netzwerke. Aktuelle Informationen über Anwendungen von Neuronalen Netzwerken und besonders über andere Arten Neuronaler Netzwerke findet man in der Zeitschrift „Neural Networks“ und in den Proceedings der jährlichen Konferenzen ([INNS89],[Eck+90]).

Die in dieser Einleitung betrachteten Neuronalen Netzwerke sind so beschaffen, daß sie eine Beschreibung durch die Statistische Physik (siehe [Reif]) ermöglichen. Dadurch werden mächtige Methoden zur Behandlung umfangreicher Systeme anwendbar. Bisher war die Aufgabe der Statistischen Physik, die Minima der Freien Energie F für physikalische Systeme zu finden. Das absolute Minimum von F wird im thermischen Gleichgewicht angenommen, während andere relative Minima mit sogenannten „lokalstabilen“ Zuständen identifiziert werden können. Eine Kenntnis von F in der Umgebung der Minima erlaubt dann physikalische Größen des Systems (wie z.B. Druck und Entropie) zu berechnen. Ist die Temperatur τ des physikalischen Systems Null, so ist F identisch mit der Energie E . Im Fall $\tau \rightarrow 0$ liefert die Statistische Physik also Aussagen über die Zustände mit niedrigsten Energien, die

wir als (stabile) Grundzustände bezeichnen.

Die Idee der Neuronalen Netzwerke kehrt nun gerade die obige Betrachtungsweise um. Hier sind die Grundzustände vorgegeben, und man fragt danach, wie ein System beschaffen sein muß, damit diese Grundzustände besonders stabil sind. Das zu den Grundzuständen passende System heißt Neuronales Netzwerk und die Grundzustände werden als Muster bezeichnet. Der Begriff Netzwerk rührt daher, daß das System aus vielen verkoppelten Ising–Spins oder Neuronen bestehen soll. Die Kopplungen zwischen den Neuronen bezeichnet man in Anlehnung an die Biologie als Synapsen. Eine weitere Rechtfertigung für das Adjektiv „neuronal“ erhalten wir, wenn wir uns die dynamische Relaxation des Systems von einem Nichtgleichgewichtszustand aus in einen stabilen Grundzustand vorstellen. Analog zu physikalischen Systemen ist wünschenswert, daß die Muster gewisse Startzustände, die in ihrer Nähe liegen, anziehen. Ein solches geschickt konstruiertes „Attraktornetzwerk“ ist also in der Lage, eine ganze Familie von Zuständen einem Muster zuzuordnen. Somit stellt es im Gegensatz zu herkömmlichen Computern einen assoziativen Speicher dar und weist damit eine Ähnlichkeit zum menschlichen Gehirn auf.

Ich konkretisiere nun diese allgemeinen Bemerkungen, stelle das Hopfield–Modell vor und führe an die Problemstellung dieser Arbeit heran. Auf eine neurophysiologische Diskussion und eine Betrachtung der geschichtlichen Entwicklung der Neuronalen Netzwerke wird weitgehend verzichtet, weil bereits umfassende Einführungen in das Gebiet existieren ([Am89], [Ge90], [Ru86], [Ri+90]). Zudem findet man in [DHS90] einen Überblick über aktuelle Forschungsergebnisse.

1.1 Vom Ising-Modell über Spin–Gläser zum Hopfield–Modell

Ich betrachte zunächst die verallgemeinerte Version des Ising–Modells ([Is25], [Huang]), um die in dieser Arbeit verwendeten Begriffe zentral einzuführen. Es handelt sich um ein makroskopisches System aus N diskreten Spins

$$S_i, \quad i \in \{1, \dots, N\}, \quad S_i \in \{-1, +1\},$$

das sich im Wärmebad der Temperatur τ befindet.

Um eine Beschreibung durch die Statistische Physik zu ermöglichen, betrachten wir später immer den „thermodynamischen Limes“ $N \rightarrow \infty$.

Je zwei Spins S_i und S_j wechselwirken miteinander über Kopplungen J_{ij} . Auf ein S_i wirkt das innere Feld h_i der anderen Spins

$$h_i = \sum_{j(\neq i)} J_{ij} S_j \quad . \quad (1.1)$$

Bei physikalischen Systemen sind Selbstkopplungen J_{ii} ausgeschlossen. Außerdem sind dort die Kopplungen J_{ij} immer symmetrisch.¹ Die Wechselwirkungspotentiale $S_i J_{ij} S_j$ für alle Paarungen, an denen S_i beteiligt ist, addiert man zur lokalen Energie

$$\lambda_i = S_i \sum_{j(\neq i)} J_{ij} S_j \quad (1.2)$$

auf². Setzt man also $J_{ij} = J_{ji}$ voraus, so ergibt sich die Gesamtenergie zu

$$E = -\frac{1}{2} \sum_{i,j(\neq i)} J_{ij} S_i S_j \quad . \quad (1.3)$$

Man kann noch äußere (Magnet-) Felder b_i betrachten, die einen weiteren Term $-\sum_{i=1}^N b_i S_i$ in E liefern. In Anlehnung an die klassische Mechanik bezeichnet man das von den Variablen S_i abhängige E auch als Hamiltonian \mathcal{H} .

Das Ising-Modell *selbst* ist ein Spezialfall des Modells Gl.(1.3) und erklärt magnetische Phänomene. So erhält man einen Ferromagneten, wenn man

$$J_{ij} \equiv \frac{J}{N} > 0$$

setzt. Man kann auch räumliche Strukturen von magnetischen Materialien modellieren, indem man jedes S_i nur mit seinen Nachbarn wechselwirken läßt, alle anderen J_{ij} also Null setzt. Ein Ferromagnet ist freilich nur ein

¹Bei Neuronalen Netzwerken können Selbstkopplungen jedoch zugelassen sein. Außerdem gibt es auch Neuronale Netzwerke mit asymmetrischen Kopplungen, wie z.B. das optimale Perzeptron.

²Verwirrenderweise wird λ_i auch oft als „inneres Feld“ bezeichnet. Ich halte mich in dieser Arbeit jedoch konsequent an die obige Bezeichnung.

schlechtes Neuronales Netz, weil er bei $\tau = 0$ nur zwei Grundzustände hat ($S_i = +1 \forall i$ oder $S_i = -1 \forall i$).

Betrachten wir nun die weitaus komplizierteren *Spingläser*, die auch in der Natur auftreten und experimentell wie theoretisch intensiv untersucht worden sind ([BiYo86],[Me+87],[KFA85]). Ihre zerklüftete Energielandschaft wird durch eine Unordnung in den Kopplungen bewirkt. In Abschnitt 2.3.2 veranschauliche ich diesen Effekt. Sherrington und Kirkpatrick haben ein „exakt“ lösbares Spinglas-Modell aufgestellt ([SK75],[SK78]), das wegen seiner Popularität auch kurz als „SK-Modell“ bezeichnet wird. In ihm sind die Kopplungen gaußverteilt mit einer Streuung $\sigma = \frac{J}{\sqrt{N}}$.³ Für die Wahrscheinlichkeitsdichte von J_{ij} gilt also

$$P(J_{ij}) = \sqrt{\frac{N}{2\pi J^2}} \exp\left(\frac{-N J_{ij}^2}{2J^2}\right) . \quad (1.4)$$

Es gilt wieder $J_{ij} = J_{ji}$ und $J_{ii} = 0$. Sherrington und Kirkpatrick berechneten die Grundzustandsenergie E_0 als Grenzwert des absoluten Minimums der Freien Energie F für $\tau \rightarrow 0$. Sie nahmen dazu an, daß die J_{ij} für den Beobachtungszeitraum konstant oder eingefroren („quenched“) bleiben und folglich als zuvor ausgewürfelt betrachtet werden können. Sie betrachteten den „thermodynamischen Limes“ $N \rightarrow \infty$ und setzten somit die *Freie Energie pro Spin f* als *selbstmittelnd* voraus (siehe dazu auch Abschnitt 2.3.3). Der „quenched average“ von f (siehe [BiYo86]) über die Wahrscheinlichkeitsverteilung (1.4) sollte einen typischen Wert für f liefern:

$$\langle\langle f(\{J_{ij}\}) \rangle\rangle = \left(\prod_{i < j} \int_{-\infty}^{+\infty} dJ_{ij} P(J_{ij}) \right) f(\{J_{ij}\}) \quad (1.5)$$

Dabei ist

$$f = \frac{F}{N} = -\frac{\tau}{N} \ln Z, \quad (1.6)$$

wobei

$$Z = \sum_{\{S_i\}} \exp(-\beta \mathcal{H}(\{S_i\})) \quad (1.7)$$

³Die Streuung muß mit $\frac{1}{\sqrt{N}}$ skaliert werden, um später die Extensivität der Freien Energie F zu garantieren. F muß nämlich für physikalische Systeme mit N skalieren.

die kanonische Zustandssumme ist und \mathcal{H} den Hamiltonian (Gl.(1.3)) darstellt. $\beta = \frac{1}{T}$ ist die reziproke Temperatur.

Um $\ln Z$ zu mitteln, verwendeten sie die Replica-Methode. Ihr Ergebnis für E_0 wurde von Parisi verbessert [Pa80].

Durch die bei den Spingläsern gesammelten Erfahrungen und insbesondere mit Hilfe der beim SK-Modell verwendeten Replica-Methode wurde das notwendige Rüstzeug gewonnen, um noch kompliziertere Modelle zu behandeln: Das Neuronale Netzwerk vom Spinglas-Typ soll p vorgegebene Muster

$$\underline{\xi}^\mu = (\xi_1^\mu, \dots, \xi_N^\mu)^T, \quad \mu \in \{1, \dots, p\}, \quad \xi_i^\mu \in \{-1, +1\},$$

als Grundzustände besitzen. Eine Kopplungsmatrix (J_{ij}) , die das unter gewissen Voraussetzungen leistet, ist schon lange bekannt und wurde von Little [Li74] und Hopfield [Ho82] aufgegriffen. Es handelt sich um die Hebbsche Lernregel ([He49]). Sie definiert die Kopplungen („Synapsen“) als

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu \quad \text{für } i \neq j \quad (1.8)$$

Das zugehörige Neuronale Netzwerk bezeichnen wir als *Hopfield-Modell*.

Amit, Gutfreund und Sompolinsky haben gezeigt, daß für endliche p und im Grenzwert $N \rightarrow \infty$ die Muster Grundzustände sind [Am+85a]. Sie gingen dabei analog zum SK-Modell vor und betrachteten die Muster als p vorgegebene, zufällig nach den unabhängigen Einzelwahrscheinlichkeiten $p\{\xi_i^\mu\} = \frac{1}{2}$ ausgewürfelte Zustände. Amit *et al.* erklärten die Stabilität der Muster mit dem Verschwinden des Rauschens in den inneren Feldern h_i der Muster [Am+87]. Für das innere Feld des ersten Musters gilt nämlich

$$h_i = \sum_{j(\neq i)} J_{ij} \xi_j^1 = \xi_i^1 (1 + \delta_i). \quad (1.9)$$

Dabei ist δ_i der Rauschterm

$$\delta_i = \frac{1}{N} \sum_{j(\neq i)} \sum_{\mu(\neq 1)} \xi_i^\mu \xi_j^\mu \xi_i^1 \xi_j^1. \quad (1.10)$$

Er stellt eine Zufallsvariable mit Varianz

$$\overline{(\delta_i)^2} = \frac{p-1}{N} \quad (1.11)$$

dar. Für endliche p zeigen die inneren Felder der Muster im Fall $N \rightarrow \infty$ damit wie gewünscht in Richtung der ξ_i^μ .

Macht man jedoch die für reale Modelle realistischere Annahme

$$\alpha = \frac{p}{N} = \text{constant}, \quad (1.12)$$

so ist zu erwarten, daß das Rauschen die Überhand gewinnt und die Muster nicht mehr stabil sind. Amit *et al.* führten eine Replica-Rechnung für diesen Fall durch und berechneten die über die Muster gemittelte Freie Energie pro Spin f ([Am+85b], [Am+87]):

$$\langle\langle f(\{\xi_i^\mu\}) \rangle\rangle = \frac{1}{2^{Np}} \sum_{\{\xi_i^\mu\}} f(\{\xi_i^\mu\}) . \quad (1.13)$$

Sie zeigten, daß das System Wiedererkennungseigenschaften hat, sofern α unter der *kritischen Speicherkapazität* von

$$\alpha_c \sim 0.14 \quad (1.14)$$

bleibt. Jedoch sind die Muster selbst wie erwartet instabil. Ihre Erkennung wird durch thermodynamisch stabile Zustände bewerkstelligt, die sich nur wenig von ihnen unterscheiden.

Neben diesen „retrieval states“ gibt es weitere Lösungen der Bestimmungsgleichungen für $\langle\langle f \rangle\rangle$. Es handelt sich um die *Spinglaszustände*. Sie sind zu allen Mustern unkorreliert, haben also *Überlapps*

$$m_\mu = \frac{1}{N} \sum_{i=1}^N S_i \xi_i^\mu = \mathcal{O}\left(\frac{1}{\sqrt{N}}\right) \quad \forall \mu . \quad (1.15)$$

Gleichzeitig weisen sie die wichtige Eigenschaft auf, daß ihr *Edwards–Anderson–Ordnungsparameter*

$$q = \frac{1}{N} \sum_{i=1}^N \langle S_i \rangle^2 \neq 0$$

ist. Dies ist das zentrale Merkmal der Spinglas-Phase. Die paramagnetische Phase

$$\langle S_i \rangle = 0 \quad \forall i$$

tritt erst bei hohen Temperaturen auf und spielt für unsere späteren Betrachtungen keine Rolle. Im Fall $\alpha \rightarrow \infty$ ergab sich als Energie der Spinglaszustände das E_0 aus der Replica-Rechnung von Sherrington und Kirkpatrick. Die J_{ij} können offenbar dann als unabhängige gaußverteilte Zufallsvariablen aufgefaßt werden. In der Tat ist schon bei endlichem α ein einzelnes J_{ij} gaußverteilt, denn der Ausdruck

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu \quad (1.16)$$

stellt für $i \neq j$ den zurückgelegten Weg eines „random walk“ [Reif] mit p Schritten dar. Der Mittelwert ist Null und die Streuung ist $\sqrt{\frac{p}{N}} = \sqrt{\alpha}$. Die Muster bewirken Korrelationen der J_{ij} , die für $\alpha \rightarrow \infty$ verschwinden.

Soviel zum Hopfield-Modell. Im Laufe der Zeit sind Neuronale Netzwerke vom Spinglas-Typ mit besseren Wiedererkennungseigenschaften und höheren Speicherkapazitäten entwickelt worden. Eines davon ist das in dieser Arbeit behandelte Projektornetzwerk (auch „Pseudoinverse“ genannt). Bei ihm sind die Muster thermodynamisch stabile Zustände und es gilt $\alpha_c = 1$. Das optimale Perzeptron ist ein weiterer prominenter Vertreter ([MiPa69],[AnBi89],[Op88]). Elizabeth Gardner hat hier mit einem neuen Rechenverfahren die kritische Speicherkapazität zu $\alpha_c = 2$ ausgerechnet [Ga88].

In der obigen Betrachtung fürs Hopfield-Modell habe ich nur die statischen Aspekte des Problems berücksichtigt. Es wurde nur mit Hilfe der Statistischen Physik gezeigt, daß die Gleichgewichtszustände den Mustern ähnlich sind. Dem Neuronalen Netzwerk fehlt noch eine Dynamik, die es auf die Gleichgewichtszustände zusteuert. Es zeigt sich (siehe Kapitel 2), daß die serielle Glauber-Dynamik Gleichgewichtszuständen des Modells (1.3) zustrebt und damit das Neuronale Netzwerk vom Spinglas-Typ zu einem assoziativen Speicher macht.

Auf dem Weg ins Gleichgewicht kann die serielle Glauber-Dynamik bei $\tau = 0$ in unerwünschten Fallen steckenbleiben. Von der Berechnung der Zahl der Fallen, die auch metastabile Zustände genannt werden, handelt diese Arbeit.

1.2 Problemstellung

Die Arbeit soll sich mit der Zahl \mathcal{N} der metastabilen Zustände bei einem Neuronalen Netzwerk mit Projektorkopplungen befassen. Der direkte Mittelwert $\langle\langle\mathcal{N}\rangle\rangle$ von \mathcal{N} über die zufällig verteilten Muster soll im thermodynamischen Limes $N \rightarrow \infty$ mit Hilfe der Sattelpunktmethode angegeben werden. $\langle\langle\mathcal{N}\rangle\rangle$ soll in Abhängigkeit vom Hamming–Abstand g zum ersten Muster betrachtet werden. Die Stabilität γ und die negative Energie pro Spin ε können als weitere Zwangsbedingungen mit in die Rechnung einbezogen werden.

1.3 Inhaltsangabe

Um deutlich zu machen, welche Rolle metastabile Zustände für die serielle $\tau = 0$ –Glauber–Dynamik spielen, wird diese in Kapitel 2 ausführlich dargestellt. Auf die Berechnung von $\langle\langle\mathcal{N}\rangle\rangle$ beim SK–Modell und beim Hopfield–Modell wird eingegangen. Im zweiten Fall wird ein neues Ergebnis für die Energieabhängigkeit von $\langle\langle\mathcal{N}\rangle\rangle$ vorgestellt.

Das dritte Kapitel befaßt sich generell mit dem Neuronalen Netzwerk mit Projektorkopplungen und führt die wichtigste dazu existierende Literatur auf. Es folgen erste analytische Rechnungen für das Modell.

Kapitel 4 enthält die gesamte Rechnung zur obigen Problemstellung. $\langle\langle\mathcal{N}\rangle\rangle$ wird mit Hilfe der Sattelpunktmethode in Abhängigkeit von den vier Parametern α, γ, g und ε berechnet. Für gewisse Kombinationen der Parameter werden die Sattelpunktgleichungen hergeleitet.

Das fünfte Kapitel stellt die Ergebnisse vor. Es handelt sich dabei meistens um kommentierte Diagramme von $\frac{1}{N} \ln(\langle\langle\mathcal{N}\rangle\rangle)$ in den verschiedenen Parameterbereichen. Die Arbeit schließt mit Kapitel 6, der Zusammenfassung und dem Ausblick.

Kapitel 2

Berechnung der Zahl der metastabilen Zustände beim SK-Modell und beim Hopfield-Modell

Ziel dieses Kapitels ist es, den Zusammenhang zwischen metastabilen Zuständen und der Dynamik des Neuronalen Netzes aufzuzeigen sowie einige Ergebnisse für das SK-Modell und das Hopfield-Modell darzustellen.

In Abschnitt 2.1 wird zunächst die hauptsächlich verwendete Dynamik für Spinsysteme vorgestellt: Die „Glauber-Dynamik“ hat den Vorteil, daß sie unter gewissen Voraussetzungen zur physikalischen Größe „Freie Energie“ in Beziehung gesetzt werden kann. Das Gleichgewicht der Glauber-Dynamik ist dann identisch mit dem thermodynamischen Gleichgewicht, und die Statistische Physik kann angewendet werden, um die stabilen Zustände der Dynamik zu finden. Es zeigt sich, daß die sequentielle $\tau = 0$ -Glauber-Dynamik *nur* metastabile Zustände als Fixpunkte besitzt. Ein Spinsystem bzw. ein Neuronales Netzwerk kann also auf dem Weg in einen Grundzustand bzw. in ein Muster in einem metastabilen Zustand gefangen werden. Um erforschen zu können, wie stark der Einfluß dieser metastabilen Zustände auf die Dynamik ist, ist es also wichtig, ihre Zahl zu kennen. Das Berechnungsverfahren wird in Abschnitt 2.2 vorgestellt. Einige Ergebnisse der Rechnungen zum SK-Modell bzw. zum Hopfield-Modell werden in Abschnitt 2.3 bzw. Abschnitt 2.4 geschildert.

2.1 Dynamik in Spinsystemen; metastabile Zustände

Betrachtet sei ein System aus N Spins $S_i \in \{-1, +1\}$ mit Kopplungen J_{ij} . Das innere Feld am Platz i ist

$$h_i = \sum_{j=1}^N J_{ij} S_j .$$

Man definiert nun einen dynamischen Prozeß mit diskreten Zeitschritten der Länge Δt für das System, die sogenannte *Glauber-Dynamik* [Gl63]. Sie stellt gewissermaßen eine Umkehrung der Monte Carlo-Methode [Me+53] dar: Die Monte Carlo-Methode wurde entwickelt, um kanonische Mittelwerte von Observablen A physikalischer Systeme näherungsweise zu berechnen. Man entwirft einen geeigneten Markoff-Prozeß [DoGr76], der nach einer Anlaufzeit typische Zustände I_{typ} für ein kanonisches Ensemble durchläuft. Der kanonische Mittelwert

$$\langle A \rangle = \sum_I \frac{e^{-\beta \mathcal{H}(I)}}{Z} A(I)$$

kann dann durch das arithmetische Mittel

$$\bar{A} = \frac{1}{M} \sum_{l=1}^M A(I_{\text{typ}}^{(l)})$$

angenähert werden.

Glauber ging umgekehrt vor, er definierte eine Monte Carlo-Dynamik für das betrachtete Spin-System im Wärmebad der Temperatur τ (zunächst sei $\tau > 0$ betrachtet): In einem Schritt des Markoff-Prozesses geht das Spin-System mit der Wahrscheinlichkeit $W(K \leftarrow I)$ vom Zustand $I = \underline{S}(t)$ in den Zustand $K = \underline{S}(t + \Delta t)$ über.

Für $P_I(n)$, die Wahrscheinlichkeit, daß sich das System zum Zeitpunkt $t = n\Delta t$ im Zustand I befindet, gilt die Mastergleichung [Am89]:

$$P_I(n+1) = P_I(n) + \sum_{K(\neq I)} [W(I \leftarrow K)P_K(n) - W(K \leftarrow I)P_I(n)] \quad (2.1)$$

Um den Zusammenhang mit der Statistischen Physik herzustellen, stellt man nun die Frage nach einer zeitlich konstanten Gleichgewichtsverteilung

P_I^{eq} . Man kann zeigen, daß im Falle der *detaillierten Balance* eine Gleichgewichtsverteilung existiert [Pe84], sofern gewisse Zusatzvoraussetzungen für $W(K \leftarrow I)$ gelten: Gibt es ein $G(I)$ mit

$$W(I \leftarrow K) G(K) = W(K \leftarrow I) G(I), \quad (2.2)$$

so gilt $P_I^{\text{eq}} \propto G(I)$ bis auf eine Normierungskonstante. Es reicht zum Nachweis nicht aus, daß G die Master-Gleichung erfüllt; man muß hauptsächlich zeigen, daß P_I^{eq} sich unvermeidlich von *jedem* Anfangszustand $\underline{S}(0)$ aus für $n \rightarrow \infty$ einstellt.

Im Gleichgewicht gilt nicht nur eine Gesamtbalance, sondern sogar die detaillierte Balance: Der „Teilfluß“ vom Zustand K zum Zustand I ist so groß wie der „Teilfluß“ vom Zustand I zum Zustand K :

$$P_K^{\text{eq}} W(I \leftarrow K) = P_I^{\text{eq}} W(K \leftarrow I) .$$

Dies entspricht Vorstellungen aus der Festkörperphysik [AsMe76] und der Theorie des schwarzen Körpers [Reif]. Es verwundert deshalb nicht, daß jetzt auch ein Hamiltonian $\mathcal{H}(I)$ existiert:

$$\mathcal{H}(I) := -\tau \ln G(I) . \quad (2.3)$$

P_I^{eq} ist die Boltzmannverteilung

$$P_I^{\text{eq}} = \frac{e^{-\beta \mathcal{H}(I)}}{Z} \quad (2.4)$$

mit der Zustandssumme

$$Z = \sum_I e^{-\beta \mathcal{H}(I)} .$$

Schließlich kann man noch zeigen, daß das Freie Energie-Funktional

$$\Phi(n) := \sum_I P_I(n) \mathcal{H}(I) + \tau \sum_I P_I(n) \ln P_I(n) \quad (2.5)$$

streng monoton mit n abnimmt, bis es sein absolutes Minimum, die Gleichgewichts-Freie-Energie

$$F = \Phi(n \rightarrow \infty) = -\tau \ln Z$$

erreicht; Φ ist also eine *Lyapunov-Funktion*, die den dynamischen Prozeß „regiert“ [Am89]. Für zwei Beispiele spezifiziere ich jetzt $W(I \leftarrow K)$ und betrachte noch den rauschfreien Fall $\tau = 0$.

2.1.1 Einzelspinflip–Dynamik nach Glauber

In jedem Schritt wird *ein* Spin S_i entweder zufällig oder zyklisch ausgesucht. Der neue Zustand $\underline{S}(t + \Delta t)$ unterscheidet sich also von $\underline{S}(t)$ höchstens durch den Spin S_i . Die Übergangswahrscheinlichkeit ist temperaturabhängig und beträgt

$$W(S_i(t + \Delta t) \leftarrow S_i(t)) = \frac{\exp(\beta S_i(t + \Delta t) h_i(t))}{\exp(\beta h_i(t)) + \exp(-\beta h_i(t))} . \quad (2.6)$$

Dabei ist $h_i(t) = \sum_{j=1}^N J_{ij} S_j(t)$. Im Fall $\tau \rightarrow 0$, d.h. $\beta \rightarrow \infty$ gilt einfach

$$S_i(t + \Delta t) = \text{sign}(h_i(t)) . \quad (2.7)$$

Die Einzelspinflipdynamik erfüllt detaillierte Balance, falls J_{ij} symmetrisch ist und keine Selbstkopplungen J_{ii} vorliegen.

$$\mathcal{H} = -\frac{1}{2} \sum_{i,j(\neq i)} J_{ij} S_i S_j \quad (2.8)$$

stellt dann einen Hamiltonian dar:

$$\begin{aligned} \frac{W(K \leftarrow I)}{W(I \leftarrow K)} &= \frac{W(-S_i(t) \leftarrow S_i(t))}{W(S_i(t) \leftarrow -S_i(t))} \\ &= \frac{\exp(-\beta S_i(t) h_i(t))}{\exp(\beta S_i(t) h_i(t))} \\ &= \exp(-2\beta S_i(t) h_i(t)) \end{aligned}$$

Es gilt also

$$\begin{aligned} \mathcal{H}(K) - \mathcal{H}(I) &= -\frac{1}{2} \left(\sum_{j(\neq i)} J_{ij} (-S_i) S_j + \sum_{k(\neq i)} J_{ki} S_k (-S_i) \right) \\ &\quad + 0 + \frac{1}{2} \sum_{j(\neq i)} J_{ij} S_i S_j + \frac{1}{2} \sum_{k(\neq i)} J_{ki} S_k S_i \\ &= +2 \sum_{j(\neq i)} J_{ij} S_i S_j, \text{ falls } J_{ij} \text{ symmetrisch ist.} \\ &\implies \mathcal{H}(K) - \mathcal{H}(I) = +2S_i h_i, \end{aligned} \quad (2.9)$$

falls keine Selbstkopplungen vorliegen. Also gilt

$$\frac{W(K \leftarrow I)}{W(I \leftarrow K)} = e^{-\beta(\mathcal{H}(K) - \mathcal{H}(I))} = \frac{G(K)}{G(I)}.$$

Liegen jedoch Selbstkopplungen vor, so ist nicht zu erwarten, daß ein Hamiltonian existiert. Dies ist auch physikalisch plausibel, da die Selbstkopplungen J_{ii} in der Physik nicht vorkommen.

Fixpunkte der energiegetriebenen $\tau = 0$ -Dynamik: Metastabile Zustände

Fixpunkte der $\tau = 0$ -Dynamik sind nach Gl.(2.7) sicher Zustände \underline{S} , für die gilt

$$S_i = \text{sign}(h_i) \quad \forall i \iff \lambda_i = S_i h_i > 0 \quad \forall i. \quad (2.10)$$

Man nennt solche Fixpunkte *metastabile Zustände*, da sie stabil gegen Einzelspin-Flips sind. Ist *zusätzlich* J_{ij} wie oben symmetrisch und sind alle $J_{ii} = 0$, so nimmt $\mathcal{H} = -\frac{1}{2} \sum_{i,j(\neq i)} J_{ij} S_i S_j$ von Schritt zu Schritt streng monoton ab und bleibt nur für metastabile Zustände konstant, denn im Falle eines Spin-Flips $S_i(t) \rightarrow S_i(t + \Delta t) = -S_i(t)$ gilt wegen Gl.(2.7) und Gl.(2.9)

$$\mathcal{H}(K(t + \Delta t)) - \mathcal{H}(I(t)) = +2S_i(t) h_i(t) = -2S_i(t + \Delta t) h_i(t) < 0.$$

Folglich ist \mathcal{H} unter den obigen Voraussetzungen eine Lyapunov-Funktion und die *einzigsten* Fixpunkte der Dynamik sind die metastabilen Zustände. Die absoluten Minima von \mathcal{H} bezeichnet man als Grundzustände. Startet man den dynamischen Prozeß von einem Grundzustand aus, so kann \mathcal{H} nicht weiter abnehmen. Fixpunkte müssen aber metastabile Zustände sein. Daraus folgt, daß Grundzustände besondere metastabile Zustände sind. Metastabile Zustände sind also sehr wichtig für die serielle $\tau = 0$ -Glauber-Dynamik. Sie sind „Fallen“ für die Dynamik: Das System gelangt in der Regel nicht bis in den Grundzustand, sondern bleibt vorher in einem metastabilen Zustand stehen.

Gibt es jedoch Selbstkopplungen $J_{ii} \neq 0$, so gilt nach Gl.(2.9)

$$\mathcal{H}(K(t + \Delta t)) - \mathcal{H}(I(t)) = 2S_i(t)h_i(t) - 2J_{ii}$$

und \mathcal{H} ist *auch* eine Lyapunov-Funktion, falls $J_{ii} \geq 0 \quad \forall i$ ist.

2.1.2 Parallelfliip–Dynamik nach Glauber

In jedem Schritt werden alle Spins S_i betrachtet. Jeder einzelne wird mit der Wahrscheinlichkeit $W(S_i(t + \Delta t) \leftarrow S_i(t))$ nach Gl.(2.6) geflippt. Somit kann der Zustand $I = \underline{S}(t)$ bei $\tau \neq 0$ in jeden anderen Zustand $K = \underline{S}(t + \Delta t)$ übergehen. Die Wahrscheinlichkeit für einen solchen Übergang ist natürlich das Produkt aller Wahrscheinlichkeiten aus Gl.(2.6). Für $\tau \rightarrow 0$ gilt analog Gl.(2.7) für jeden Spin S_i . Detaillierte Balance ist nach [Pe84] erfüllt, sobald die J_{ij} symmetrisch sind. $J_{ii} = 0$ ist nicht gefordert.

$$\mathcal{H}(I) := -\tau \sum_{i=1}^N \ln(\cosh \beta h_i) \quad (2.11)$$

ist dann der zugehörige Hamiltonian.

Fixpunkte der energiegetriebenen $\tau = 0$ –Dynamik: Metastabile Zustände und Zweierzyklen

Nach dem Satz von L’Hospital gilt

$$\lim_{\beta \rightarrow \infty} \mathcal{H}(I) = - \sum_{i=1}^N h_i \operatorname{sign}(h_i) =: E$$

Amit hat gezeigt [Am89], daß E eine Lyapunov–Funktion der parallelen Dynamik ist: E nimmt streng monoton mit t ab, bis das System *entweder* einen metastabilen Zustand *oder* einen Zweierzyklus trifft. Zweier–Zyklen sind „Blinkzustände“

$$\underline{S}(t + 2k\Delta t) = \underline{S}^{(1)}; \quad \underline{S}(t + (2k + 1)\Delta t) = \underline{S}^{(2)} \quad \forall k \in N_0$$

Es gibt hier keine höheren Zyklen. In Abschnitt 2.3.5 wird ein Hinweis darauf gegeben, daß Zweierzyklen für die parallele Dynamik von größerer Relevanz sind als metastabile Zustände.

Der Prozeß endet wiederum spätestens in einem der Zustände des Systems, für die E ein absolutes Minimum annimmt, also in einem der Grundzustände. Ein Grundzustand ist also entweder ein metastabiler Zustand, oder er wechselt sich mit einem anderen Grundzustand in einem Zweierzyklus ab.

2.2 Allgemeines Berechnungsverfahren für die Zahl der metastabilen Zustände bei $\tau = 0$

Die Berechnung der Zahl \mathcal{N} der metastabilen Zustände geschieht mit Hilfe der Heaviside-Funktion

$$\Theta(x) = \begin{cases} 1 & \text{für } x \geq 0 \\ 0 & \text{für } x < 0 \end{cases}$$

Man zählt alle Zustände \underline{S} , die $\lambda_i = S_i h_i \geq 0 \quad \forall i$ haben, mit Gewicht 1, alle anderen mit Gewicht 0 und bildet

$$\mathcal{N} = \sum_{\{\underline{S}_i\}} \prod_{i=1}^N \Theta(\lambda_i) \quad (2.12)$$

Um die diskreten Ausdrücke auf analytische Funktionen umzuschreiben, verwendet man die δ -Funktions-Darstellung der Θ -Funktion und setzt wiederum die Fourier-Darstellung der δ -Funktion ein (siehe z.B. [TaEd80]):

$$\Theta(x) = \int_0^{\infty} \delta(\lambda - x) d\lambda = \int_0^{\infty} \int_{-i\infty}^{+i\infty} \frac{d\mu}{2\pi i} e^{\mu(\lambda-x)} d\lambda. \quad (2.13)$$

Gegen Ende der Rechnung bildet man den thermodynamischen Limes und wertet die verbleibenden Integrale mit Hilfe der Sattelpunktmethode aus.

Des weiteren ist es möglich, in Gl.(2.12) zusätzliche Zwangsbedingungen einzubauen. Beispielsweise kann man danach fragen, wieviele metastabile Zustände im Energieintervall $[E, E + dE]$ liegen, falls für das betrachtete Modell ein Hamiltonian existiert. Analog zu der in [Reif] geschilderten Darwin-Fowler-Methode schreibt man für die Energiedichte $\mathcal{N}(E)$

$$\mathcal{N}(E) dE = \sum_{\{\underline{S}_i\}} \left(\prod_{i=1}^N \Theta(\lambda_i) \right) \delta(E - \mathcal{H}(\underline{S})) dE \quad (2.14)$$

Man verwendet dann wieder die Fourierdarstellung der δ -Funktion. Im folgenden Abschnitt 2.3 wird in \mathcal{N} immer der Einzelspinflip-Hamiltonian (2.8) verwendet. Es ist also wichtig darauf hinzuweisen, daß die *dann* folgenden Aussagen nur für serielle Dynamik gelten. Da bei der seriellen Dynamik Grundzustände spezielle metastabile Zustände sind, erwartet man, daß sie durch $\mathcal{N}(E)$ erfaßt werden.

2.3 Metastabile Zustände beim SK–Modell

Bei Spingläsern und neuronalen Netzwerken sind die J_{ij} zufällig verteilt (siehe Abschnitt 1.1), so daß man über die Verteilung mitteln muß, um eine analytische Aussage über die Zahl der metastabilen Zustände machen zu können. Die Mittelung über die J_{ij} wird wieder mit $\langle\langle \dots \rangle\rangle$ bezeichnet. Der Gedankengang in diesem Punkt wird zunächst kurz skizziert:

In Abschnitt 2.3.1 wird das Ergebnis von Tanaka und Edwards vorgestellt. Der direkte Mittelwert $\langle\langle \mathcal{N} \rangle\rangle$ hat exponentielle Form. Ursache dafür ist der Effekt der Frustration (siehe Abschnitt 2.3.2). Die Untersuchungen von Toulouse und Vannimenus führen auf den Ansatz $\mathcal{N} = e^{Nh}$, wobei h als intensiv angenommen wird. In Abschnitt 2.3.3 wird die Selbstmittelungseigenschaft von h behandelt und aufgezeigt, daß $e^{N\langle\langle h \rangle\rangle}$ anstelle von $\langle\langle \mathcal{N} \rangle\rangle$ eine bessere Aussage über \mathcal{N} macht. Die wichtige Jensensche Ungleichung

$$\langle\langle \mathcal{N} \rangle\rangle \geq e^{N\langle\langle h \rangle\rangle}$$

wird plausibilisiert. Nach dieser Vorbereitung schildere ich dann in Abschnitt 2.3.4 einige Ergebnisse von Bray und Moore und von Roberts für das SK–Modell. Abschließend wird in Abschnitt 2.3.5 kurz auf Zweierzyklen eingegangen und mitgeteilt, daß diese die entscheidenden Fallen für die parallele Dynamik darstellen.

2.3.1 Das Ergebnis von Tanaka und Edwards für das SK–Modell

Tanaka und Edwards haben unter anderem $\langle\langle \mathcal{N} \rangle\rangle$ und $\langle\langle \mathcal{N}(E) \rangle\rangle$ für Einzelspinflip¹ bei $\tau = 0$ nach dem in Abschnitt 2.2 geschilderten Verfahren für das SK–Modell mit symmetrischen Kopplungen berechnet [TaEd80]. Ihr Ergebnis ist

$$\langle\langle \mathcal{N} \rangle\rangle \propto e^{N \cdot 0.1992}, \quad (2.15)$$

und für $\frac{1}{N} \ln \langle\langle \mathcal{N}(\varepsilon) \rangle\rangle$ erhalten sie in etwa eine verschobene Gauß–Verteilung. ε bezeichnet dabei die negative normierte Energie pro Spin

$$\varepsilon := -\frac{E}{J/\sqrt{N}}/N, \quad (2.16)$$

¹Man beachte, daß dann für E gemäß der Bemerkung am Ende von Abschnitt 2.2 der Einzelspinflip–Hamiltonian (Gl.(2.8)) genommen werden muß.

wobei $\frac{J}{\sqrt{N}}$ die Standardabweichung für J_{ij} in der Wahrscheinlichkeitsverteilung (Gl.(1.4))

$$P(J_{ij}) = \sqrt{\frac{N}{2\pi J^2}} \exp\left(-N \frac{J_{ij}^2}{2J^2}\right)$$

ist.

Das Maximum in der Verteilung für ε ergibt sich zu $\varepsilon = 0.506$.

2.3.2 Frustration

Die exponentielle Form von $\langle\langle \mathcal{N} \rangle\rangle$ ist auf die *Frustration* zurückzuführen: Unter einem frustrierten Spin versteht man einen Spin, der nicht mit allen seinen Nachbarn im Einklang steht, das heißt S_i ist nicht zu allen Partialfeldern $J_{ij}S_j$ der Nachbarspins S_j parallel. Ein ganzes Spinsystem wird als frustriert bezeichnet, wenn die Kopplungen J_{ij} so vorgegeben sind, daß es in *jedem* Zustand \underline{S} des Systems einen frustrierten Spin gibt. Es darf also keinen Zustand \underline{S} des Systems geben, der zu allen Kopplungen J_{ij} energetisch günstig steht [Ki Vorl]. Spingläser weisen in der Regel Frustration auf, weil hier alle Kopplungen (unabhängig voneinander) ausgewürfelt werden.

Man kann sich nun anschaulich vorstellen, daß ein frustriertes System mehr Grundzustände als ein nichtfrustriertes System besitzt:

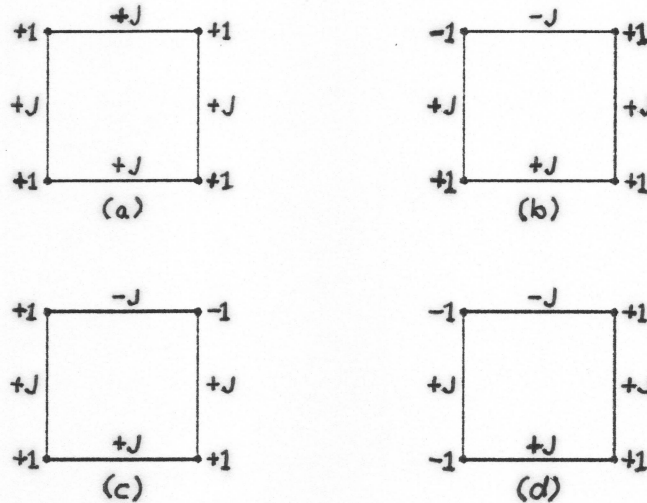


Abbildung 2.1: Veranschaulichung der Entartung des Grundzustandes:
a) Bei einem nichtfrustrierten Viereck-Spin-System: $1 \cdot 2$ Grundzustände mit $\mathcal{H} = -4J$
(b)-(d): Bei einem frustrierten Viereck-Spin-System: $3 \cdot 2$ Grundzustände mit $\mathcal{H} = -2J$

Weitergehende anschauliche Überlegungen dieser Art hat G. Toulouse angestellt [To77]. Er hat weiterhin zusammen mit J. Vannimenus die frustrationsbedingte Entartung W des Grundzustandes² in einem $\pm J$ -Spinglas mit Hilfe von Computersimulationen quantitativ untersucht [VaTo77]:

Steigt die Konzentration x der negativen Bindungen und damit der Grad der Frustration von $x = 0$ bis $x = \frac{1}{2}$ an, so steigt auch W an. Für W konnten Vannimenus und Toulouse eine exponentielle Form nachweisen, und sie erhielten im Maximum $x = \frac{1}{2}$:

$$W(x = \frac{1}{2}) = e^{N \cdot 0.07}. \quad (2.17)$$

Man erwartet nun, daß nicht nur die Zahl der Grundzustände, sondern auch die Zahl der metastabilen Zustände \mathcal{N} exponentielle Form hat. Der Ansatz

²für den Einzelspinflip-Hamiltonian

für das (nicht gemittelte) \mathcal{N} des SK–Spinglases ist damit

$$\mathcal{N} = e^{Nh}, \quad (2.18)$$

wobei h intensiv ist, also nicht mit der Systemgröße N wächst.

2.3.3 Selbstmittelungseigenschaft von $h = \frac{1}{N} \ln \mathcal{N}$

Die Form

$$h(\{J_{ij}\}) = \frac{1}{N} \ln \mathcal{N}$$

ähnelt der Beziehung

$$f(\{J_{ij}\}) = -\frac{\tau}{N} \ln Z$$

zwischen der Freien Energie pro Spin f und der kanonischen Zustandssumme Z des SK–Spinglases. Nf und Nh sind beide extensiv, wohingegen \mathcal{N} und Z beide exponentiell wachsen. Um eine analytische Aussage für große Systeme machen zu können, führt man jetzt die *Selbstmittelungseigenschaft* von f bzw. h ins Feld (siehe [BiYo86]):

Wählt man ein großes N und würfelt die J_{ij} einmal aus, so liefert h den gleichen Wert, als wenn man für dasselbe System über alle möglichen Kopplungen J_{ij} gemittelt hätte, also den „quenched average“ (siehe Abschnitt 1.1) gebildet hätte:

$$\lim_{N \rightarrow \infty} h(\{J_{ij}\}) = \langle\langle h(\{J_{ij}\}) \rangle\rangle \quad (2.19)$$

h wird also für große N immer schärfer und nimmt seinen Mittelwert an. „Das ist nichts anderes als das Gesetz der großen Zahlen“ [Am89]. Für Systeme mit kurzreichweitigen Wechselwirkungen kann man sich dies anschaulich klar machen. Die Selbstmittelungseigenschaft läßt sich für gewisse solche Systeme beweisen [EnHe84].

Um eine Aussage über die Zahl der metastabilen Zustände $\mathcal{N} = e^{Nh}$ für große N machen zu können, muß man also $h = \frac{1}{N} \ln \mathcal{N}$ mitteln. Die Mittelung des \ln geschieht mit Hilfe der Replica–Methode. Roberts hat dies fürs SK–Modell durchgeführt [Ro81]. Die direkte Mittelung $\langle\langle \mathcal{N} \rangle\rangle$ umgeht die aufwendige Replica–Rechnung, liefert jedoch nur eine *obere Schranke* für $e^{N\langle\langle h \rangle\rangle}$. Man macht sich dies folgendermaßen klar [BiYo86]:

Man nimmt an, daß die selbstmittelnde Größe h gaußverteilt ist; h habe die Wahrscheinlichkeitsdichte

$$P(h) \propto \exp\left(-N \frac{(h - h_0)^2}{2\sigma^2}\right), \quad (2.20)$$

wobei $h_0 = \langle\langle h \rangle\rangle = \lim_{N \rightarrow \infty} h$ und $\frac{\sigma}{\sqrt{N}}$ die für $N \rightarrow \infty$ verschwindende Streuung von h ist, $P(h)$ wird dann δ -förmig. Für den Mittelwert von \mathcal{N} folgt dann mit Hilfe der Gauß-Integral-Formel (Gl.(A.1)):

$$\langle\langle \mathcal{N} \rangle\rangle = \int_{-\infty}^{+\infty} dh P(h) e^{Nh} \propto \int_{-\infty}^{+\infty} dh_1 \exp\left(-N \frac{h_1^2}{2\sigma^2} + N(h_1 + h_0)\right) \quad (2.21)$$

$$\langle\langle \mathcal{N} \rangle\rangle = \langle\langle e^{Nh} \rangle\rangle = e^{N(h_0 + \frac{1}{2}\sigma^2)} \geq e^{Nh_0} = e^{N\langle\langle h \rangle\rangle} \quad . \quad (2.22)$$

Das Ergebnis stellt einen Spezialfall der Jensenschen Ungleichung dar (siehe z.B. [Hi75]). Feynman veranschaulichte die Ungleichung mit der Konvexität der e -Funktion; die e -Funktion gewichtet höhere Werte von h stärker, so daß $\frac{1}{N} \ln \langle\langle e^{Nh} \rangle\rangle$ größer wird als $\langle\langle h \rangle\rangle$ [Fe72]. Die Ungleichung beruht im Endeffekt darauf, daß die Streuung von h als mit $\frac{1}{\sqrt{N}}$ skalierend angenommen wird. In diesem Fall kann man das Integral in Gl.(2.21) auch als Sattelpunktintegral verstehen [Ki Vorl]. $\langle\langle \mathcal{N} \rangle\rangle$ wird dann nicht vom Mittelwert h_0 beherrscht, sondern vom Sattelpunktwert des Ausdruckes

$$k = -\frac{1}{2} \cdot \frac{h_1^2}{\sigma^2} + h_1 + h_0 \ ,$$

der sich mit $h_{1S} = \sigma^2$ zu $k = h_0 + \frac{1}{2}\sigma^2$ ergibt.

Die Jensensche Ungleichung besagt also auch, daß der Sattelpunktwert höher als der Mittelwert h_0 ist. Fällt die Streuung von h stärker mit N ab, so gilt

$$\langle\langle \mathcal{N} \rangle\rangle \propto \exp\left(N\left(h_0 + \mathcal{O}(N^{-\alpha})\right)\right)$$

mit $\alpha > 0$. Für große N stimmen dann die beiden Seiten in der Ungleichung

$$\frac{1}{N} \ln \langle\langle \mathcal{N} \rangle\rangle \geq \langle\langle h \rangle\rangle \quad (2.23)$$

mit beliebiger Genauigkeit überein. Man sagt dann *ebenfalls* \mathcal{N} sei „selbstmittelnd“. Dies ist *nur* in Zusammenhang damit zu verstehen, daß in Ungleichung (2.23) für $N \rightarrow \infty$ das Gleichheitszeichen steht, genauso, wie in

der „richtigen“ Selbstmittelungsgleichung (2.19) für $N \rightarrow \infty$ das Gleichheitszeichen steht. In Ungleichung (2.23) steht genau dann das Gleichheitszeichen, wenn alle Nebendiagonalparameter in der Replica-Rechnung Null sind. Das ist nämlich ein Zeichen dafür, daß zwischen den einzelnen Replica keine Korrelationen bestehen und man sich die Replica-Rechnung hätte sparen können. Diese Bemerkung wird später in Abschnitt 3.3 an einem Beispiel veranschaulicht.

2.3.4 Ergebnisse von Bray und Moore und von Roberts für das SK-Modell

Bray und Moore ([BrMo80],[BrMo81]) haben die Dichte $\mathcal{N}_S(f)$ der Lösungen der TAP-Gleichungen [TAP77] bei vorgegebener TAP-Freier Energie pro Spin f für das SK-Spinalgas berechnet. Sie berechneten zunächst den direkten Mittelwert $\langle\langle \mathcal{N}_S(f) \rangle\rangle$ in Abhängigkeit von der Temperatur und erhielten für $\tau \rightarrow 0$ wieder das Ergebnis von Tanaka und Edwards (siehe Abschnitt 2.3.1). Des weiteren führten sie eine Replica-Rechnung unter Annahme der Replica-Symmetrie durch und ermittelten die selbstmittelnde Größe $\frac{1}{N} \langle\langle \ln \mathcal{N}_S(f) \rangle\rangle$, lösten jedoch nicht die zugehörigen schwierigen Sattelpunktgleichungen. Für $\tau \rightarrow 0$ berechneten sie wie Tanaka und Edwards den direkten Mittelwert $\langle\langle \mathcal{N}(\varepsilon) \rangle\rangle$ der Dichte der metastabilen Zustände bei vorgegebenem ε . Da $\langle\langle \mathcal{N}(\varepsilon) \rangle\rangle$ eine obere Schranke für

$$\langle\langle h \rangle\rangle = \frac{1}{N} \langle\langle \ln \mathcal{N}(\varepsilon) \rangle\rangle$$

liefert, fanden sie den Energie-Bereich, in dem metastabile Zustände erwartet werden:

$$0.286 \leq \varepsilon \leq 0.791, \quad (2.24)$$

wobei die Konstanten die Nullstellen von $f^* := \frac{1}{N} \ln \langle\langle \mathcal{N}(\varepsilon) \rangle\rangle$ sind.³ Innerhalb des ε -Bereichs ist nämlich $f^*(\varepsilon) > 0$ und außerhalb gilt $f^*(\varepsilon) < 0$, so daß die obere Schranke für $\langle\langle h \rangle\rangle$ dort beliebig klein wird. Wenn wir also die Selbstmittelungseigenschaft (Gl.(2.19)) für h voraussetzen, also annehmen, daß $\langle\langle h \rangle\rangle$ ein typischer Wert für h ist, können wir die Existenz von metastabilen Zuständen und damit auch von Grundzuständen im äußeren Bereich

³Der Stern dient zur Unterscheidung von der Freien Energie pro Spin.

ausschließen. $\varepsilon = 0.791$ ist damit eine obere Schranke für die (negative normierte) Grundzustandsenergie ⁴ des Spinglases.

Roberts [Ro81] führte die Replica-symmetrische Rechnung für $\frac{1}{N}\langle\langle\ln\mathcal{N}(\varepsilon)\rangle\rangle$ durch und erhielt bis $\varepsilon_0 = 0.672$ die gleiche Kurve wie Bray und Moore, die diesen Wert auch schon voraussagten; $\mathcal{N}(\varepsilon)$ ist also „selbstmittelnd“, sofern $\varepsilon \leq 0.672$. Ab ε_0 werden die Nebendiagonalparameter in der Replica-Rechnung von Null verschieden, und es ist

$$\frac{1}{N}\langle\langle\ln\mathcal{N}(\varepsilon)\rangle\rangle < \frac{1}{N}\ln(\langle\langle\mathcal{N}(\varepsilon)\rangle\rangle),$$

die Kurve knickt ab. Für die (negative normierte) Grundzustandsenergie erhielt Roberts somit eine bessere obere Schranke:

$$\varepsilon_G = 0.77,$$

die in „guter Übereinstimmung“ mit vorhergehenden Simulationen und dem Parisi-Ergebnis [Pa80] war.

Unter der Annahme der Selbstmittelungseigenschaft von h kann man also die Replica-symmetrische Theorie von Sherrington und Kirkpatrick ([SK75], [SK78]) widerlegen, denn dort war $\varepsilon_G = 0.798 > 0.77$. Eine ähnliche Argumentation führe ich mit Hilfe meiner Ergebnisse bezüglich der Rechnung von Kanter und Sompolinsky [KaSo87] (siehe Abschnitt 5.3).

Allgemeiner müßte also die obere Nullstelle von $\frac{1}{N}\langle\langle\ln\mathcal{N}_S(f)\rangle\rangle$ aus der Rechnung von Bray und Moore die Freie Energie pro Spin im thermischen Gleichgewicht bei Temperatur τ liefern. Bray und Moore argumentierten, daß die Replica-symmetrische Theorie nicht für alle τ gute Werte liefern würde, so daß man die Brechung der Replica-Symmetrie beachten muß und die Rechnung sehr schwierig wird.

Abschließend sei noch auf das Problem hingewiesen, daß die metastabilen Zustände in der thermodynamischen Rechnung von Sherrington und Kirkpatrick nicht auftreten; man findet dort nur die absoluten Minima der Freien Energie bzw. die Energie-Grundzustände und die relativen Minima von F bzw. von E . Zur Aufklärung könnte man das Argument vorbringen, daß diese Minima stabil gegen eine makroskopische Zahl von Spin-Flips sind, weil sie

⁴Ich erinnere nocheinmal daran, daß der Begriff Energie hier untrennbar mit der Einzelspinflip-Dynamik verknüpft ist.

Lösungen von Ordnungsparameter-Gleichungen darstellen, wohingegen metastabile Zustände nur gegen Einzelspinflips stabil sind.

Die Existenz exponentiell vieler TAP-Lösungen wirft dann wieder ein *neues* Problem auf: Die TAP-Lösungen erfüllen $\nabla f(\vec{m}) = 0$, viele von ihnen sind wohl lokale Minima der TAP-freien Energie pro Spin f . Da die Zahl der TAP-Lösungen für $\tau \rightarrow 0$ stetig in die Zahl der metastabilen Zustände übergeht, kann man vermuten, daß exponentiell viele Zustände existieren, die auch gegen eine makroskopische Zahl von Spin-Flips stabil sind. Es bleibt also die Frage, welche „besonderen lokalstabilen“ Zustände die thermodynamische (Replica)-Rechnung herauszieht. Diese Frage ist bisher ungeklärt.

2.3.5 Zweierzyklen beim SK-Spینگlas

In der Arbeit von Gutfreund, Reger und Young [Gu+88] wird erwähnt, daß für das SK-Spینگlas die Zahl der Zweierzyklen ⁵ \mathcal{N}_Z im Vergleich zu der Zahl der metastabilen Zustände \mathcal{N}_M überwiegt. Es gilt überraschenderweise $\langle\langle\mathcal{N}_Z\rangle\rangle \sim (\langle\langle\mathcal{N}_M\rangle\rangle)^2$, das heißt

$$\frac{1}{N} \ln \langle\langle\mathcal{N}_Z\rangle\rangle = 2 \cdot 0.1992 \quad (2.25)$$

Ein paralleler $\tau = 0$ Glauber-Prozeß wird also fast immer in einem Zweier-Zyklus landen.

Die Zahl der metastabilen Zustände ist folglich für den Parallel-Prozeß nicht wichtig. Da zudem in den vorausgegangenen Überlegungen stets der Einzelspinflip-Hamiltonian verwendet wurde, hat man bisher nur Aufschlüsse über die Einzelspinflip-Dynamik gewonnen. Dies wird auch für den folgenden Abschnitt 2.4 sowie für die Hauptkapitel der vorliegenden Arbeit, Kapitel 4 und 5, der Fall sein.

⁵Ein Berechnungsverfahren für die Zahl der Zweierzyklen ist in Anhang D angegeben.

2.4 Metastabile Zustände beim Hopfield–Modell

Bei Neuronalen Netzen mit $\tau = 0$ –Glauber–Dynamik sind die metastabilen Zustände unerwünschte Fallen für die sequentielle Dynamik: Das System soll ja nicht mitten auf dem Weg stehenbleiben, sondern nach Möglichkeit in ein Muster laufen. Es ist deshalb von besonderer Wichtigkeit zu wissen, wo im Phasenraum (in welcher „Entfernung“ von den Mustern) die Fallen liegen. Elizabeth Gardner hat analog zu Tanaka und Edwards die Rechnung für das Hopfield–Modell durchgeführt [Ga86]. In dem Neuronalen Netzwerk mit N Spins S_i sind $p = \alpha N$ Muster $\underline{\xi}^\nu$, $\nu \in \{1, \dots, p\}$, durch die Kopplungen

$$J_{ij} = \frac{1}{N} \sum_{\nu=1}^p \xi_i^\nu \xi_j^\nu \quad (2.26)$$

gespeichert (Hebbsche Lernregel, siehe Abschnitt 1.1). Da J_{ij} symmetrisch ist, stellt

$$\mathcal{H} = -\frac{1}{2} \sum_{i,j(\neq i)} J_{ij} S_i S_j \quad (2.27)$$

nach Abschnitt 2.1.1 wieder einen Hamiltonian für die $\tau = 0$ –Glauber–Einzelspinflip–Dynamik ohne Selbstkopplungen dar.

Die Mittelung erfolgt wieder über Zufallsmuster, die zuvor mit den Wahrscheinlichkeiten $p(\xi_i^\nu = \pm 1) = \frac{1}{2}$ ausgewürfelt worden sind. An dieser Stelle wird betont, daß die $\binom{N}{2}$ vielen Kopplungen J_{ij} über die p vielen Muster stark miteinander korreliert sind (siehe auch Abschnitt 1.1). Es ist deshalb unumgänglich, über die Muster zu mitteln. Eine Mittelung über die fälschlicherweise als unabhängig angenommenen Kopplungen würde wieder zum Ergebnis von Tanaka und Edwards führen (siehe Abschnitt 2.3.1).

Zusätzlich zu $\varepsilon = -\frac{\mathcal{H}}{N}$ kann man noch den *Hamming–Abstand* g der betrachteten metastabilen Zustände zu irgendeinem Muster in die Rechnung einbauen. Da wegen der Mittelung kein Muster ausgezeichnet ist, kann man g bezüglich des ersten Musters betrachten. Unter dem Hamming–Abstand versteht man den Bruchteil der Spins, die im Zustand \underline{S} bezüglich $\underline{\xi}^1$ falsch stehen. Für den Überlapp m_1 zwischen $\underline{\xi}^1$ und \underline{S} besteht dann die Beziehung

$$m_1 = \frac{1}{N} \sum_{i=1}^N S_i \xi_i^1$$

$$\begin{aligned}
&= \frac{1}{N} ((\text{Zahl der Plätze } i, \text{ wo } S_i = \xi_i^1) \\
&\quad - (\text{Zahl der Plätze } i, \text{ wo } S_i = -\xi_i^1)) \\
&= \frac{1}{N} (N(1-g) - Ng) = 1 - 2g \iff g = \frac{1-m_1}{2}. \quad (2.28)
\end{aligned}$$

Die mit den Einschränkungen g und E versehene Dichte $\mathcal{N}(\alpha, g, E)$ von metastabilen Zuständen kann man auf zwei Arten berechnen:

1.

$$\mathcal{N}(\alpha, g, E) = \sum'_{\{S_i\}} \left(\prod_{i=1}^N \Theta(S_i h_i) \right) \delta(E - \mathcal{H}(\underline{S})),$$

wobei der Strich bedeutet, daß die Summe auf die Zustände mit festem Hamming-Abstand g zum ersten Muster beschränkt bleibt. Diese Methode wird in [Ga86] und auch in der vorliegenden Arbeit verwendet (siehe Abschnitt 4.1).

2. Man baut eine zweite δ -Funktion für m_1 ein und summiert über alle Zustände \underline{S} :

$$\mathcal{N}(\alpha, m_1, E) = \sum_{\{S_i\}} \left(\prod_{i=1}^N \Theta(S_i h_i) \right) \delta(E - \mathcal{H}(\underline{S})) \delta\left(m_1 - \frac{1}{N} \sum_{i=1}^N \xi_i^1 S_i\right).$$

An dieser Stelle soll betont werden, daß der Autor dieser Arbeit der Meinung ist, daß das in [Ga86] angegebene Ergebnis für $\langle\langle \mathcal{N}(\alpha, g = \frac{1}{2}, E) \rangle\rangle$ nicht stimmt. Diesbezügliche Ergebnisse sind also eigene Ergebnisse. Die Rechnung findet sich in Anhang B.

Das asymptotische Ergebnis für $N \rightarrow \infty$ ist

$$\langle\langle \mathcal{N}(\alpha, g, E) \rangle\rangle \propto e^{N \cdot f(\alpha, g, E)}, \quad (2.29)$$

wobei f der Sattelpunktwert bezüglich b der folgenden Größe ist:

$$\begin{aligned}
f(b, \alpha, g, E) &= -g \ln g - (1-g) \ln(1-g) + g \ln \Phi(t_1) + (1-g) \ln \Phi(t_2) + \\
&\quad + \alpha \left[\frac{(1-b)^2}{2a} - \frac{1}{2} + b + \frac{1}{2} \ln a \right] \quad ; \quad (2.30)
\end{aligned}$$

dabei ist die Φ -Funktion gegeben durch

$$\Phi(x) = \int_{-\infty}^x \frac{d\lambda}{\sqrt{2\pi}} e^{-\frac{1}{2}\lambda^2}.$$

Ihre Argumente sind

$$t_1 = \frac{1}{\sqrt{\alpha a}}(-b\alpha - (1 - 2g)); \quad t_2 = \frac{1}{\sqrt{\alpha a}}(-b\alpha + 1 - 2g); \quad (2.31)$$

a ist bei vorgegebenem E eine Konstante und ergibt sich aus

$$\frac{E}{N} = -\varepsilon = -\frac{1}{2}(1 - 2g)^2 + \frac{\alpha}{2}(1 - a) = -\frac{1}{2}(m_1)^2 + \frac{\alpha}{2}(1 - a). \quad (2.32)$$

Wegen

$$\begin{aligned} \frac{E}{N} &= -\frac{1}{2} \sum_{\mu=1}^p \left(\frac{1}{N} \sum_{i=1}^N \xi_i^\mu S_i \right)^2 + \frac{\alpha}{2} \\ &= -\frac{1}{2} \sum_{\mu=1}^p (m_\mu)^2 + \frac{\alpha}{2} \end{aligned} \quad (2.33)$$

können wir

$$a = \frac{1}{\alpha} \sum_{\mu=2}^p (m_\mu)^2$$

identifizieren. Gl.(2.32) weist somit verblüffende Ähnlichkeit zu Amits Formel für die Grundzustandsenergie beim Hopfield-Modell auf [Am+87].

Es reicht im folgenden aus, $g \in (0, \frac{1}{2}]$ zu betrachten. Dies ist zum einen anschaulich klar, da mit \underline{S} auch $-\underline{S}$ metastabil ist, zum anderen sieht man auch

$$f\left(\alpha, \frac{1}{2} + x, E\right) = f\left(\alpha, \frac{1}{2} - x, E\right)$$

nach Lösung der Sattelpunktgleichung für f . Das Maximum von f bezüglich ε errechnet sich wegen der linearen Abhängigkeit von a und ε aus

$$\frac{\partial f}{\partial a} = 0.$$

Die für diese Arbeit wichtigsten Ergebnisse *an* diesem Maximum finden sich in [Ga86] und lauten:

- Für kleine α -Werte gibt es in der Nähe der Muster, also bei kleinen Hamming-Abständen g , ein schmales Band von metastabilen Zuständen, die durch eine Lücke von den unkorrelierten Zuständen um $g = \frac{1}{2}$ getrennt sind. Unter der Lücke wird ein g -Bereich mit $f < 0$ verstanden. Analog Abschnitt 2.3.4 für den ε -Bereich schließt man hier für den g -Bereich, daß dort keine Zustände liegen können. Bei $\alpha_0 = 0.113$ verschwindet die Lücke. Das schmale Band ist für die Wiedererkennung des Musters verantwortlich. In der Tat enthält es die zuvor von Amit *et al.* [Am+87] bestimmten lokalen Minima der Energie, die die wegen des Ansatzes $p = \alpha N$ instabilen Muster ersetzen (siehe Abschnitt 1.1). Nicht verstanden ist bisher, ob und wie das Verschwinden der Lücke bei 0.113 mit der kritischen Speicherkapazität $\alpha_c \sim 0.14$ zusammenhängt: Bei $\alpha_c \sim \alpha_0$ verschwinden nämlich die oben erwähnten lokalen Minima in der Nähe des Musters; *nur* noch Spinglaszustände mit $g = \frac{1}{2}$ sind dann thermodynamisch stabil.
- Für $\alpha \rightarrow \infty$ verschwinden die Korrelationen zwischen den J_{ij} , und das Hopfield-Modell wird zum SK-Modell. $f(\alpha, g = \frac{1}{2})$ liefert dann erwartungsgemäß den Spinglaswert 0.1992 aus Abschnitt 2.3.1. In *diesem* Fall hätte man also auch wie eingangs erwähnt über die Wahrscheinlichkeitsverteilung der J_{ij} mitteln können.

Das Ergebnis für $\alpha \rightarrow 0$ ist

$$f(\alpha, g = \frac{1}{2}) \rightarrow -\frac{1}{2}\alpha \ln \alpha + \frac{1}{2}\alpha \left(\ln \left(\frac{2}{\pi} \right) - 1 \right) + \mathcal{O}(\alpha^2). \quad (2.34)$$

Neben dem oben erwähnten Maximum ergibt sich für $f(\alpha, g = \frac{1}{2}, \varepsilon)$ folgendes:

Für jeden α -Wert kann man $f(\varepsilon)$ auftragen. Es ergeben sich ähnliche Kurven wie bei Bray und Moore (siehe Abschnitt 2.3.4). Die obere Nullstelle von f ist wieder eine obere Schranke für die (negative normierte) Spinglas-Grundzustandsenergie ε_{SG} , die die Form

$$\varepsilon_{SG} = +\frac{1}{\pi} + \sqrt{\frac{2\alpha}{\pi}} \quad (2.35)$$

hat [Am+87].

Für $\alpha \rightarrow \infty$ liefert $f(\varepsilon)$ die Kurve von Bray und Moore fürs SK–Spinglas; dabei ist zu beachten, daß auf der Abszisse ε durch die Streuung $J = \sqrt{\alpha}$ der Kopplungen dividiert werden muß, damit man die beiden Kurven vergleichen kann. So erhält man z.B. für das Maximum der $f(\varepsilon)$ –Kurve im Falle $\alpha \rightarrow \infty$

$$\varepsilon_{max} \rightarrow 0.506 \sqrt{\alpha} , \quad (2.36)$$

was genau dem Ergebnis von Bray und Moore entspricht.

Für $\alpha \rightarrow 0$ rücken die beiden Nullstellen und ε_{max} der $f(\varepsilon)$ –Kurve zusammen, und es gilt

$$\varepsilon_{max} \rightarrow \frac{1}{\pi} + \frac{\alpha}{2} \left(3 - \frac{4}{\pi} \right) + \mathcal{O}(\alpha^2) . \quad (2.37)$$

Die Rechnung für diesen Fall findet sich im Anhang B. Es gilt also stets (nach Gl.(2.35))

$$\varepsilon_{max} < \varepsilon_{SG} ,$$

und im Grenzfall stimmen beide überein.

Kapitel 3

Ein Neuronales Netzwerk mit Projektorkopplungen

In diesem Kapitel wird ein Neuronales Netzwerk mit p linear unabhängigen N -dimensionalen Mustervektoren

$$\underline{\xi}^\nu = (\xi_1^\nu, \dots, \xi_N^\nu)^T$$

betrachtet, wobei

$$\nu \in \{1, \dots, p\}, \quad \xi_i^\nu \in \{-1, +1\} \quad \text{und} \quad p < N.$$

Später wird $p = \alpha N$ gesetzt, $\alpha \in (0, 1)$ konstant.

Personnaz, Guyon und Dreyfus ([Pe+85],[Pe+86]) führten für die Kopplungen (J_{ij}) eine Projektormatrix mit folgenden Eigenschaften ein:

- Für die inneren Felder der Muster gilt einfach

$$h_i^\mu = \sum_{j=1}^N J_{ij} \xi_j^\mu = \xi_i^\mu \quad \forall i, \mu \quad (3.1)$$

- Die Zustände $\underline{\delta S}$ aus dem Orthogonalraum der Muster haben innere Felder Null

$$h_i = \sum_{j=1}^N J_{ij} \delta S_j = 0 \quad \forall i, \text{ falls } \underline{\delta S} \cdot \underline{\xi}^\mu = 0 \quad \forall \mu \quad (3.2)$$

In der $\tau = 0$ -Glauber-Dynamik sind die Muster dann nicht nur metastabil, sondern es gilt sogar

$$\lambda_i^\mu = \xi_i^\mu h_i^\mu = 1 > 0 \quad \forall i, \mu . \quad (3.3)$$

Bevor ich das Modell näher erläutere, stelle ich zunächst den Zusammenhang zwischen der obigen Projektormatrix und der „pseudoinversen“ Matrix dar.

3.1 Die Projektormatrix

Die Mustervektoren seien linear unabhängig. Man bildet die *Mustermatrix*

$$\mathbf{A} := \frac{1}{\sqrt{N}} (\underline{\xi}^1, \underline{\xi}^2, \dots, \underline{\xi}^p) . \quad (3.4)$$

\mathbf{A} ist eine $N \times p$ -Matrix. Die $\underline{\xi}^\nu$ sind Spaltenvektoren.

Die symmetrische $p \times p$ -Matrix

$$\mathbf{C} := \mathbf{A}^T \mathbf{A} \quad \text{mit} \quad C_{\mu\nu} = \frac{1}{N} \sum_{i=1}^N \xi_i^\mu \xi_i^\nu \quad (3.5)$$

enthält die Überlapps der Muster untereinander und heißt demzufolge *Korrelationsmatrix*. \mathbf{C} ist invertierbar und sogar positiv definit [ZuFa84]:

$$\text{Mit } \vec{y} = \mathbf{A} \vec{x}$$

folgt nämlich

$$Q(\vec{x}) = \vec{x}^T \mathbf{C} \vec{x} = \vec{y}^T \vec{y} \geq 0 .$$

Da die $\underline{\xi}^\nu$ linear unabhängig sind, ist

$$\vec{y} = 0 \iff \vec{x} = 0 \implies Q(\vec{x}) > 0 \quad \forall \vec{x} \neq 0 .$$

Die Linksinverse von \mathbf{A} ist nach [ZuFa84] nicht eindeutig bestimmt, falls $p < N$. Man definiert eine besondere Linksinverse, die „Pseudoinverse“ von \mathbf{A} :

$$\hat{\mathbf{A}}^T := \mathbf{C}^{-1} \mathbf{A}^T \implies \hat{\mathbf{A}}^T \mathbf{A} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{A} = \mathbf{E}_p$$

Mit der Transponierten $\hat{\mathbf{A}}$ der Pseudoinversen definiert man schließlich den gesuchten *Projektor*:

$$\mathbf{J} := \hat{\mathbf{A}}\mathbf{A}^{\mathbf{T}} = \mathbf{A}(\mathbf{C}^{-1})^{\mathbf{T}}\mathbf{A}^{\mathbf{T}} = \mathbf{A}\mathbf{C}^{-1}\mathbf{A}^{\mathbf{T}} \quad (3.6)$$

ist eine symmetrische $N \times N$ -Matrix.

Nachweis der Projekteigenschaften:

$$\begin{aligned} \mathbf{J}^2 &= \hat{\mathbf{A}}\mathbf{A}^{\mathbf{T}} \cdot \hat{\mathbf{A}}\mathbf{A}^{\mathbf{T}} = \mathbf{A}\mathbf{C}^{-1}\mathbf{A}^{\mathbf{T}}\mathbf{A}\mathbf{C}^{-1}\mathbf{A}^{\mathbf{T}} \\ &= \mathbf{A}\mathbf{C}^{-1}\mathbf{A}^{\mathbf{T}} = \mathbf{J} \end{aligned} \quad (3.7)$$

$$\mathbf{J}\mathbf{A} = \mathbf{A}\mathbf{C}^{-1}\mathbf{A}^{\mathbf{T}}\mathbf{A} = \mathbf{A} \quad . \quad (3.8)$$

Man liest also in den Spalten dieser Matrixgleichung ab:

$$\mathbf{J}\underline{\xi}^{\nu} = \underline{\xi}^{\nu} ;$$

alle Muster sind wie gefordert Eigenvektoren zum Eigenwert 1. Für die Zustände $\underline{\delta S}$ aus dem Orthogonalraum der Muster gilt

$$\mathbf{J}\underline{\delta S} = \mathbf{A}\mathbf{C}^{-1}\mathbf{A}^{\mathbf{T}}\underline{\delta S} = \mathbf{A}\mathbf{C}^{-1}\vec{0} = \vec{0}.$$

Da der Projektor auf einer Basis des \mathbf{R}^N erklärt ist, ist er somit durch \mathbf{J} *eindeutig bestimmt*.

Äquivalent zur obigen Behandlung kommt man auf den Projektor, wenn man fordert

$$\mathbf{J}\underline{\xi}^{\nu} = \underline{\xi}^{\nu} \quad \forall \nu \quad \text{mit minimaler Norm} \quad \frac{1}{N} \sum_{i,j} J_{ij}^2 .$$

Die Methode der Lagrange-Parameter [Fl87] kommt hier zur Anwendung:

$$F(\{J_{ij}\}) = \frac{1}{N} \sum_{i,j} J_{ij}^2 + \sum_{i,\mu} \lambda_i^{\mu} \left(\sum_{j=1}^N J_{ij} \xi_j^{\mu} - \xi_i^{\mu} \right)$$

ist zu minimieren. Wir bilden den Gradienten bezüglich der $\{J_{ij}\}$ und erhalten N^2 Gleichungen der Form

$$\frac{2}{N} J_{ij} + \sum_{\nu=1}^p \lambda_i^{\nu} \xi_j^{\nu} = 0, \quad (3.9)$$

sowie $N \cdot p$ Nebenbedingungen

$$\sum_{j=1}^N J_{ij} \xi_j^\mu = \xi_i^\mu . \quad (3.10)$$

Setzt man Gl.(3.9) in Gl.(3.10) ein, so folgen $N \cdot p$ Gleichungen für die Lagrange-Parameter. Dies ergibt N entkoppelte lineare Gleichungssysteme

$$\mathbf{C} \vec{\lambda}_i = -\frac{2}{N^2} \vec{\xi}_i \quad \forall i \iff \vec{\lambda}_i = -\frac{2}{N^2} \mathbf{C}^{-1} \vec{\xi}_i \quad \forall i .$$

Nach Einsetzen in Gl.(3.9) erhält man dann

$$J_{ij} = -\frac{N}{2} \vec{\lambda}_i^T \cdot \vec{\xi}_j = \frac{1}{N} \vec{\xi}_i^T \mathbf{C}^{-1} \vec{\xi}_j .$$

Das sind aber gerade die Matrixelemente von \mathbf{J} aus Gl.(3.6). Die Lösung ist eindeutig und sie stellt ein Minimum von $F\{J_{ij}\}$ dar, weil die Matrix der zweiten Ableitungen von F proportional zur Einheitsmatrix und positiv definit ist (siehe Theorem 9.3.15 in [Fl87] (Satz über die hinreichende Bedingung für Extrema unter Zwangsbedingungen)).

Man gewinnt nun die beiden Formeln

$$\frac{1}{N} \sum_{i,j} J_{ij}^2 = \alpha \quad (3.11)$$

$$\frac{1}{N} \sum_{i=1}^N J_{ii} = \alpha , \quad (3.12)$$

denn

$$\begin{aligned} \frac{1}{N} \sum_{i,j} J_{ij}^2 &= \frac{1}{N} \text{Tr}(\mathbf{J} \cdot \mathbf{J}) = \frac{1}{N} \text{Tr}(\mathbf{J}) \\ &= \frac{1}{N^2} \sum_{i=1}^N \sum_{\mu,\nu} (C^{-1})_{\mu\nu} \xi_i^\mu \xi_i^\nu = \frac{1}{N} \text{Tr}(\mathbf{C}^{-1} \mathbf{C}) = \alpha \end{aligned}$$

gilt wegen der Projektoreigenschaft von \mathbf{J} und den Symmetrieeigenschaften von \mathbf{J} und \mathbf{C} .

Zum Abschluß sei noch erwähnt, daß die Pseudoinverse im Falle $p > N$ das Gaußsche Problem der kleinsten Quadrate löst [ZuFa84]; weitergehende Betrachtungen zur Pseudoinversen findet man im Standardwerk [Al72] und in [Ko88].

3.2 Eigenschaften des Modells

Das betrachtete Projektornetzwerk (kurz: PN) hat also Kopplungen

$$J_{ij} = \frac{1}{N} \sum_{\mu, \nu} \xi_i^\mu (C^{-1})_{\mu\nu} \xi_j^\nu, \quad (3.13)$$

wobei

$$C_{\mu\nu} = \frac{1}{N} \sum_{i=1}^N \xi_i^\mu \xi_i^\nu; \quad \xi_i^\mu \in \{-1, +1\}. \quad (3.14)$$

Für nicht korrelierte Muster ist $\mathbf{C} = \mathbf{E}_p$; J_{ij} stellt dann die Hebbische Lernregel dar:

$$J_{ij}^H = \frac{1}{N} \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu.$$

Bei der Hebbischen Regel handelt es sich um eine „lokale“ Lernregel, denn J_{ij}^H hängt nur von den Werten der ξ_k^μ an den Plätzen $k = i$ und $k = j$ ab. Das PN stellt mit den oben definierten Kopplungen hingegen ein nichtlokales Modell dar. Zur Berechnung der J_{ij} benötigt man Informationen von anderen Stellen des Netzwerks, hier sogar von allen anderen Stellen. In biologischen Systemen tritt dies nicht auf. Auch für spätere technische Anwendungen dürften nichtlokale Lernregeln von Nachteil sein.

Deshalb stellt die lokale „Adaline-Lernregel“ (siehe [DiOp87]) eine Alternative dar: Man legt dem Netzwerk der Reihe nach alle Muster vor und verändert die Kopplungen (Startwerte $J_{ij}^0 = 0$) gemäß

$$\delta J_{ij} = \frac{1}{N} \left(1 - \xi_i^\nu \sum_{k=1}^N J_{ik} \xi_k^\nu \right) \xi_i^\nu \xi_j^\nu. \quad (3.15)$$

Dann konvergiert das Verfahren gegen die obigen Kopplungen des PN (3.13). K.W. Berryman *et al.* bewiesen später, daß das Adaline-Verfahren auch konvergiert, wenn man linear abhängige Muster lernt. J_{ij} ist dann die Projektormatrix einer maximal linear unabhängigen Untermenge der präsentierten Muster [Be+90].

Betrachtet sei nun wieder das PN mit Kopplungen aus Gl.(3.13). Für die späteren Rechnungen ist wichtig, daß die Muster zufällig mit unabhängigen Wahrscheinlichkeiten $p(\xi_i^\mu = \pm 1) = \frac{1}{2} \quad \forall i, \mu$ verteilt sind. Der Überlapp

eines solchen Zufallsmusters μ_0 mit *allen* anderen Mustern kann dann als „random walk“ der Länge $(p-1) \cdot N$ aufgefaßt werden (siehe [KaSo87]):

$$\sum_{\nu(\neq\mu_0)} C_{\mu_0\nu} = \frac{1}{N} \sum_{\nu(\neq\mu_0)} \sum_{i=1}^N \xi_i^{\mu_0} \xi_i^\nu = \mathcal{O}\left(\sqrt{\frac{p}{N}}\right) . \quad (3.16)$$

Im Falle $p = \alpha N$ sind die Muster also korreliert. A.Odlyzko hat weiterhin nachgewiesen [KaSo87], daß die p Zufallsmuster (trotz der Korrelation) im Fall $N \rightarrow \infty$ und $\alpha < 1$ mit verschwindender Wahrscheinlichkeit linear abhängig sind. Für fast alle Mengen von p Zufallsmustern ist die Mustermatrix \mathbf{A} aus Gl.(3.4) also spaltenregulär und die Korrelationsmatrix \mathbf{C} positiv definit.

Wie im Vorspann beschrieben, sind die Muster zumindest metastabile Zustände bezüglich der $\tau = 0$ -Glauber-Dynamik. Sie sollten aber auch stabil gegen mehrere Spin-Flips sein. Dies ist nicht der Fall für das betrachtete PN *mit* Selbstkopplungen J_{ii} ; die folgende Überlegung aus [KaSo87] macht dies plausibel: Man betrachtet einen Zustand, der außer am 1. Platz mit allen Spins des μ -ten Musters identisch ist. Die Glauber-Dynamik bei $\tau = 0$ besagt dann

$$\begin{aligned} S_1(t+1) &= \text{sign} \left[\sum_{j(\neq 1)} J_{1j} \xi_j^\mu + J_{11} S_1(t) \right] \sim \text{sign} [(1-\alpha)\xi_1^\mu + \alpha S_1] \\ &\sim \text{sign} [(1-2\alpha)\xi_1^\mu] = -\xi_1^\mu = S_1(t) \quad \text{für } \alpha \geq \frac{1}{2} . \end{aligned} \quad (3.17)$$

Dies gilt wegen Formel 3.1 im Vorspann und wegen $\langle\langle J_{ii} \rangle\rangle = \alpha$ mit Streuung $\mathcal{O}\left(\frac{1}{\sqrt{N}}\right)$ (wird in Abschnitt 3.4 noch bewiesen). Für $\alpha \geq \frac{1}{2}$ sind die Muster also nicht einmal stabil gegen Ein-Spin-Flips. In numerischen Simulationen bestätigten Kanter und Sompolinsky [KaSo87], daß die kritische Speicherkapazität $\alpha_c \leq \frac{1}{2}$ ist. Die vorliegende Arbeit liefert einen Nachweis für

$$\alpha_c = \frac{1}{2}$$

(siehe Abschnitt 5.2.2).

Um α_c zu erhöhen, betrachteten Kanter und Sompolinsky das Modell ohne die Hauptdiagonalelemente J_{ii} :

$$\mathcal{H} = -\frac{1}{2} \sum_{i,j(\neq i)} J_{ij} S_i S_j - N \frac{\alpha}{2} = -\frac{1}{2} \sum_{i,j} J_{ij} S_i S_j . \quad (3.18)$$

In die Dynamik gehen die Selbstkopplungen nicht ein:

$$h_i = \sum_{j(\neq i)} J_{ij} S_j = \sum_{j=1}^N J_{ij} S_j - J_{ii} S_i .$$

Damit wurde sichergestellt, daß das Freie Energie–Funktional Φ aus Gl.(2.5) Lyapunov–Funktion der *seriellen* Dynamik ist ¹.

Die inneren Felder der Muster sind dann nach Gl.(3.1)

$$\xi_i^\mu h_i^\mu = 1 - J_{ii} \cdot 1 \sim 1 - \alpha > 0 \quad \text{für} \quad \alpha < 1 .$$

Kanter und Sompolinsky zeigten für dieses Modell

$$\alpha_c = 1 .$$

Skaliert man die Kopplungen mit

$$J'_{ij} = \frac{J_{ij}}{1 - J_{ii}}$$

um, so erhält man wieder $\xi_i^\mu h_i^\mu = 1$. Man sieht leicht, daß die J'_{ij} Fixpunkte der Adaline–Regel (Gl.(3.15)) *ohne* Hauptdiagonale sind. Wir schreiben für $i \neq j$:

$$\begin{aligned} \delta J'_{ij} &= \frac{1}{N} \left(1 - \xi_i^\nu \sum_{k(\neq i)} J'_{ik} \xi_k^\nu \right) \xi_i^\nu \xi_j^\nu \\ &= \frac{1}{N} \left(1 - \xi_i^\nu \sum_{k(\neq i)} \frac{J_{ik}}{1 - J_{ii}} \xi_k^\nu \right) \xi_i^\nu \xi_j^\nu = 0 . \end{aligned} \quad (3.19)$$

Es ist also egal, ob man $J_{ij} - \delta_{ij} J_{ii}$ nach Gl.(3.15) oder nach Gl.(3.19) lernt.

Zum Schluß seien noch die für die vorliegende Arbeit wichtigsten Ergebnisse der Molekularfeldtheorie für $\tau = 0$ für das Modell (3.18) zusammengefaßt ([KaSo87]):

1. Die Muster sind die stabilen Grundzustände für alle $\alpha < 1$.

¹Das Modell mit dem Hamiltonian (2.11) für die Paralleldynamik wurde bisher nicht betrachtet. Dies geschah lediglich beim Hopfield–Modell (siehe [FoKo87]). Dort ergab sich sogar, daß das serielle und das parallele Modell die gleichen Gleichgewichtseigenschaften haben.

2. Bei kleineren α -Werten gibt es lokalstabile Zustände, die Überlapps $m_\nu = \frac{1}{N} \sum_{i=1}^N S_i \xi_i^\nu$ mit mehreren Mustern ν haben:

$$S_i = \text{sign} \left(\sum_{\nu=1}^s a_\nu \xi_i^\nu \right) . \quad (3.20)$$

3. Für

$$\alpha < \alpha_g = 1 - \frac{2}{\pi} \sim 0.363 \quad (3.21)$$

gibt es lokalstabile („Spinglas“-) Zustände, die Überlapp $m_\nu = 0$ (oder Hamming-Abstand $g = \frac{1}{2}$) mit jedem Muster haben.² Für $\alpha \geq \alpha_g$ läßt die Molekularfeldtheorie keine solchen Spinglaszustände mehr als thermodynamisch stabile Zustände zu.³ Für ihre Energie gilt an der Grenze des erlaubten Bereichs

$$\varepsilon = -\frac{E_g(\alpha)}{N} \rightarrow \frac{1}{2} \quad (\alpha \rightarrow \alpha_g) . \quad (3.22)$$

Die vorliegende Arbeit liefert hingegen eine untere und eine obere Schranke für die Energie der Spinglaszustände und beweist, daß das obige Ergebnis der Molekularfeldtheorie falsch sein muß (siehe Abschnitt 5.3).

3.3 Selbstmittelungseigenschaft der Determinanten der Korrelationsmatrix \mathbf{C}

In Abschnitt 2.3.3 wurde erwähnt, daß Ausdrücke der Form e^{Nk} selbstmittelnd genannt werden, falls in der Ungleichung (2.23) für $N \rightarrow \infty$ das Gleichheitszeichen steht:

$$\lim_{N \rightarrow \infty} \langle \langle k \rangle \rangle = \lim_{N \rightarrow \infty} \frac{1}{N} \ln \langle \langle e^{Nk} \rangle \rangle . \quad (3.23)$$

Nun wird k durch

$$(\det \mathbf{C})^{-\frac{1}{2}} = e^{Nk}$$

²Ich erinnere daran, daß die Spinglas-Zustände außerdem den Edwards-Anderson-Parameter $q \neq 0$ aufweisen.

³Es gibt natürlich auch keine thermodynamisch stabilen paramagnetischen Zustände mit $q = 0$.

definiert, und später werden die beiden Seiten in Gl.(3.23) verglichen. $\langle\langle \dots \rangle\rangle$ bezeichnet wieder die Mittelung über die Zufallsmuster $\{\xi_i^\mu\}$. In Abschnitt 3.3.1 wird der direkte Mittelwert berechnet, und in Abschnitt 3.3.2 wird die Replica-Rechnung für die linke Seite in Gl.(3.23) durchgeführt. Es wird sich zeigen, daß der Edwards-Anderson-Ordnungsparameter verschwindet, die Replica sind nicht korreliert und Gl.(3.23) wird bewiesen.

Analog zur Betrachtung, die Ungleichung (2.22) vorausgeht, nehmen wir dann für die Wahrscheinlichkeitsverteilung von k eine Gaußverteilung an:

$$P(k) \propto \exp \left[-\frac{(k - k_0)^2}{2(\sigma N^{-\alpha})^2} \right].$$

Dabei muß $\alpha > \frac{1}{2}$ sein, damit in der Jensenschen Ungleichung (2.23) das Gleichheitszeichen steht. Wir erwarten also, daß dann

$$\det \mathbf{C} = e^{-2Nk} \quad (3.24)$$

auch selbstmittelnd ist. Die Selbstmittelungseigenschaft von $\det \mathbf{C}$ ist von zentraler Bedeutung für die Rechnung in Kapitel 4.

3.3.1 Der direkte Mittelwert $\langle\langle (\det \mathbf{C})^{-\frac{1}{2}} \rangle\rangle$

Ich berechne

$$I = \langle\langle (\det \mathbf{C})^{-\frac{1}{2}} \rangle\rangle$$

nach Gl.(A.2) und erhalte mit $C_{\mu\nu} = \frac{1}{N} \sum_{i=1}^N \xi_i^\mu \xi_i^\nu$

$$I = \langle\langle \left(\prod_{\nu=1}^p \int_{-\infty}^{+\infty} \frac{dx_\nu}{\sqrt{2\pi}} \right) \exp \left[-\frac{1}{2} \sum_{i=1}^N \left(\sum_{\mu=1}^p \frac{1}{\sqrt{N}} \xi_i^\mu x_\mu \right)^2 \right] \rangle\rangle$$

Es wird eine Hubbard-Stratonovich-Transformation nach Gl.(A.1) durchgeführt. Man erhält

$$I = \left(\prod_{\nu=1}^p \int_{-\infty}^{+\infty} \frac{dx_\nu}{\sqrt{2\pi}} \right) \left(\prod_{j=1}^N \int_{-\infty}^{+\infty} \frac{dp_j}{\sqrt{2\pi}} \right) \exp \left(-\frac{1}{2} \sum_{j=1}^N p_j^2 \right) \cdot \langle\langle \exp \left[\frac{i}{\sqrt{N}} \sum_{j,\mu} \xi_j^\mu x_\mu p_j \right] \rangle\rangle \quad (3.25)$$

Die Mittelung über die Muster vollzieht sich wie in Anhang C.1. Dort ist eine analoge Mittelung durchgeführt, die auf Kapitel 4 Bezug nimmt. Anschließend schreibe ich

$$R^2 = \sum_{i=1}^N p_i^2$$

und führe die Integrationen über die x_μ aus.

$$I = \left(\prod_{i=1}^N \int_{-\infty}^{+\infty} \frac{dp_i}{\sqrt{2\pi}} \right) \exp\left(-\frac{1}{2}R^2\right) \cdot \left(\frac{N}{R^2}\right)^{\frac{p}{2}}. \quad (3.26)$$

Jetzt wird eine Transformation in Kugelkoordinaten durchgeführt (siehe dazu [Wa71]). Man erhält

$$I = (2\pi)^{-N/2} N^{p/2} \cdot S(N) \cdot \int_0^R dR R^{N-p-1} e^{-\frac{1}{2}R^2}. \quad (3.27)$$

Dabei ist $S(N)$ die Oberfläche der N -dimensionalen Einheitskugel:

$$S(N) = 2 \cdot \frac{\pi^{N/2}}{\Gamma\left(\frac{N}{2}\right)}. \quad (3.28)$$

Da $\alpha < 1$, also $N > p$ ist, konvergiert das Gauß-Integral. Das Ergebnis ist

$$\langle\langle (\det \mathbf{C})^{-\frac{1}{2}} \rangle\rangle = \left(\frac{N}{2}\right)^{\frac{p}{2}} \cdot \frac{\Gamma\left(\frac{N}{2}(1-\alpha)\right)}{\Gamma\left(\frac{N}{2}\right)}. \quad (3.29)$$

Dabei wurde schon bei der Mittelung über die Muster N als groß angenommen. Mit der Stirling-Formel erhält man den asymptotischen Ausdruck

$$\langle\langle (\det \mathbf{C})^{-\frac{1}{2}} \rangle\rangle \rightarrow (1-\alpha)^{-\frac{1}{2}} \cdot \exp\left[N\left(\frac{\alpha}{2} + \frac{1-\alpha}{2} \ln(1-\alpha)\right)\right] \quad (3.30)$$

3.3.2 Die Replica-Rechnung für $\frac{1}{N} \langle \langle \ln(\det \mathbf{C})^{-\frac{1}{2}} \rangle \rangle$

Ich setze

$$Z = (\det \mathbf{C})^{-\frac{1}{2}} \quad (3.31)$$

und wende den Replica-Trick (siehe [SK75] und [SK78]) an:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \langle \langle \ln Z \rangle \rangle = \lim_{N \rightarrow \infty} \frac{1}{N} \lim_{n \rightarrow 0} \frac{\langle \langle Z^n \rangle \rangle - 1}{n}. \quad (3.32)$$

Analog zu Gl.(3.25) erhält man nach Einführung des Replica-Index ρ , $\rho \in \{1, \dots, n\}$:

$$\begin{aligned} \langle \langle Z^n \rangle \rangle &= \left(\prod_{\rho, \nu} \int_{-\infty}^{+\infty} \frac{dx_\nu^\rho}{\sqrt{2\pi}} \right) \left(\prod_{\rho, j} \int_{-\infty}^{+\infty} \frac{dp_j^\rho}{\sqrt{2\pi}} \right) \exp \left[-\frac{1}{2} \sum_{\rho, j} (p_j^\rho)^2 \right] \\ &\cdot \left\langle \left\langle \exp \left[\frac{i}{\sqrt{N}} \sum_{j, \mu} \xi_j^\mu \left(\sum_{\rho} p_j^\rho x_\mu^\rho \right) \right] \right\rangle \right\rangle. \end{aligned} \quad (3.33)$$

Die Mittelung über die Muster führt jetzt zur Doppelsumme über ρ und σ . Ich führe die Matrix

$$\mathbf{P} = (P^{\rho\sigma}) \quad \text{mit} \quad P^{\rho\sigma} = \frac{1}{N} \sum_{j=1}^N p_j^\rho p_j^\sigma$$

ein. Die $P^{\rho\sigma}$ werden mit Hilfe von δ -Funktionen ersetzt. Die zu $P^{\rho\sigma}$ konjugierten Größen heißen $\hat{P}^{\rho\sigma}$. Nach einigen elementaren Rechnungen ergibt sich

$$\begin{aligned} \langle \langle Z^n \rangle \rangle &= \left(\prod_{\rho, \sigma} \int_{-\infty}^{+\infty} dP^{\rho\sigma} \int_{-i\infty}^{+i\infty} \frac{d\hat{P}^{\rho\sigma}}{2\pi i N/2} \right) \cdot \\ &\cdot \exp \left(N \left[-\frac{1}{2} \text{Tr} \mathbf{P} + \frac{1}{2} \text{Tr}(\hat{\mathbf{P}} \mathbf{P}) \right] \right) \cdot \\ &\cdot \left(\left(\prod_{\rho} \int_{-\infty}^{+\infty} \frac{dp^\rho}{\sqrt{2\pi}} \right) \exp \left(-\frac{1}{2} \sum_{\rho, \sigma} p^\rho \hat{P}^{\rho\sigma} p^\sigma \right) \right)^N \cdot \\ &\cdot \left(\left(\prod_{\rho} \int_{-\infty}^{+\infty} \frac{dx^\rho}{\sqrt{2\pi}} \right) \exp \left(-\frac{1}{2} \sum_{\rho, \sigma} x^\rho P^{\rho\sigma} x^\sigma \right) \right)^{N\alpha} \end{aligned}$$

Ich mache nun die Standard–Annahmen der Replica–symmetrischen Rechnung:

1. Die Grenzprozesse $N \rightarrow \infty$ und $n \rightarrow 0$ vertauschen. Genau genommen wurde dies bereits bei der Mittelung über die Muster vorausgesetzt.
2. Die Sattelpunktwerte der $P^{\rho\sigma}$ und $\hat{P}^{\rho\sigma}$ weisen folgende „Replica–Symmetrie“ auf:

$$\begin{aligned} P^{\rho\sigma} &= q, & \hat{P}^{\rho\sigma} &= \hat{q} \quad \text{für } \rho \neq \sigma, \\ P^{\rho\rho} &= Q, & \hat{P}^{\rho\rho} &= \hat{Q}. \end{aligned} \quad (3.34)$$

Es wird sich herausstellen, daß \mathbf{P} und $\hat{\mathbf{P}}$ dann positiv definit sind. Wir wenden also Formel (A.2) an und führen die Integrationen über die p^ρ und x^ρ aus.

\mathbf{P} hat also am Sattelpunkt die Eigenwerte

$$\lambda_1 = Q - q \quad \text{und} \quad \lambda_2 = Q + (n - 1)q .$$

λ_1 ist $(n - 1)$ –fach entartet und λ_2 ist einfach. Dann gilt

$$\ln(\det \mathbf{P}) = \text{Tr}(\ln \mathbf{P}) = (n - 1) \ln \lambda_1 + \ln \lambda_2 .$$

Die Ergebnisse für $\hat{\mathbf{P}}$ sind analog. Nach einigen Rechnungen erhält man mit Hilfe der obigen Formeln

$$\langle\langle Z^n \rangle\rangle \propto e^{Nnf} . \quad (3.35)$$

Dabei ist f der Sattelpunktwert bezüglich Q, q, \hat{Q}, \hat{q} von

$$\begin{aligned} f(Q, q, \hat{Q}, \hat{q}, \alpha) &= -\frac{1}{2} \left(Q + q\hat{q} - Q\hat{Q} + \ln(\hat{Q} - \hat{q}) + \frac{\hat{q}}{\hat{Q} - \hat{q}} + \right. \\ &\quad \left. + \alpha \ln(Q - q) + \alpha \frac{q}{Q - q} \right) + \mathcal{O}(n) . \end{aligned} \quad (3.36)$$

Nach Ableiten und Lösen des einfachen Gleichungssystems erhält man

$$q = \hat{q} = 0 . \quad (3.37)$$

Die Nebendiagonalelemente verschwinden also. Beim SK–Modell handelt es sich bei q um den Edwards–Anderson–Parameter. Ist er Null, so gibt es keine

Korrelationen zwischen den Replica. Der direkte Mittelwert der Zustands-
summe und der Mittelwert über die Freie Energie pro Spin liefern dann ver-
gleichbare Ergebnisse.

In der Tat erhalten wir aus Gl.(3.37) die Werte

$$Q = 1 - \alpha \quad \text{und} \quad \hat{Q} = \frac{1}{1 - \alpha} ,$$

und nach Einsetzen in f und Ausführen des Grenzwertes $n \rightarrow 0$ in Gl.(3.35)
ergibt sich schließlich aus Gl.(3.32)

$$\lim_{N \rightarrow \infty} \frac{1}{N} \langle \langle \ln (\det \mathbf{C})^{-\frac{1}{2}} \rangle \rangle = \frac{\alpha}{2} + \frac{1 - \alpha}{2} \ln(1 - \alpha) .$$

Durch Vergleich mit Gl.(3.30) erhalten wir das Endergebnis.

Es gilt

$$\begin{aligned} \lim_{N \rightarrow \infty} \langle \langle \frac{1}{N} \ln (\det \mathbf{C}) \rangle \rangle &= \lim_{N \rightarrow \infty} \frac{1}{N} \ln (\langle \langle \det \mathbf{C} \rangle \rangle) \\ &= -\alpha - (1 - \alpha) \ln(1 - \alpha) . \end{aligned} \quad (3.38)$$

Aus Gl.(3.29) lesen wir die Vorfaktoren ab:

$$\langle \langle \det \mathbf{C} \rangle \rangle = \exp (\langle \langle \ln (\det \mathbf{C}) \rangle \rangle) = \left(\frac{2}{N} \right)^p \left(\frac{\Gamma(\frac{N}{2})}{\Gamma(\frac{N}{2}(1 - \alpha))} \right)^2 . \quad (3.39)$$

3.4 Berechnung der Wahrscheinlichkeitsverteilung für J_{ii}

Die Kenntnis der Verteilung ist relevant für Kapitel 5. Das Projektornetzwerk ohne Selbstkopplungen kann dort nämlich nicht exakt behandelt werden. Für eine Diskussion der Näherung

$$J_{ii} \sim \langle\langle J_{ii} \rangle\rangle$$

muß man den Mittelwert von J_{ii} und die Standardabweichung kennen. Der Mittelwert von J_{ii} ergibt sich sofort aus Gl.(3.12) zu

$$\langle\langle J_{ii} \rangle\rangle = \left\langle \left\langle \frac{1}{N} \sum_{i=1}^N J_{ii} \right\rangle \right\rangle = \alpha . \quad (3.40)$$

Der erlaubte Bereich für J_{ii} ist nach [KaSo87] ersichtlich aus

$$\begin{aligned} \sum_{i,j} J_{ij}^2 = J_{ii} \geq 0 &\iff J_{ii} - J_{ii}^2 = \sum_{i,j(\neq i)} J_{ij}^2 \geq 0 \\ &\implies 0 \leq J_{ii} \leq 1 . \end{aligned} \quad (3.41)$$

Da wir zufällig verteilte Muster voraussetzen, reicht es aus, die Wahrscheinlichkeitsverteilung von J_{11} zu betrachten. Für die Verteilungsfunktion gilt

$$\begin{aligned} F(J) &= \int_0^J P(J_{11}) dJ_{11} = \int_0^1 P(J_{11}) \cdot \Theta(J - J_{11}) dJ_{11} \\ &= \langle\langle \Theta(J - J_{11}) \rangle\rangle_{J_{11}} = \langle\langle \Theta(J - J_{11}(\{\xi_i^\mu\})) \rangle\rangle_{\{\xi_i^\mu\}} . \end{aligned}$$

Durch Ableiten erhält man also die Wahrscheinlichkeitsdichte

$$P(J) = \langle\langle \delta(J - J_{11}) \rangle\rangle . \quad (3.42)$$

Dabei sind

$$J_{11} = \frac{1}{N} \sum_{\mu,\nu} \xi_1^\mu (C^{-1})_{\mu\nu} \xi_1^\nu$$

aus Gl.(3.13) und

$$C_{\mu\nu} = \frac{1}{N} \sum_{i=1}^N \xi_i^\mu \xi_i^\nu$$

aus Gl.(3.14) bekannt.

Nach Einführung der Fourier-Darstellung für die δ -Funktion ist der Fourier-Koeffizient $a(k)$ in

$$P(J) = \int_{-i\infty}^{+i\infty} \frac{dk}{2\pi i} a(k) e^{kJ} \quad (3.43)$$

zu berechnen. Nach Einsetzen von J_{11} erhält man mit Gl.(A.2)

$$\begin{aligned} a(k) &= \left\langle \left\langle \exp\left(\frac{1}{2}\vec{m}^T \mathbf{C}^{-1} \vec{m}\right) \right\rangle \right\rangle \\ &= \left\langle \left\langle \sqrt{\det \mathbf{C}} \left(\prod_{\nu=1}^p \int_{-\infty}^{+\infty} \frac{dx_\nu}{\sqrt{2\pi}} \right) \exp\left(-\frac{1}{2}\vec{x}^T \mathbf{C} \vec{x} + \vec{m}^T \vec{x}\right) \right\rangle \right\rangle \quad (3.44) \end{aligned}$$

mit

$$\vec{m} = i\sqrt{\frac{2k}{N}} \underline{\xi}_-^1.$$

$\det \mathbf{C}$ ist nach Abschnitt 3.3 selbstmittelnd und kann aus der Mittelung über die Muster herausgezogen werden. Nach einer Hubbard-Stratonovich-Transformation ist

$$\begin{aligned} a(k) &= \left\langle \left\langle \sqrt{\det \mathbf{C}} \right\rangle \right\rangle \cdot \left(\prod_{\nu=1}^p \int_{-\infty}^{+\infty} \frac{dx_\nu}{\sqrt{2\pi}} \right) \left(\prod_{j=1}^N \int_{-\infty}^{+\infty} \frac{dp_j}{\sqrt{2\pi}} \right) \\ &\quad \exp\left(-\frac{1}{2} \sum_{j=1}^N p_j^2\right) \cdot \left\langle \left\langle \exp\left(\frac{i}{\sqrt{N}} \sum_{j,\mu} \xi_j^\mu (p_j x_\mu + \delta_{j1} \sqrt{2k} x_\mu)\right) \right\rangle \right\rangle. \end{aligned}$$

δ_{j1} ist das Kronecker-Delta.

Die weitere Rechnung geht analog zu Abschnitt 3.3.1 vor sich. Nach der Mittelung über die Muster werden die Gauß-Integrale über die p_j ausgeführt. Anschließend führt man Kugelkoordinaten

$$R^2 = \sum_{\nu=1}^p \frac{x_\nu^2}{N}$$

ein und erhält

$$\begin{aligned} a(k) &= \left\langle \left\langle \sqrt{\det \mathbf{C}} \right\rangle \right\rangle \cdot N^{p/2} (2\pi)^{-p/2} \\ &\quad \int_0^\infty dR R^{p-1} S(p) \cdot (1 + R^2)^{-N/2} \exp\left(-k \frac{R^2}{1 + R^2}\right). \quad (3.45) \end{aligned}$$

$S(p)$ ist durch Gl.(3.28) gegeben.
Nach der Transformation

$$J' = \frac{R^2}{1 + R^2} \iff R = \sqrt{\frac{J'}{1 - J'}} \quad (3.46)$$

sieht man, daß wie erwartet $0 \leq J' \leq 1$ gilt. Nach Einsetzen in Gl.(3.43) erhalten wir

$$P(J) = \langle\langle \sqrt{\det \mathbf{C}} \rangle\rangle \cdot S(p) \cdot N^{p/2} (2\pi)^{-p/2} \cdot \frac{1}{2} \cdot \int_0^1 dJ' (J')^{p/2-1} \cdot (1 - J')^{\frac{N-p}{2}-1} \delta(J - J') . \quad (3.47)$$

Wir setzen das zu Gl.(3.39) analoge Ergebnis für $\langle\langle \sqrt{\det \mathbf{C}} \rangle\rangle$ sowie $S(p)$ ein. Man erhält dann das folgende Endergebnis, das *sogar* die Normierungskonstante für $P(J)$ richtig wiedergibt.

Endergebnis

Die Wahrscheinlichkeitsdichte $P(J)$ für die Selbstkopplungen beim Neuronalen Netzwerk mit Projektorkopplungen ist eine β -Verteilung (siehe [AbSt65]) mit Mittelwert

$$\langle\langle J_{ii} \rangle\rangle = \alpha \quad (3.48)$$

und Streuung

$$\sigma = \sqrt{\frac{\alpha(1 - \alpha)}{\frac{N}{2} + 1}} . \quad (3.49)$$

Es gilt

$$P(J) = \frac{\Gamma\left(\frac{N}{2}\right)}{\Gamma\left(\frac{N}{2}\alpha\right)\Gamma\left(\frac{N}{2}(1 - \alpha)\right)} J^{\frac{N}{2}\alpha-1} (1 - J)^{\frac{N}{2}(1-\alpha)-1} . \quad (3.50)$$

Voraussetzung ist zum einen die zufällige Verteilung der Muster mit $p(\xi_i^\mu) = \frac{1}{2}$ und zum anderen der thermodynamische Limes $N \rightarrow \infty$.

Der Vollständigkeit halber geben wir noch Mittelwert und Streuung für J_{ij} , $i \neq j$, an. Hier können wir näherungsweise eine Gaußverteilung mit Mittelwert

$$\langle\langle J_{ij} \rangle\rangle = 0 \quad (3.51)$$

annehmen.

Die Standardabweichung erhält man aus der Betrachtung

$$\langle\langle J_{ij}^2 \rangle\rangle = \frac{1}{N-1} \left\langle \left\langle \sum_{j(\neq i)} J_{ij}^2 \right\rangle \right\rangle = \frac{1}{N-1} (\alpha - \langle\langle J_{ii}^2 \rangle\rangle).$$

Aus Gl.(3.49) wissen wir

$$\langle\langle J_{ii}^2 \rangle\rangle = \alpha^2 + \frac{\alpha(1-\alpha)}{\frac{N}{2} + 1}.$$

Also gilt

$$\sigma(J_{ij}) \sim \sqrt{\frac{\alpha(1-\alpha)}{N}}. \quad (3.52)$$

Die Berechnung der exakten Wahrscheinlichkeitsverteilung für J_{ij} ist viel schwieriger als im obigen Fall für J_{ii} .

Kapitel 4

Berechnung der Zahl der metastabilen Zustände bei einem Neuronalen Netzwerk mit Projektorkopplungen

Die Berechnung des direkten Mittelwertes $\langle\langle\mathcal{N}\rangle\rangle$ der Zahl \mathcal{N} der metastabilen Zustände vollzieht sich hier ähnlich wie beim Hopfield-Modell (siehe Abschnitt 2.4 und Anhang B). Aufgabe dieses Kapitels ist, die wichtigsten Schritte der Rechnung vorzuführen und zu erläutern. In Anhang C finden sich viele Nebenrechnungen. In Kapitel 5 werden die Ergebnisse der Rechnung vorgestellt.

\mathcal{N} wird wie in Abschnitt 2.4 in Abhängigkeit von α, g und ε betrachtet. Außerdem kommt noch die Stabilität γ hinzu. Zur Erinnerung: α ist die Proportionalitätskonstante in $p = \alpha N$, wobei p die Zahl der Muster ist und N die Zahl der Neuronen darstellt. Unter g versteht man den Hamming-Abstand, den der betrachtete Zustand \underline{S} zum ersten Muster $\underline{\xi}^1$ hat, $N \cdot g =$ Zahl der falschstehenden Bits. ε ist die negative Energie pro Spin, $\varepsilon = -\frac{E}{N}$.

Die Stabilität γ schränkt die metastabilen Zustände auf diejenigen Zustände ein, für die die „lokalen Energien“

$$\lambda_i = S_i h_i = S_i \sum_{j=1}^N J_{ij} S_j$$

die Bedingungen

$$\lambda_i > \gamma \quad \forall i \quad (4.1)$$

mit $\gamma \geq 0$ erfüllen. Die λ_i sind alle größer als eine vorgegebene Schranke, und mit steigendem γ sinken die oberen Schranken für die Gesamtenergie

$$E = -\frac{1}{2} \sum_{i=1}^N \lambda_i .$$

Man erwartet also, daß mit steigendem γ die „stabileren“ metastabilen Zustände übrigbleiben und die „instabileren“ aus dem Raster herausfallen. Die Einführung von γ an dieser Stelle geschieht vor allem, weil man später in Kapitel 5 den Fall $\gamma = \alpha$ mit dem Modell ohne die Selbstkopplungen J_{ii} in Verbindung bringen will. Bei jenem Modell fehlen in den inneren Feldern h_i die J_{ii} :

$$h_i = \sum_{j(\neq i)} J_{ij} S_j = \sum_{j=1}^N (J_{ij} - \delta_{ij} J_{jj}) S_j \quad \forall i . \quad (4.2)$$

In Kapitel 5 wird sich zeigen, daß das Modell

$$h_i' = \sum_{j=1}^N (J_{ij} - \delta_{ij} \alpha) S_j \quad \forall i \quad (4.3)$$

wegen $\langle\langle J_{ii} \rangle\rangle = \alpha$ (siehe Abschnitt 3.4) wichtige Eigenschaften aufweist, die man vom Modell in Gl.(4.2) erwarten würde. „Einfache“ metastabile Zustände des Modells in Gl.(4.3) erfüllen

$$\lambda_i' = h_i' S_i > 0 \quad \forall i$$

und stellen selbstverständlich metastabile Zustände des Modells in Gl.(4.1) mit Stabilität $\gamma = \alpha$ dar.

Man stellt sich nun zunächst vor, daß $p = \alpha N$ Zufallsmuster (mit $p(\xi_i^\mu) = \frac{1}{2}$ vorher ausgewürfelt) *fest* vorgelegt seien. Im Fall $\alpha < 1$ sind diese dann für $N \rightarrow \infty$ fast sicher linear unabhängig und die Projektormatrix existiert (siehe Abschnitt 3.2). Später mittelt man über die Verteilung der Zufallsmuster und bildet den direkten Mittelwert $\langle\langle \mathcal{N}(\alpha, g, \gamma, \varepsilon) \rangle\rangle$.

Wie in Abschnitt 2.3 und 2.4 wird sich zeigen, daß

$$\langle\langle \mathcal{N}(\alpha, g, \gamma, \varepsilon) \rangle\rangle \propto \exp(N f(\alpha, g, \gamma, \varepsilon))$$

asymptotisch für $N \rightarrow \infty$ gilt, wobei $f(\alpha, g, \gamma, \varepsilon)$ der Sattelpunktwert einer Funktion f bezüglich mehrerer Sattelpunktvariablen sein wird. In Abschnitt 4.1 wird zuerst das allgemeine Ergebnis $f(\alpha, g, \gamma, \varepsilon)$ hergeleitet. In den restlichen Punkten werden die Sattelpunktgleichungen für Spezialfälle aufgestellt und einige Entwicklungen ($\alpha \rightarrow 0$ und $g \rightarrow 0$) durchgeführt.

4.1 Der Rechenweg bis zum allgemeinen Ergebnis $\langle\langle \mathcal{N}(\alpha, \mathbf{g}, \gamma, \varepsilon) \rangle\rangle$

$\mathcal{N}(\alpha, g, \gamma, E) dE$ repräsentiert wieder die Zahl der metastabilen Zustände im Energieintervall $[E, E + dE]$. E war gegeben durch

$$E = -\frac{1}{2} \sum_{j,k} J_{jk} S_j S_k .$$

Zudem sollen nur Zustände mit $\lambda_i = S_i \sum_{j=1}^N J_{ij} S_j > \gamma \quad \forall i$ gezählt werden, so daß $\Theta(\lambda_i - \gamma)$ anstelle von $\Theta(\lambda_i)$ in Gl.(2.14) gesetzt wird. Man beachte zudem, daß in der vorliegenden Rechnung die Selbstkopplungen in E und λ_i mitgezählt werden. Berechnet wird zunächst $\mathcal{N}(\alpha, g, \gamma, E)$; analog Gl.(2.14) ergibt sich unter Verwendung der Darstellung (2.13) der Θ -Funktion

$$\begin{aligned} \mathcal{N}(\alpha, g, \gamma, E) &= \\ &= \sum'_{\{S_i\}} \left(\prod_{j=1}^N \Theta \left(S_j \sum_{k=1}^N J_{jk} S_k - \gamma \right) \right) \cdot \delta \left(E - \left(-\frac{1}{2} \sum_{j,k} J_{jk} S_j S_k \right) \right) \\ &= \sum'_{\{S_i\}} \left(\prod_{j=1}^N \int_0^{\infty} d\lambda'_j \delta \left(\lambda'_j - \left(S_j \sum_{k=1}^N J_{jk} S_k - \gamma \right) \right) \right) \cdot \\ &\quad \cdot \delta \left(E - \left(-\frac{1}{2} \sum_{j,k} J_{jk} S_j S_k \right) \right) \\ &= \sum'_{\{S_i\}} \left(\prod_{j=1}^N \int_{\gamma}^{\infty} d\lambda_j \int_{-i\infty}^{+i\infty} \frac{d\mu_j}{2\pi i} \right) \exp \left(\sum_{j=1}^N \mu_j \left(\lambda_j - S_j \sum_{k=1}^N J_{jk} S_k \right) \right) \cdot \\ &\quad \cdot \int_{-i\infty}^{+i\infty} \frac{dw}{2\pi i} \exp \left(wE + \frac{1}{2} w \sum_{j=1}^N \lambda_j \right) , \end{aligned} \tag{4.4}$$

wobei $\lambda_j = \lambda'_j + \gamma$ substituiert wurde und $\delta(\lambda_j - S_j \sum_{k=1}^N J_{jk} S_k)$ verwendet wurde, um $\frac{1}{2} \sum_{j,k} J_{jk} S_j S_k$ durch $\frac{1}{2} \sum_{j=1}^N \lambda_j$ zu ersetzen. Der Strich an der Summe deutet wieder an, daß nur über Zustände mit vorgegebenem Hamming-Abstand g zum 1.Muster summiert wird.

4.1.1 Rechenweg bis zur Mittelung über die Muster

Setzt man die Projektorkopplungen

$$J_{jk} = \frac{1}{N} \sum_{\mu,\nu} \xi_j^\mu (C^{-1})_{\mu\nu} \xi_k^\nu$$

in Gl.(4.4) ein, so stellt die Beseitigung von C^{-1} aus dem asymmetrischen Exponenten

$$- \sum_{\mu,\nu} \left(\frac{1}{\sqrt{N}} \sum_{j=1}^N \mu_j S_j \xi_j^\mu \right) (C^{-1})_{\mu\nu} \left(\frac{1}{\sqrt{N}} \sum_{k=1}^N S_k \xi_k^\nu \right)$$

ein Problem dar. Dieses wurde zunächst durch Symmetrisierung des Exponenten und anschließender Verwendung der Gauß-Integral-Formel (A.2) gelöst. Beim Zusammenstellen dieser Arbeit fand ich schließlich folgenden eleganten Lösungsweg, der mit den Entwicklungskoeffizienten a_μ und den Überlapps m_ν im Ausdruck

$$\begin{aligned} \sum_{k=1}^N J_{jk} S_k &= \sum_{k=1}^N \frac{1}{N} \sum_{\mu,\nu} \xi_j^\mu (C^{-1})_{\mu\nu} \xi_k^\nu S_k = \sum_{\mu,\nu} (C^{-1})_{\mu\nu} \left(\frac{1}{N} \sum_{k=1}^N \xi_k^\nu S_k \right) \xi_j^\mu \\ &= \sum_{\mu=1}^p \left(\sum_{\nu=1}^p (C^{-1})_{\mu\nu} m_\nu \right) \xi_j^\mu = \sum_{\mu=1}^p a_\mu \xi_j^\mu \end{aligned} \quad (4.5)$$

arbeitet und die gleichen Ergebnisse liefert.

Wir ersetzen die m_ν in Gl.(4.4) mit Hilfe von δ -Funktionen und schreiben

$$\begin{aligned} \mathcal{N}(\alpha, g, \gamma, E) &= \\ &\sum'_{\{S_i\}} \left(\prod_{j=1}^N \int_{\gamma}^{\infty} d\lambda_j \int_{-i\infty}^{+i\infty} \frac{d\mu_j}{2\pi i} \right) \left(\prod_{\nu=1}^p \int_{-i\infty}^{+i\infty} dm_\nu \int_{-i\infty}^{+i\infty} \frac{dl_\nu}{2\pi i} \right) \\ &\exp \left[\sum_{j=1}^N \mu_j \left(\lambda_j - S_j \sum_{\mu,\nu} (C^{-1})_{\mu\nu} m_\nu \xi_j^\mu \right) \right] \cdot \exp \left[\sum_{\mu=1}^p l_\mu \left(m_\mu - \frac{1}{N} \sum_{k=1}^N \xi_k^\mu S_k \right) \right]. \end{aligned}$$

$$\cdot \int_{-i\infty}^{+i\infty} \frac{dw}{2\pi i} \exp \left(wE + \frac{1}{2} w \sum_{j=1}^N \lambda_j \right) \quad (4.6)$$

An dieser Stelle wird die Transformation

$$m_\mu = \sum_{\nu=1}^p C_{\mu\nu} a_\nu \quad \forall \mu$$

durchgeföhrt, die wir aus Gl.(4.5) ablesen können. Die Jacobi-Determinante der Transformation ist $\det \mathbf{C} = |\det \mathbf{C}|$. Unter Verwendung von Gl.(4.5) erhält man mit der Korrelationsmatrix

$$(C_{\mu\nu}) = \left(\frac{1}{N} \sum_{j=1}^N \xi_j^\mu \xi_j^\nu \right) = \left(\frac{1}{N} \sum_{j=1}^N \xi_j^\mu S_j \xi_j^\nu S_j \right)$$

die Gleichung

$$\begin{aligned} \mathcal{N}(\alpha, g, \gamma, E) = & \sum'_{\{S_i\}} \left(\prod_{j=1}^N \int_{\gamma}^{\infty} d\lambda_j \int_{-i\infty}^{+i\infty} \frac{d\mu_j}{2\pi i} \right) \left(\prod_{\nu=1}^p \int_{-\infty}^{+\infty} da_\nu \int_{-i\infty}^{+i\infty} \frac{dl_\nu}{2\pi i} \right) \det \mathbf{C} \cdot \\ & \cdot \exp \left[\sum_{j=1}^N \mu_j \left(\lambda_j - \sum_{\mu=1}^p a_\mu S_j \xi_j^\mu \right) \right] \cdot \int_{-i\infty}^{+i\infty} \frac{dw}{2\pi i} \exp \left(wE + \frac{1}{2} w \sum_{j=1}^N \lambda_j \right) \cdot \\ & \cdot \exp \left[\sum_{\mu=1}^p l_\mu \left(\sum_{\nu=1}^p \frac{1}{N} \sum_{j=1}^N S_j \xi_j^\mu S_j \xi_j^\nu a_\nu - \frac{1}{N} \sum_{k=1}^N \xi_k^\mu S_k \right) \right] \end{aligned} \quad (4.7)$$

Man identifiziert

$$\sum_{\nu=1}^p \frac{1}{N} \sum_{j=1}^N S_j \xi_j^\mu S_j \xi_j^\nu a_\nu = \frac{1}{N} \sum_{j=1}^N S_j \xi_j^\mu \left(\sum_{\nu=1}^p a_\nu S_j \xi_j^\nu \right) = \frac{1}{N} \sum_{j=1}^N S_j \xi_j^\mu \lambda_j \quad (4.8)$$

und entkoppelt so die Muster.

In Abschnitt 3.3 wurde gezeigt, daß $\det \mathbf{C}$ „selbstmittelnd“ ist. Man nimmt also N als hinreichend groß an und ersetzt $\det \mathbf{C}$ durch ihren Mittelwert (siehe Gl.(3.38)). So wird der Ausdruck $\det \mathbf{C}$ zu einer Konstanten. Es wird also nicht über die Muster gemittelt, sondern lediglich die Selbstmittelungseigenschaft für große N ausgenutzt. Die Muster selbst sind noch fest vorgegeben.

Jetzt wird die Summe aller $\{S_j\}$ mit Hamming-Abstand g zum ersten Muster ausgeführt. Für jeden solchen Summanden denkt man sich das mehrdimensionale Integral hingeschrieben. Die j -Variablen werden dann jeweils so permutiert, daß die ersten Ng zu $S_j \xi_j^1 = -1$ gehören und die restlichen $N(1-g)$ zu $S_j \xi_j^1 = +1$ passen. Das Vertauschen der Reihenfolge in den anderen Mustern $\underline{\xi}^\nu$, $\nu \geq 2$, ist unerheblich, da später über diese gemittelt wird.

Um die Mittelung über die Muster vorzubereiten, wird zunächst das *Skalierungsverhalten* der a_μ und l_μ untersucht. Die zum „kondensierten“ ersten Muster gehörenden Variablen a_1 und l_1 skalieren anders als ihre Brüder:

$$a_1 = \mathcal{O}(1), \quad l_1 = \mathcal{O}(N), \quad \text{aber}$$

$$a_\mu = \mathcal{O}\left(\frac{1}{\sqrt{N}}\right) \quad \text{und} \quad l_\mu = \mathcal{O}(\sqrt{N}) \quad \text{für} \quad \mu \geq 2.$$

Für die a_μ ist dies auch anschaulich klar, denn sie sind die Koeffizienten in der Entwicklung (4.5); so erwartet man das oben angegebene Verhalten, je nachdem, ob eine Korrelation zum ersten Muster vorliegt oder nicht.

Wir skalieren so um, daß stets das Verhalten $\mathcal{O}(1)$ vorliegt und setzen

$$l_1' = \frac{l_1}{N} \quad \text{und} \quad a_\mu' = \sqrt{N} a_\mu, \quad l_\mu' = \frac{1}{\sqrt{N}} l_\mu \quad \text{für} \quad \mu \geq 2.$$

Die Striche werden im folgenden wieder weggelassen. Nach Einsetzen von Gl.(4.8) erhält man also aus Gl.(4.7)

$$\begin{aligned} \mathcal{N}(\alpha, g, \gamma, E) = & \sum'_{\{S_i\}} \left(\prod_{j=1}^N \int_{\gamma}^{\infty} d\lambda_j \int_{-i\infty}^{+i\infty} \frac{d\mu_j}{2\pi i} \right) \left(\prod_{\nu=1}^p \int_{-i\infty}^{+i\infty} da_\nu \int_{-i\infty}^{+i\infty} \frac{dl_\nu}{2\pi i} \right) \int_{-i\infty}^{+i\infty} \frac{dw}{2\pi i} \cdot \\ & \exp \left[\sum_{j=1}^N \mu_j \lambda_j + a_1 \left(\sum_{j=1}^{Ng} \mu_j - \sum_{j=Ng+1}^N \mu_j \right) + l_1 \left(\sum_{j=1}^{Ng} (1 - \lambda_j) + \sum_{j=Ng+1}^N (\lambda_j - 1) \right) + \right. \\ & \quad \left. + wE + \frac{1}{2}w \sum_{j=1}^N \lambda_j \right] \cdot N \cdot \langle\langle \det \mathbf{C} \rangle\rangle \cdot \\ & \cdot \exp \left(\frac{1}{\sqrt{N}} \sum_{j=1}^N \sum_{\mu=2}^p S_j \xi_j^\mu (l_\mu (\lambda_j - 1) - a_\mu \mu_j) \right) \end{aligned} \quad (4.9)$$

Muster 1 ist jetzt verschwunden. Man nimmt nun an, daß die Zustände $\{S_j\}$ keine Korrelation zu den anderen Mustern besitzen. Das obige Skalierungsverhalten $a_\mu' = \mathcal{O}(1)$, $\mu \geq 2$, trifft dann zu; es gibt also keine weiteren „kondensierten Muster“. Die verbleibenden Muster werden als Zufallsvariablen mit $p(\xi_i^\mu) = \frac{1}{2}$ aufgefaßt;

$$\sigma_j^\mu = S_j \xi_j^\mu \quad , \quad \mu \geq 2,$$

ist dann ebenfalls eine Zufallsvariable, über die nun anstelle von ξ_i^μ gemittelt wird.

An dieser Stelle wird betont, daß Zustände, die mit weiteren Mustern korreliert sind, in der Rechnung nicht berücksichtigt werden. Dies liegt an der obigen Annahme für das Skalierungsverhalten der a_μ !

Wir haben also die Summe $\sum_{\{S_i\}'}$ über die Zustände *weiter* eingeschränkt; eine Untermenge vom Maß Null fällt aus der Summe heraus. Der Mittelwert $\langle\langle \mathcal{N}(\alpha, g, \gamma, E) \rangle\rangle$ macht somit nur Aussagen über die unkorrelierten Zustände. Dies kann man später (siehe Kapitel 5) auch an der Form der Sattelpunktvariable Z ablesen. Bereits im vorliegenden Fall, in dem nur die Korrelation zu einem Muster behandelt wird, treten zwei zusätzliche Sattelpunktvariablen oder Ordnungsparameter a_1 und l_1 auf. Im Fall $g = \frac{1}{2}$ sind diese natürlich Null. Will man Korrelationen zu mehreren Mustern berücksichtigen (= „mehrere kondensierte Muster“), so schränkt sich die Summe über die \underline{S} weiter ein. Man umschiffet diese Schwierigkeit, indem man, wie in Abschnitt 2.4 geschildert, δ -Funktionen einbaut und die Dichte

$$\mathcal{N}(\alpha, m_1, \dots, m_s) = \sum_{\{\underline{S}\}} \left(\prod_{i=1}^N \Theta(\lambda_i) \right) \cdot \delta \left(m_1 - \frac{1}{N} \sum_{i=1}^N S_i \xi_i^1 \right) \cdot \dots \cdot \delta \left(m_s - \frac{1}{N} \sum_{i=1}^N S_i \xi_i^s \right)$$

berechnet. Es treten dann $2s$ zusätzliche Ordnungsparameter auf. Setzt man $m_k = 0$, so sind die zugehörigen Parameter natürlich Null, sie brauchen dann nicht berücksichtigt zu werden.

Die Mittelung über die σ_j^μ wird ausgeführt. Dabei entwickelt man den auftretenden cosh bis zu $\mathcal{O}\left(\frac{1}{N}\right)$. Es handelt sich hierbei um ein Standardverfahren, das bei analytischen Rechnungen mit Neuronalen Netzen ständig angewendet wird. Mathematische Überlegungen dazu findet man in [HeKu90]. Man erhält (siehe C.1)

$$\begin{aligned} \left\langle \left\langle \exp \left[\frac{1}{\sqrt{N}} \sum_{j=1}^N \sum_{\mu=2}^p \sigma_j^\mu (-a_\mu \mu_j + l_\mu (\lambda_j - 1)) \right] \right\rangle \right\rangle = \\ \exp \left[\sum_{j=1}^N \frac{1}{2} \left(X (\lambda_j - 1)^2 - 2Y \mu_j (\lambda_j - 1) + Z \mu_j^2 \right) \right], \end{aligned} \quad (4.10)$$

wobei X, Y, Z vorerst noch Abkürzungen für

$$X = \frac{1}{N} \sum_{\mu=2}^p l_\mu^2, \quad Y = \frac{1}{N} \sum_{\mu=2}^p a_\mu l_\mu, \quad Z = \frac{1}{N} \sum_{\mu=2}^p a_\mu^2 \quad (4.11)$$

sind. Es ergibt sich schließlich als Endergebnis dieses Unterpunktes 4.1.1 für den direkten Mittelwert von \mathcal{N} :

$$\begin{aligned} \langle \langle \mathcal{N}(\alpha, g, \gamma, E) \rangle \rangle = \\ \left(\sum'_{\{S_i\}} \cdot 1 \right) \left(\prod_{j=1}^N \int_{\gamma}^{\infty} d\lambda_j \int_{-i\infty}^{+i\infty} \frac{d\mu_j}{2\pi i} \right) \left(\prod_{\nu=1}^p \int_{-\infty}^{+\infty} da_\nu \int_{-i\infty}^{+i\infty} \frac{dl_\nu}{2\pi i} \right) \int_{-i\infty}^{+i\infty} \frac{dw}{2\pi i} \cdot \\ N \cdot \langle \langle \det \mathbf{C} \rangle \rangle \cdot \\ \exp \left[\sum_{j=1}^N \mu_j \lambda_j + a_1 \left(\sum_{j=1}^{Ng} \mu_j - \sum_{Ng+1}^N \mu_j \right) + l_1 \left(\sum_{j=1}^{Ng} (1 - \lambda_j) + \sum_{Ng+1}^N (\lambda_j - 1) \right) + \right. \\ \left. + wE + \frac{1}{2} w \sum_{j=1}^N \lambda_j + \sum_{j=1}^N \left(\frac{1}{2} X (\lambda_j - 1)^2 - Y \mu_j (\lambda_j - 1) + \frac{1}{2} Z \mu_j^2 \right) \right] \end{aligned}$$

mit X, Y, Z aus Gl.(4.11). Mit Hilfe der Stirling-Formel

$$\ln M! \sim M \ln M - M$$

für große M erhält man noch

$$\ln \left(\sum'_{\{S_i\}} \cdot 1 \right) = \ln \binom{N}{Ng} \sim N[-g \ln g - (1-g) \ln (1-g)] \quad (4.12)$$

4.1.2 Entkopplung der Variablen in i und μ ; mehrdimensionale Sattelpunktintegration

Jetzt werden X, Y, Z aus Gl.(4.11) mit Hilfe dreier δ -Funktionen ausgedrückt. Die konjugierten Größen in den Fourier-Darstellungen der δ -Funktionen sind

$$x' = Nx, \quad y' = Ny, \quad z' = Nz .$$

Dies bewirkt eine Entkopplung der Variablen mit Index i von den Variablen mit Index μ , und es entstehen Produkte von Gauß-Integralen.

$$\begin{aligned} \langle\langle \mathcal{N}(\alpha, g, \gamma, E) \rangle\rangle = & N \cdot N^3 \cdot \binom{N}{Ng} \cdot \langle\langle \det \mathbf{C} \rangle\rangle \cdot \int_{-\infty}^{+\infty} da_1 \int_{-\infty}^{+\infty} \frac{dl_1}{2\pi i} \int_{-\infty}^{+\infty} \frac{dw}{2\pi i} \int_{-\infty}^{+\infty} dX \int_{-\infty}^{+\infty} \frac{dx}{2\pi i} \\ & \int_{-\infty}^{+\infty} \frac{dY}{i} \int_{-\infty}^{+\infty} \frac{dy}{2\pi} \int_{-\infty}^{+\infty} dZ \int_{-\infty}^{+\infty} \frac{dz}{2\pi i} \exp[wE + N(xX + yY + zZ)] \cdot \\ & \cdot \left(\int_{-\infty}^{+\infty} da \int_{-\infty}^{+\infty} \frac{dl}{2\pi i} \exp[-xl^2 - yla - za^2] \right)^{p-1} \cdot \\ & \cdot \left(\int_{\gamma}^{\infty} d\lambda \int_{-\infty}^{+\infty} \frac{d\mu}{2\pi i} \exp \left[\mu\lambda + a_1\mu + \frac{1}{2}w\lambda + l_1(1-\lambda) + \right. \right. \\ & \quad \left. \left. + \frac{1}{2}X(\lambda-1)^2 - Y\mu(\lambda-1) + \frac{1}{2}Z\mu^2 \right] \right)^{Ng} \cdot \\ & \cdot \left(\int_{\gamma}^{\infty} d\lambda \int_{-\infty}^{+\infty} \frac{d\mu}{2\pi i} \exp \left[\mu\lambda - a_1\mu + \frac{1}{2}w\lambda - l_1(1-\lambda) + \right. \right. \\ & \quad \left. \left. + \frac{1}{2}X(\lambda-1)^2 - Y\mu(\lambda-1) + \frac{1}{2}Z\mu^2 \right] \right)^{N(1-g)} \end{aligned} \quad (4.13)$$

Dabei unterscheidet sich die Basis im Ausdruck $(\dots)^{Ng}$ nur durch das Vorzeichen vor a_1 und l_1 von der Basis im Ausdruck $(\dots)^{N(1-g)}$. Die Gauß-Integrale werden mit Hilfe der Formel (A.1) berechnet (siehe Anhang C.2 und C.3). Die Variablen $x, y, z, X, Y, Z, a_1, l_1$ und w bleiben übrig und man erhält mit $\varepsilon = -\frac{E}{N}$ ein neundimensionales Integral

$$\begin{aligned}
\langle\langle \mathcal{N}(\alpha, g, \gamma, \varepsilon) \rangle\rangle = & \\
N^4 \int_{-\infty}^{+\infty} da_1 & \int_{-i\infty}^{+i\infty} \frac{dl_1}{2\pi i} \int_{-i\infty}^{+i\infty} \frac{dw}{2\pi i} \int_{-\infty}^{+\infty} dX \int_{-i\infty}^{+i\infty} \frac{dx}{2\pi i} \int_{-i\infty}^{+i\infty} \frac{dY}{i} \int_{-\infty}^{+\infty} \frac{dy}{2\pi} \\
& \int_{-\infty}^{+\infty} dZ \int_{-i\infty}^{+i\infty} \frac{dz}{2\pi i} \exp [N f(\alpha, g, \gamma, \varepsilon, x, y, z, X, Y, Z, a_1, l_1, w)] \quad (4.14)
\end{aligned}$$

mit

$$\begin{aligned}
f(\alpha, g, \gamma, \varepsilon, x, y, z, X, Y, Z, a_1, l_1, w) = & \\
& -g \ln g - (1-g) \ln(1-g) - \alpha - (1-\alpha) \ln(1-\alpha) + \\
& -w\varepsilon + xX + yY + zZ - \frac{\alpha}{2} \ln(y^2 - 4xz) - \frac{1}{2} \ln K + \frac{w^2 Z}{8K} + \\
& + \frac{w}{2K} [(K-1+Y) + (2g-1) \cdot (-l_1 Z - a_1(1-Y))] + \\
& + \frac{1}{2K} [l_1^2 Z + 2l_1 a_1(1-Y) + a_1^2 X + X + (2g-1)(2l_1(1-Y) + 2a_1 X)] + \\
& + g \ln \Phi(t_1) + (1-g) \ln \Phi(t_2) , \quad (4.15)
\end{aligned}$$

dabei sind

$$\begin{aligned}
t_1 &= \sqrt{\frac{K}{Z}} \left(1 - \gamma + \frac{1}{K} \left(\frac{1}{2} w Z - l_1 Z - a_1(1-Y) + Y - 1 \right) \right) , \\
t_2 &= \sqrt{\frac{K}{Z}} \left(1 - \gamma + \frac{1}{K} \left(\frac{1}{2} w Z + l_1 Z + a_1(1-Y) + Y - 1 \right) \right) , \\
K &= (1-Y)^2 - XZ , \quad \Phi(t) = \int_{-\infty}^t \frac{d\lambda}{\sqrt{2\pi}} e^{-\frac{1}{2}\lambda^2} .
\end{aligned}$$

An dieser Stelle kommt die Sattelpunktmethode zur Geltung:
Für $N \rightarrow \infty$ ist

$$\langle\langle \mathcal{N}(\alpha, g, \gamma, \varepsilon) \rangle\rangle \propto \exp(N f(\alpha, g, \gamma, \varepsilon)) ,$$

wobei $f(\alpha, g, \gamma, \varepsilon)$ den Sattelpunktwert der Funktion f in Gl.(4.15) über die Variablen $x, y, z, X, Y, Z, a_1, l_1$ und w darstellt. Es interessiert also nur noch das asymptotische Verhalten und somit der Wert von f am Sattelpunkt. Im folgenden wendet man zwei Methoden an, um das Problem zu vereinfachen.

1. Ist es möglich, mit Hilfe der Sattelpunktgleichungen eine Sattelpunktvariable *vollständig* durch die anderen Variablen auszudrücken, so kann man in f die betreffende Variable ersetzen und dann den Sattelpunkt bezüglich der restlichen Variablen suchen. Nach Lösung der *neuen* Sattelpunktgleichungen erhält man dann den gesuchten Sattelpunktwert von f .

Das eben Beschriebene folgt aus der Kettenregel:

Ist etwa der Sattelpunktwert von $g(x, y)$ gesucht und gilt $y = y(x)$, so ist am Sattelpunkt

$$\frac{dg(x, y(x))}{dx} = \frac{\partial g}{\partial x} + \frac{\partial g}{\partial y} \cdot y' = 0 \quad (4.16)$$

erfüllt, wenn

$$\frac{\partial g}{\partial x} = 0 \quad \text{und} \quad \frac{\partial g}{\partial y} = 0$$

erfüllt sind. Der gesuchte Sattelpunktwert für x befindet sich also unter den Lösungen von Gl.(4.16). y folgt dann aus $y = y(x)$.

2. Auch durch Variablensubstitutionen kann man f vereinfachen. Dies wird in Abschnitt 4.2 von Bedeutung sein. Ist die Jacobi-Matrix der Transformation am Sattelpunkt invertierbar, so ist die Suche nach dem Sattelpunkt bezüglich der neuen Variablen äquivalent zur Suche bezüglich der alten Variablen. Dies folgt wieder aus der Kettenregel: Betrachtet sei wieder $g(x, y)$. Man transformiert

$$x = x(u, v) \quad \text{und} \quad y = y(u, v) .$$

Nach Kettenregel gilt am Sattelpunkt bezüglich x und y

$$\frac{\partial g}{\partial x} = \frac{\partial g}{\partial u} \cdot \frac{\partial u}{\partial x} + \frac{\partial g}{\partial v} \cdot \frac{\partial v}{\partial x} = 0$$

$$\frac{\partial g}{\partial y} = \frac{\partial g}{\partial u} \cdot \frac{\partial u}{\partial y} + \frac{\partial g}{\partial v} \cdot \frac{\partial v}{\partial y} = 0 .$$

Falls also

$$\det \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix} \neq 0$$

am Sattelpunkt, so kann man äquivalent das System

$$\left(\frac{\partial g}{\partial u}, \frac{\partial g}{\partial v} \right) = (0, 0)$$

lösen.

Um den Abschnitt 4.1 abzuschließen, ersetzen wir die Variablen x, y, z mit Methode 1 durch die Variablen X, Y, Z (siehe Anhang C.4).

4.1.3 Das allgemeine Endergebnis $\langle\langle \mathcal{N}(\alpha, \mathbf{g}, \gamma, \varepsilon) \rangle\rangle$

Für den direkten Mittelwert der Zahl der metastabilen Zustände $\mathcal{N}(\alpha, g, \gamma, \varepsilon)$ gilt asymptotisch für $N \rightarrow \infty$

$$\langle\langle \mathcal{N}(\alpha, g, \gamma, \varepsilon) \rangle\rangle \propto \exp(N f(\alpha, g, \gamma, \varepsilon)) . \quad (4.17)$$

Dabei ist $f(\alpha, g, \gamma, \varepsilon)$ der Sattelpunktwert über die Variablen X, Y, Z, a_1, l_1, w von

$$\begin{aligned} f(\alpha, g, \gamma, \varepsilon, X, Y, Z, a_1, l_1, w) = & \\ & -g \ln g - (1-g) \ln(1-g) - \alpha \ln \alpha - (1-\alpha) \ln(1-\alpha) + \\ & -w\varepsilon + \frac{\alpha}{2} \ln(Y^2 - XZ) - \frac{1}{2} \ln K + \frac{w^2 Z}{8K} + \\ & + \frac{w}{2K} [(K-1+Y) + (2g-1) \cdot (-l_1 Z - a_1(1-Y))] + \\ & + \frac{1}{2K} [l_1^2 Z + 2l_1 a_1(1-Y) + a_1^2 X + X] + \\ & + \frac{2g-1}{K} (l_1(1-Y) + a_1 X) + g \ln \Phi(t_1) + (1-g) \ln \Phi(t_2) , \end{aligned} \quad (4.18)$$

wobei K die Abkürzung

$$K = (1-Y)^2 - XZ$$

ist. Φ ist erklärt als

$$\Phi(t) = \int_{-\infty}^t \frac{d\lambda}{\sqrt{2\pi}} e^{-\frac{1}{2}\lambda^2} .$$

Die Argumente der Φ -Funktionen sind

$$t_1 = \sqrt{\frac{K}{Z}} \left(1 - \gamma + \frac{1}{K} \left(\frac{1}{2} wZ - l_1 Z - a_1(1-Y) + Y - 1 \right) \right) ,$$

$$t_2 = \sqrt{\frac{K}{Z}} \left(1 - \gamma + \frac{1}{K} \left(\frac{1}{2} wZ + l_1 Z + a_1(1-Y) + Y - 1 \right) \right) .$$

Die Auflösung der Sattelpunktgleichungen wurde in der vorliegenden Arbeit für zwei Spezialfälle vorgenommen.

4.2 Das Ergebnis $\langle\langle\mathcal{N}(\alpha, \mathbf{g}, \gamma, \varepsilon_{\max})\rangle\rangle$

4.2.1 Rechenweg

ε_{\max} ist der Wert, an dem $f(\alpha, g, \gamma, \varepsilon)$ bezüglich ε maximal wird. Formal folgt dann $w = 0$, wenn man mit Gl.(4.18) die Bedingung

$$\frac{\partial f}{\partial \varepsilon} = 0$$

aufstellt. $f(\alpha, g, \gamma, \varepsilon_{\max})$ ist also der Wert, den man erhalten hätte, wenn man die Energieabhängigkeit der metastabilen Zustände gar nicht berücksichtigt hätte, denn die zu E gehörende Sattelpunktvariable w ist Null.

Mit den verbleibenden Sattelpunktvariablen wird eine Variablentransformation durchgeführt (siehe C.5):

$$(X, Y, Z, l_1, a_1) \rightarrow (K, r, Z, s, a_1)$$

s ist Null, wenn a_1 und l_1 Null sind und zeigt anstelle von l_1 an, ob $g = \frac{1}{2}$ ist oder nicht (näheres siehe unter C.5).

Nach der in Abschnitt 4.1.2 geschilderten Methode 1 wird dann a_1 in $f(\alpha, g, \gamma, \varepsilon_{\max}, K, r, Z, s, a_1)$ durch $1 - 2g$ ersetzt. Man sucht anschließend den Sattelpunkt bezüglich den verbleibenden Variablen K, r, Z und s und löst das Gleichungssystem

$$\left(\frac{\partial f}{\partial K}, \frac{\partial f}{\partial r}, \frac{\partial f}{\partial Z}, \frac{\partial f}{\partial s} \right) = (0, 0, 0, 0)$$

auf (siehe C.6).

Es gelingt, K und r als einfache Funktionen von Z und s auszudrücken. Z und s hingegen müssen aus Iterationsgleichungen gewonnen werden. Die Substitutionen

$$U = s\sqrt{K} \quad \text{und} \quad V = \frac{Z}{4g(1-g)}$$

vereinfachen die numerische Behandlung des Problems. Man iteriert schließlich U und V . K und r lassen sich ebenfalls als Funktionen von U und V schreiben. Zusätzlich treten noch die Abkürzungen oder Hilfsvariablen

$$m, t_1 \quad \text{und} \quad t_2$$

auf, die wieder nur von U und V abhängen. Damit folgt das Endergebnis.

4.2.2 Das Endergebnis $\langle\langle\mathcal{N}(\alpha, \mathbf{g}, \gamma, \varepsilon_{\max})\rangle\rangle$, Sattelpunktgleichungen und Liste aller Hilfsvariablen

Der direkte Mittelwert von $\mathcal{N}(\alpha, g, \gamma, \varepsilon_{\max})$ ist im Fall $N \rightarrow \infty$ durch den Ausdruck

$$\langle\langle\mathcal{N}(\alpha, g, \gamma, \varepsilon_{\max})\rangle\rangle \propto \exp(N f(\alpha, g, \gamma, \varepsilon_{\max}))$$

bestimmt. Dabei ist f gegeben durch

$$\begin{aligned} f(\alpha, g, \gamma, \varepsilon_{\max}) &= - (1 - \alpha) \ln(1 - \alpha) - \alpha \ln \alpha - g \ln g - (1 - g) \ln(1 - g) + \\ &+ \frac{\alpha}{2} \ln(m + 1) - \frac{1}{2} \ln K + \\ &+ \frac{1}{2Z} \left(r^2 - 2r\sqrt{K}(1 - \gamma) + (1 - \gamma)^2 K - 1 + s^2 \right) + \\ &+ \frac{(2g - 1)^2}{2Z} + (2g - 1) \frac{rs - \sqrt{K}s(1 - \gamma)}{Z} + \\ &+ g \ln \Phi(t_1) + (1 - g) \ln \Phi(t_2) . \end{aligned} \quad (4.19)$$

Φ steht wieder für die Φ -Funktion (siehe auch A.3). Die Hilfsvariablen

$$Z, m, K, r, s, t_1 \quad \text{und} \quad t_2$$

hängen wie folgt von den Iterationsvariablen U und V ab:

$$Z = V \cdot 4g(1 - g) \quad (4.20)$$

$$m = \frac{V - 1}{1 - V(1 - \alpha)} \quad (4.21)$$

$$K = \frac{Z + 1 - (2g - 1)^2 + (1 - \gamma)m + 2U(2g - 1)(1 - \gamma)}{(1 - 2\gamma) \cdot (Z(1 - \alpha) + (2g - 1)^2) + \gamma} \quad (4.22)$$

$$r = \frac{1}{2\sqrt{K}}(m + (1 - 2\gamma)K) \quad (4.23)$$

$$s = \frac{U}{\sqrt{K}} \quad (4.24)$$

$$t_1 = \frac{r + s}{\sqrt{Z}} \quad (4.25)$$

$$t_2 = \frac{r-s}{\sqrt{Z}}. \quad (4.26)$$

Die Variablen U und V ergeben sich schließlich aus den Iterationsgleichungen

$$V = \frac{1}{1-\alpha} \cdot \left[1 + \frac{1}{2(1-g)} \left(\frac{m}{2K} - \frac{1}{2} + \frac{U}{K} \right) + \frac{1}{2(1-g)} \sqrt{\frac{Z}{K}} \cdot \frac{\Phi'(t_1)}{\Phi(t_1)} \right], \quad (4.27)$$

$$U = 2gK(1-V(1-\alpha)) + \frac{m-K}{2} + \sqrt{ZK} \cdot \frac{\Phi'(t_2)}{\Phi(t_2)}. \quad (4.28)$$

Alternativ zu Gl.(4.27) und Gl.(4.28) kann man auch die beiden Iterationsgleichungen (C.30) und (C.31) benutzen.

Die Gleichungen wurden mit der Regula-falsi-Methode auf dem Computer gelöst. Am günstigsten erwies sich dabei das Verfahren, V in einer äußeren Schleife und U in einer inneren Schleife zu iterieren. Mit jedem Iterationsschritt für V wurde U in der inneren Schleife an den neuen Wert für V vollständig heraniteriert, siehe auch [Schwarz].

4.2.3 Grenzwerte und Näherungen im Fall $g \rightarrow 0$

Bei der Lösung der obigen Sattelpunktgleichungen stellt man fest, daß sich die Argumente t_1 , t_2 der Φ -Funktionen folgendermaßen für $g \rightarrow 0$ verhalten:

1. $(\gamma = 0 \quad \text{und} \quad \alpha < \frac{1}{2})$ oder $(\gamma = \alpha \quad \text{und} \quad 0 < \alpha < 1)$:
 $t_1 \rightarrow -\infty$ und $t_2 \rightarrow +\infty$.
2. $\gamma = 0$ und $\alpha = \frac{1}{2}$:
 $t_1 = \mathcal{O}(1)$ und $t_2 \rightarrow +\infty$.
3. $\gamma = 0$ und $\frac{1}{2} < \alpha < 1$:
 $t_1 \rightarrow +\infty$ und $t_2 \rightarrow +\infty$.

Man macht in allen drei Fällen selbstkonsistente Ansätze, um analytische Näherungen durchzuführen. Wegen $t_2 \rightarrow +\infty$ vernachlässigt man stets

$$\Phi'(t_2) = \frac{1}{\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2}t_2^2\right)$$

in Gl.(4.28). Man kann also U als Funktion von V ausdrücken und braucht nur noch die Iterationsgleichung (4.27) für V zu lösen.

Im Fall 1 wird die Näherung (A.6) für $\Phi(t_1)$ verwendet. Setzt man V als Potenzreihe in g an, so kann man alle Koeffizienten der Reihe durch Koeffizientenvergleich in der Sattelpunktgleichung bestimmen.

Im Fall 3 wird auch $\Phi'(t_1)$ aus Gl.(4.27) gestrichen. Es gelingt, Gl.(4.28) nach V aufzulösen. Die Variablen unterscheiden sich nur durch exponentiell kleine Fehler von ihren Grenzwerten für $g \rightarrow 0$. Der exponentiell kleine Fehler der Ordnung $\mathcal{O}\left(\exp\left(-\frac{\varepsilon}{g}\right)\right)$ wurde beim Wegstreichen der Terme $\Phi'(t_1)$ und $\Phi'(t_2)$ begangen.

Die Grenzwerte der Variablen bezüglich $g \rightarrow 0$ im Fall 2 kann man bestimmen, indem man für die entsprechenden Grenzwerte der Fälle 1 bzw. 3 $\gamma = 0$ und $\alpha = \frac{1}{2}$ einsetzt. Damit das in der Numerik beobachtete $t_1 = \mathcal{O}(1)$ erfüllt ist, muß der Fehler in V von der Ordnung $\mathcal{O}(\sqrt{g})$ sein.

Ergebnisse für $g \rightarrow 0$

1. ($\gamma = 0$ und $\alpha < \frac{1}{2}$) oder ($\gamma = \alpha$ und $0 < \alpha < 1$):

$$V = \frac{1}{2(1-\alpha)} \left(1 + \gamma + g(1-\gamma) \frac{1+2\alpha+\gamma}{1-2\alpha+\gamma} \right) + \mathcal{O}(g^2) \quad (4.29)$$

$$U = \frac{2\alpha + \gamma - 3}{2(1-\alpha)(1-\gamma)} + \mathcal{O}(g)$$

$$t_1 = \frac{2\alpha - 1 - \gamma}{2\sqrt{(1+\gamma)(1-\gamma)}} \cdot \frac{1}{\sqrt{g}} + \mathcal{O}(\sqrt{g})$$

$$t_2 = \sqrt{\frac{1-\gamma}{1+\gamma}} \cdot \frac{1}{\sqrt{g}} + \mathcal{O}(\sqrt{g}) .$$

Einsetzen von V und den anderen Variablen in Gl.(4.19) liefert

$$\begin{aligned} f(\alpha, g \rightarrow 0, \gamma, \varepsilon_{max}) = & \\ & - \frac{1-\alpha}{2} \ln(1-\alpha) - \frac{\alpha}{2} \ln \alpha - \frac{1}{2} \ln 2 + \\ & + \frac{1-\alpha}{2} \ln(1-\gamma) + \frac{\alpha}{2} \ln(1+\gamma) - \frac{1}{2} g \ln g + \mathcal{O}(g) . \end{aligned} \quad (4.30)$$

Für $\gamma = \alpha$ gilt dann für $\alpha < 1$

$$t_1 = -\frac{1}{2} \cdot \sqrt{\frac{1-\alpha}{1+\alpha}} \cdot \frac{1}{\sqrt{g}} + \mathcal{O}(\sqrt{g}) \rightarrow -\infty,$$

$$t_2 = \sqrt{\frac{1-\alpha}{1+\alpha}} \cdot \frac{1}{\sqrt{g}} + \mathcal{O}(\sqrt{g}) \rightarrow +\infty,$$

und für $\gamma = 0$ erhält man für $\alpha < \frac{1}{2}$

$$t_1 = \frac{2\alpha - 1}{2} \cdot \frac{1}{\sqrt{g}} + \mathcal{O}(\sqrt{g}) \rightarrow -\infty; \quad t_2 = \frac{1}{\sqrt{g}} + \mathcal{O}(\sqrt{g}) \rightarrow +\infty.$$

Das paßt mit dem obigen Ansatz zusammen.

2. $\gamma = 0$ und $\alpha = \frac{1}{2}$

$$V = 1 + \mathcal{O}(\sqrt{g}); \quad (4.31)$$

$$U = -2 + \mathcal{O}(\sqrt{g})$$

als Ansatz liefert konsistent

$$t_1 = \frac{\mathcal{O}(\sqrt{g})}{2\sqrt{g}} = \mathcal{O}(1),$$

$$t_2 = \frac{2 + \mathcal{O}(\sqrt{g})}{2\sqrt{g}} = \frac{1}{\sqrt{g}} + \mathcal{O}(1) \rightarrow +\infty.$$

Es gilt

$$f\left(\alpha = \frac{1}{2}, g \rightarrow 0, \gamma = 0, \varepsilon_{max}\right) = -g \ln g + \mathcal{O}(\sqrt{g}). \quad (4.32)$$

3. $\gamma = 0$ und $\frac{1}{2} < \alpha < 1$

Die Fehler in U, V und f sind von der Ordnung $\mathcal{O}\left(\exp\left(-\frac{c}{g}\right)\right)$ (c ist eine Konstante), also verschwindend klein.

$$V \rightarrow \frac{\alpha}{1-\alpha}, \quad (4.33)$$

$$\begin{aligned}
U &\rightarrow \frac{2g-1}{1-\alpha} \\
t_1 &= \frac{2\alpha-1}{2\sqrt{\alpha(1-\alpha)}} \cdot \frac{1}{\sqrt{g}} + \mathcal{O}(\sqrt{g}) \rightarrow +\infty, \\
t_2 &= \frac{1}{2\sqrt{\alpha(1-\alpha)}} \cdot \frac{1}{\sqrt{g}} + \mathcal{O}(\sqrt{g}) \rightarrow +\infty. \\
f\left(\alpha > \frac{1}{2}, g \rightarrow 0, \gamma = 0, \varepsilon_{max}\right) &\rightarrow -g \ln g - (1-g) \ln(1-g). \quad (4.34)
\end{aligned}$$

4.3 Das Ergebnis $\langle\langle \mathcal{N}(\alpha, \mathbf{g} = \frac{1}{2}, \gamma, \varepsilon_{max}) \rangle\rangle$

Wir schränken den in Abschnitt 4.2 behandelten Spezialfall weiter ein und setzen $g = \frac{1}{2}$. Die zum 1. Muster gehörende Sattelpunktvariable s ist dann Null. Wie schon zuvor erwähnt, tritt s in einer Rechnung ohne Berücksichtigung des Hamming-Abstandes nicht auf. Dies sieht man auch formal, wenn man die in Abschnitt 4.2.2 aufgeführten Gleichungen für $g = \frac{1}{2}$ löst; $U = 0$ und damit $s = 0$ sind die Ergebnisse der Iteration.

4.3.1 Das Ergebnis $\langle\langle \mathcal{N}(\alpha, \mathbf{g} = \frac{1}{2}, \gamma, \varepsilon_{max}) \rangle\rangle$; Sattelpunktgleichung und Liste aller Hilfsvariablen

Man setzt in Abschnitt 4.2.2 $g = \frac{1}{2}$ und erhält für $N \rightarrow \infty$

$$\langle\langle \mathcal{N}\left(\alpha, g = \frac{1}{2}, \gamma, \varepsilon_{max}\right) \rangle\rangle \propto \exp\left(N f\left(\alpha, g = \frac{1}{2}, \gamma, \varepsilon_{max}\right)\right)$$

mit

$$\begin{aligned}
f\left(\alpha, g = \frac{1}{2}, \gamma, \varepsilon_{max}\right) &= -(1-\alpha) \ln(1-\alpha) - \alpha \ln \alpha + \ln 2 + \frac{\alpha}{2} \ln(m+1) - \frac{1}{2} \ln K + \\
&\quad + \frac{1}{2Z} \left(r^2 - 2r\sqrt{K}(1-\gamma) + (1-\gamma)^2 K - 1 \right) + \ln \Phi(t). \quad (4.35)
\end{aligned}$$

Φ steht wieder für die Φ -Funktion (siehe auch A.3). Die Hilfsvariablen

$$m, K, r \quad \text{und} \quad t$$

hängen wie folgt von der Iterationsvariable Z ab:

$$m = \frac{Z - 1}{1 - Z(1 - \alpha)} \quad (4.36)$$

$$K = \frac{Z + 1 + (1 - \gamma)m}{(1 - 2\gamma) \cdot Z(1 - \alpha) + \gamma} \quad (4.37)$$

$$r = \frac{1}{2\sqrt{K}}(m + (1 - 2\gamma)K) \quad (4.38)$$

$$t = \frac{r}{\sqrt{Z}}. \quad (4.39)$$

Die Variable Z ergibt sich schließlich aus der Iterationsgleichung

$$Z = \frac{1}{1 - \alpha} \cdot \left[\frac{1}{2} + \frac{m}{2K} + \sqrt{\frac{Z}{K}} \cdot \frac{\Phi'(t)}{\Phi(t)} \right], \quad (4.40)$$

die mit der Regula-falsi-Methode gelöst wurde.

4.3.2 Grenzwerte im Fall $\alpha \rightarrow 0$

Wir betrachten wieder die Spezialfälle 1. $\gamma = 0$ und 2. $\gamma = \alpha$.

In beiden Fällen beobachtet man für $\alpha \rightarrow 0$, daß das Argument t der Φ -Funktion gegen Null geht. Man macht einen selbstkonsistenten Taylorreihen-Ansatz für Z und verwendet Gl.(A.4) für $\Phi(t \rightarrow 0)$ (siehe C.7). Durch Koeffizientenvergleich in der Sattelpunktgleichung (4.40) erhält man die Ergebnisse

1. $\gamma = 0$:

$$Z = \frac{2}{\pi} + 1 \cdot \alpha + \mathcal{O}(\alpha^2) \quad (4.41)$$

$$t = \sqrt{\frac{\pi}{2}} \cdot \frac{\alpha}{1 - \frac{2}{\pi}} + \mathcal{O}(\alpha^2) \quad (4.42)$$

$$f(\alpha) = -\frac{\alpha}{2} \ln \alpha + \frac{\alpha}{2} \left(\ln \left(\frac{2}{\pi} \right) + 1 - \ln \left(1 - \frac{2}{\pi} \right) \right) + \mathcal{O}(\alpha^2). \quad (4.43)$$

2. $\gamma = \alpha$:

$$Z = \frac{2}{\pi} + \alpha \left(3 - \frac{2}{\pi} + \frac{1 - \frac{4}{\pi}}{1 - \frac{2}{\pi}} \right) + \mathcal{O}(\alpha^2) \quad (4.44)$$

$$t = \sqrt{\frac{\pi}{2}} \cdot \frac{\alpha}{1 - \frac{2}{\pi}} + \mathcal{O}(\alpha^2) \quad (4.45)$$

$$f(\alpha) = -\frac{\alpha}{2} \ln \alpha + \frac{\alpha}{2} \left(\ln \left(\frac{2}{\pi} \right) - 1 - \ln \left(1 - \frac{2}{\pi} \right) \right) + \mathcal{O}(\alpha^2) . \quad (4.46)$$

In beiden Fällen liefert der Taylorreihen-Ansatz also das richtige Skalierungsverhalten für t .

4.3.3 Grenzwerte im Fall $\alpha \rightarrow 1$

Hier geht t gegen ∞ . Für den Fall $\gamma < \alpha$ kann man einen selbstkonsistenten Skalierungsansatz machen. Im Fall $\gamma = \alpha$ ist dies nicht möglich, und man ist letztendlich auf numerische Rechnungen angewiesen.

Vernachlässigt man $\Phi'(t)$ in Gl.(4.40), so erhält man die Gleichung

$$Z = \frac{1}{1 - \alpha} \left[\frac{1}{2} + \frac{m}{2K} \right]. \quad (4.47)$$

Nach Einsetzen von m und K aus den Gleichungen (4.36) und (4.37) erhält man nach einfachen Umformungen eine Bestimmungsgleichung für Z :

$$\begin{aligned} Z^3 - \frac{1 + \alpha}{1 - \alpha} Z^2 + \frac{\alpha}{(1 - \alpha)^2} Z &= 0 \\ \iff Z = \frac{1}{1 - \alpha} \vee Z = \frac{\alpha}{1 - \alpha} \vee Z = 0 . \end{aligned} \quad (4.48)$$

$Z = 0$ entfällt, da in diesem Fall $r \sim -\gamma$ ist und $t \rightarrow +\infty$ nicht erfüllt ist.

1. Selbstkonsistenter Skalierungsansatz im Fall $\gamma < \alpha$ mit der Zusatzvoraussetzung $\gamma \leq q < 1$, wobei q eine Konstante ist:

$$Z = \frac{\alpha}{1 - \alpha} .$$

Es folgen nach Einsetzen

$$m = \frac{2\alpha - 1}{(1 - \alpha)^2}, \quad K = \frac{1}{(1 - \alpha)^2}, \quad r = \frac{\alpha - \gamma}{1 - \alpha} .$$

Insbesondere ist

$$t = \frac{\alpha - \gamma}{\sqrt{\alpha(1 - \alpha)}} . \quad (4.49)$$

Für $\alpha \rightarrow 1$ ist also $t \rightarrow \infty$ unter der obigen Zusatzvoraussetzung erfüllt. Die Fehler in den Variablen sind somit von der Ordnung des vernachlässigten

$$\Phi'(t) \propto \exp\left(-\frac{1}{2}t^2\right) = \exp\left(-\frac{1}{2} \cdot \frac{(\alpha - \gamma)^2}{\alpha(1 - \alpha)}\right) .$$

Wir setzen die Variablen in Gl.(4.35) für f ein und erhalten

$$f\left(\alpha, g = \frac{1}{2}, \gamma, \varepsilon_{max}\right) = \ln 2 - \mathcal{O}\left(\exp\left(-\frac{1}{2} \cdot \frac{(\alpha - \gamma)^2}{\alpha(1 - \alpha)}\right)\right) \quad (4.50)$$

für $\alpha \rightarrow 1$ und $\gamma < \alpha$. Der Fehler ist also in diesem Fall exponentiell klein und für große α ist $f \sim \ln 2$.

2. Der Fall $\gamma = \alpha$:

Die Lage ist hier nicht so einfach wie im Fall 1, weil der Fehler nicht exponentiell klein ist. Im Grenzfall gilt

$$Z \rightarrow \frac{1}{1 - \alpha} - \delta , \quad (4.51)$$

wobei

$$\delta \rightarrow 0 \quad (\alpha \rightarrow 1) .$$

Nach der Berechnung der Hilfsvariablen erhält man mit Gl.(4.35)

$$f\left(\alpha, g = \frac{1}{2}, \gamma, \varepsilon_{max}\right) = \ln 2 + \mathcal{O}\left(\frac{1 - \alpha}{2} \ln \delta\right) , \quad (4.52)$$

was vermuten läßt, daß f hier viel schwächer gegen $\ln 2$ strebt. δ selbst ist nur durch numerisches Lösen der Sattelpunktgleichung (4.40) zu gewinnen. $f \rightarrow \ln 2$ wurde bestätigt, wie die Ergebnisse in Abschnitt 5.1.1 zeigen werden.

4.4 Das Ergebnis $\langle\langle\mathcal{N}(\alpha, \mathbf{g} = \frac{1}{2}, \gamma, \varepsilon)\rangle\rangle$

Wir lassen ε jetzt frei und betrachten das *allgemeine* Endergebnis (4.18) für $f(\alpha, g, \gamma, \varepsilon, X, Y, Z, a_1, l_1, w)$. Es werden wieder nur Zustände betrachtet, die keine Korrelation zum ersten Muster haben; setzt man also $g = \frac{1}{2}$, so folgt wieder, daß die zum ersten Muster gehörenden Sattelpunktvariablen a_1 und l_1 Null sind.

Um das Problem weiter zu vereinfachen, wendet man die in Abschnitt 4.1.2 geschilderten Methoden 1 und 2 wiederholt an (siehe C.8). Man erhält für $N \rightarrow \infty$ das folgende Endergebnis:

$$\langle\langle\mathcal{N}(\alpha, g = \frac{1}{2}, \gamma, \varepsilon)\rangle\rangle \propto \exp\left(N f\left(\alpha, g = \frac{1}{2}, \gamma, \varepsilon\right)\right). \quad (4.53)$$

Dabei ist

$$\begin{aligned} f\left(\alpha, g = \frac{1}{2}, \gamma, \varepsilon\right) = & \\ & - (1 - \alpha) \ln(1 - \alpha) - \alpha \ln \alpha + \ln 2 - \frac{\alpha}{2} + \frac{\alpha}{2} \ln \frac{\alpha}{2} - \frac{\alpha}{2} \ln \left(\frac{1}{2} - \varepsilon\right) + \\ & + \ln \Phi(t) + t \cdot \sqrt{P}(-2\varepsilon + \gamma) + \frac{t^2}{2} - \frac{\varepsilon}{Z} + \\ & + P\left(\frac{1}{2}(1 - \gamma)^2 - \left(\frac{1}{2} - \varepsilon\right) \cdot (1 - 2\gamma)\right) - \frac{1}{2} \ln P - \frac{1 - \alpha}{2} \ln Z. \end{aligned} \quad (4.54)$$

Es gilt

$$Z = \frac{2\varepsilon}{1 - \alpha}. \quad (4.55)$$

P und t hängen von einer Sattelpunktvariable S ab:

$$P = \frac{\frac{1}{2} - S}{\frac{1}{2}\gamma^2 + \varepsilon(1 - 2\gamma)} \quad (4.56)$$

$$t = \frac{2S}{\sqrt{P}} \cdot \frac{1}{\gamma - 2\varepsilon}. \quad (4.57)$$

S erhält man schließlich aus der Iterationsgleichung

$$S = -\frac{1}{2}(\gamma - 2\varepsilon) \cdot \left(\sqrt{P} \cdot \frac{\Phi'(t)}{\Phi(t)} + P(\gamma - 2\varepsilon)\right). \quad (4.58)$$

Φ steht wieder für die Φ -Funktion (A.3). Die Iterationsgleichung (4.58) wurde wieder mit der Regula falsi-Methode gelöst. Dabei erweist es sich als günstig,

$$S \rightarrow \frac{S}{1 - 2\varepsilon}$$

umzuskalieren. Die Ergebnisse für $\gamma = \alpha$ finden sich in Abschnitt 5.3.

Das Maximum ε_{max} von $f(\alpha, g = \frac{1}{2}, \gamma, \varepsilon)$ bezüglich ε errechnet sich aus

$$\frac{\partial f}{\partial \varepsilon} = 0 .$$

Ersetzt man analog zu Anhang B in Gl.(4.54) ε durch Z , so kann man wegen des linearen Zusammenhangs (4.55) äquivalent

$$\frac{\partial f}{\partial Z} = 0$$

lösen. Mit dem Sattelpunktwert für Z aus Abschnitt 4.3 kann man folglich sofort eine Kurve

$$\varepsilon_{max}(\alpha) = \frac{1}{2}(1 - \alpha)Z \tag{4.59}$$

zeichnen. Es gelingt also, erste wichtige Erkenntnisse über die Energieabhängigkeit der metastabilen Zustände zu gewinnen, ohne Gl.(4.58) lösen zu müssen.

Kapitel 5

Zahl der metastabilen Zustände bei einem Netzwerk mit Projektorkopplungen; Auswertung der Ergebnisse

Dieses Kapitel stellt das Herz der vorliegenden Arbeit dar. Es präsentiert die Lösungen der Sattelpunktgleichungen aus Kapitel 4. Die Ergebnisse werden interpretiert; zudem wird auf noch ungelöste Probleme hingewiesen. Wir nehmen außerdem den in Kapitel 2 begonnenen Gedankengang wieder auf und stellen den Zusammenhang der metastabilen Zustände mit der $\tau = 0$ -Glauber-Einzelspinflipdynamik dar.

Aus Kapitel 4 ergab sich für den direkten Mittelwert der Zahl der metastabilen Zustände beim Netzwerk mit Projektorkopplungen die Form

$$\langle\langle \mathcal{N}(\alpha, g, \gamma, \varepsilon) \rangle\rangle \propto \exp(N f(\alpha, g, \gamma, \varepsilon)) .$$

Dabei ist der Exponent $f(\alpha, g, \gamma, \varepsilon)$ nach der Jensenschen Ungleichung (siehe Abschnitt 2.3.3) eine obere Schranke für den typischen Wert

$$\frac{1}{N} \langle\langle \ln \mathcal{N}(\alpha, g, \gamma, \varepsilon) \rangle\rangle ;$$

die meisten Abbildungen in diesem Kapitel stellen deshalb f bei vorgegebenen Parametern α, g, γ und ε dar.

α braucht nur im Intervall $(0,1)$ betrachtet zu werden, weil die Projektormatrix

$$J_{ij} = \frac{1}{N} \sum_{\mu,\nu} \xi_i^\mu (C^{-1})_{\mu\nu} \xi_j^\nu$$

nur in diesem Bereich existiert.

Die Stabilität γ wird vorrangig für die beiden Werte $\gamma = 0$ und $\gamma = \alpha$ betrachtet. Im ersten Fall liefert die Rechnung die Zahl der metastabilen Zustände beim Projektornetzwerk mit Selbstkopplungen. Im Fall $\gamma = \alpha$ schränken wir diese Zahl auf die Zustände ein, deren lokale Energien λ_i zusätzlich die Bedingung

$$\lambda_i = S_i \sum_{j=1}^N J_{ij} S_j > \gamma = \alpha \quad \forall i \quad (5.1)$$

$$\iff \lambda'_i = S_i \sum_{j=1}^N (J_{ij} - \delta_{ij} \alpha) S_j > 0 \quad \forall i \quad (5.2)$$

erfüllen.

Wie in Kapitel 4 bereits erwähnt, wollen wir den Fall $\gamma = \alpha$ mit dem Netzwerk ohne Selbstkopplungen J_{ii} identifizieren. Dies stützt sich hauptsächlich auf die Erkenntnis

$$J_{ii} \rightarrow \alpha + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right) \quad (N \rightarrow \infty).$$

Bei einer exakten Einbeziehung von J_{ii} in Gl.(5.2) (ersetze α durch J_{ii}) stößt man auf Schwierigkeiten, und eine Rechnung analog zu Kapitel 4 ist nicht möglich. Es wird sich jedoch zeigen, daß das Modell $\gamma = \alpha$ zwei wichtige Eigenschaften des Modells ohne Selbstkopplungen aufweist. Zum einen hat es nach Abschnitt 5.2.2 auch ein kritisches α von $\alpha_c = 1$, zum anderen liefert es in Abschnitt 5.3 die richtigen unteren und oberen Schranken für eine Simulation der Endenergien eines seriellen $\tau = 0$ -Glauber-Prozesses für unkorrelierte Zustände.

Der Hamming-Abstand g wird wie beim Hopfield-Modell nur im Intervall $(0, \frac{1}{2}]$ betrachtet, weil f bezüglich $g = \frac{1}{2}$ symmetrisch ist:

$$f\left(g = \frac{1}{2} - x\right) = f\left(g = \frac{1}{2} + x\right).$$

Dies ist anschaulich klar, weil mit \underline{S} auch $-\underline{S}$ metastabil ist. Es folgt aber auch formal aus der Lösung der Sattelpunktgleichungen.

Um die einleitenden Bemerkungen abzuschließen, sei noch daran erinnert, daß ε für die vorgegebene negative Energie pro Spin steht:

$$\varepsilon = \frac{1}{2N} \sum_{i=1}^N \sum_{j=1}^N J_{ij} S_i S_j .$$

Wir stellen jetzt die Lösung der Sattelpunktgleichungen aus Kapitel 4 in den einzelnen Parameterbereichen vor. Die in Kapitel 4 aufgeführten Näherungen für $g \rightarrow 0$, $\alpha \rightarrow 0$ und $\alpha \rightarrow \infty$ wurden jeweils sorgfältig numerisch überprüft. Ebenso wurde stets festgestellt, daß die Sattelpunktvariable

$$Z = \sum_{\mu=2}^p a_{\mu}^2 \quad (5.3)$$

(siehe Gl.(4.11) mit umskalierten a_{μ}) von der Ordnung $\mathcal{O}(1)$ ist. Dies stimmt mit der in Abschnitt 4.1.1 gemachten Voraussetzung

$$a_{\mu} = \mathcal{O}\left(\frac{1}{\sqrt{N}}\right) \quad \forall \mu \geq 2 \quad (5.4)$$

für die Entwicklungskoeffizienten in

$$\underline{S} = \sum_{\mu=1}^p a_{\mu} \underline{\xi}^{\mu} + \delta \underline{S} \quad (5.5)$$

überein.

5.1 Ergebnisse für $\langle\langle \mathcal{N}(\alpha, \mathbf{g} = \frac{1}{2}, \gamma, \varepsilon_{\max}) \rangle\rangle$

$g = \frac{1}{2}$ ist der Sattelpunktwert von $f(\alpha, g, \gamma, \varepsilon_{\max})$ bezüglich g :

$$\begin{aligned} \int \langle\langle \mathcal{N}(\alpha, g, \gamma, \varepsilon_{\max}) \rangle\rangle dg &\propto \\ &\int \exp(N f(\alpha, g, \gamma, \varepsilon_{\max})) dg \\ &\propto \exp\left(N f\left(\alpha, g = \frac{1}{2}, \gamma, \varepsilon_{\max}\right)\right) . \end{aligned}$$

Wir erhalten für $g = \frac{1}{2}$ also die Gesamtzahl der metastabilen Zustände. Die in Abschnitt 4.3.1 aufgeführte Sattelpunktgleichung wurde für verschiedene γ -Werte gelöst.

5.1.1 Das Ergebnis für $f(\alpha, g = \frac{1}{2}, \gamma, \varepsilon_{max})$ in den Fällen $\gamma = 0$ und $\gamma = \alpha$

In Abb. 5.1 sieht man, daß in beiden Fällen das vorhergesagte Grenzverhalten

$$f \rightarrow 0 \quad (\alpha \rightarrow 0) \quad \text{und} \quad f \rightarrow \ln 2 \quad (\alpha \rightarrow 1)$$

auftritt (siehe Abschnitt 4.3). Wie erwartet ist f kleiner, wenn die Stabilität γ größer wird:

$$f(\gamma = \alpha) < f(\gamma = 0) .$$

Die Kurve für $f(\gamma = \alpha)$ ist in etwa mit der Kurve für die remanente Magnetisierung [HeOp90] beim Projektornetzwerk ohne Selbstkopplungen vergleichbar (siehe Abb. 5.2). Sie wurde ermittelt, indem man die serielle $\tau = 0$ -Glauber-Dynamik für einen zufällig ausgewürfelten Anfangszustand \underline{S}^a bis in einen (metastabilen) Endzustand \underline{S}^e laufen läßt, die remanente Magnetisierung

$$m_r = \frac{1}{N} \underline{S}^a \cdot \underline{S}^e$$

bildet und über viele Läufe mittelt. m_r ist vermutlich um so größer, je mehr Fallen für die Dynamik existieren, je größer also f ist. Eine Simulation für $\alpha \sim 1$ ist äußerst schwierig und wurde nicht durchgeführt; falls

$$f(\gamma = \alpha) \rightarrow \ln 2 \quad (\alpha \rightarrow 1)$$

richtig ist, ist wohl

$$m_r \rightarrow 1 \quad (\alpha \rightarrow 1)$$

zu erwarten.

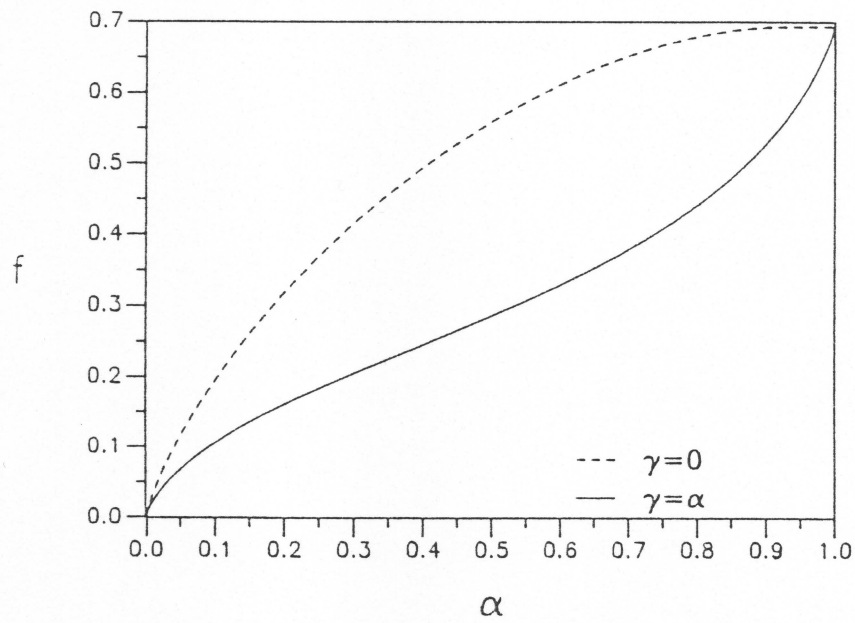


Abbildung 5.1: $f(\alpha, g = \frac{1}{2}, \gamma, \varepsilon_{max})$ für $\gamma = 0$ und $\gamma = \alpha$.

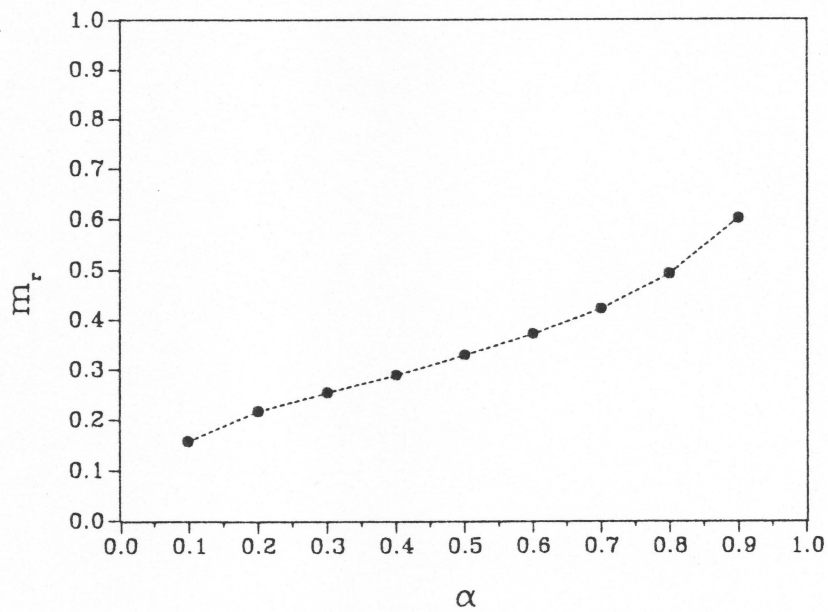


Abbildung 5.2: Die remanente Magnetisierung m_r beim Projektornetzwerk (ohne J_{ii}) für die serielle Dynamik.

Kehren wir nun zur Diskussion von Abb. 5.1 zurück. $f \rightarrow 0$ ($\alpha \rightarrow 0$) ist einleuchtend, denn man erwartet immer weniger metastabile Zustände, wenn es auch weniger Grundzustände gibt.

Um uns $f \rightarrow \ln 2$ ($\alpha \rightarrow 1$) im Fall $\gamma < \alpha$ plausibel zu machen, betrachten wir

$$-\frac{2E}{N} = 2\varepsilon = \frac{1}{N} \sum_{i=1}^N \lambda_i ,$$

wobei E die Gesamtenergie des Netzwerks ist und λ_i die lokale Energie

$$\lambda_i = S_i \sum_{j=1}^N J_{ij} S_j$$

darstellt. Wegen der Projekteigenschaft von (J_{ij}) gilt

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \lambda_i^2 &= \frac{1}{N} \sum_{i=1}^N S_i^2 \sum_{j,k} J_{ij} J_{ik} S_j S_k = \frac{1}{N} \sum_{j,k} \left(\sum_{i=1}^N J_{ji} J_{ik} \right) S_j S_k \\ &= \frac{1}{N} \sum_{j,k} (\mathbf{J}^2)_{jk} S_j S_k = \frac{1}{N} \sum_{i=1}^N \lambda_i = 2\varepsilon . \end{aligned} \quad (5.6)$$

$2\varepsilon = \lambda$ ist also der Mittelwert der lokalen Energien λ_i für *einen* Zustand \underline{S} . Mit Gl.(5.6) ergibt sich die Streuung der λ_i zu

$$\Delta\lambda_i(\underline{S}) = \sqrt{\frac{1}{N} \sum_{i=1}^N \lambda_i^2 - \left(\frac{1}{N} \sum_{i=1}^N \lambda_i \right)^2} = \sqrt{2\varepsilon - (2\varepsilon)^2} . \quad (5.7)$$

Für zufällig verteilte („gewöhnliche“) Zustände \underline{S} ist λ selbstmittelnd:

$$\begin{aligned} \lambda &= \frac{1}{N} \sum_{i=1}^N J_{ii} + \frac{1}{N} \sum_{i,j(\neq i)} J_{ij} S_i S_j \\ &= \alpha + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right) . \end{aligned} \quad (5.8)$$

Daraus folgt

$$\Delta\lambda_i \sim \sqrt{\alpha(1-\alpha)} . \quad (5.9)$$

Für einen *metastabilen* Zustand \underline{S}^m macht nach Abschnitt 4.4 die Sattelpunktvariable Z eine Aussage über λ :

$$2\varepsilon = Z(1 - \alpha) .$$

Mit dem in Abschnitt 4.3.3 hergeleiteten Grenzverhalten für Z erhält man mit Gl.(5.7)

1. $\gamma < \alpha$ mit der Zusatzvoraussetzung $\gamma \leq q < 1$, wobei q eine Konstante ist:

$$\begin{aligned} 2\varepsilon &= \lambda = \frac{\alpha}{1 - \alpha} \cdot (1 - \alpha) + \mathcal{O}\left(\exp\left(-\frac{(\alpha - \gamma)^2}{2\alpha(1 - \alpha)}\right)\right) \\ &= \alpha + \mathcal{O}\left(\exp\left(-\frac{(\alpha - \gamma)^2}{2\alpha(1 - \alpha)}\right)\right) \end{aligned} \quad (5.10)$$

$$\Delta\lambda_i(\underline{S}^m) \sim \sqrt{\alpha(1 - \alpha)} . \quad (5.11)$$

2. $\gamma = \alpha$:

$$2\varepsilon = \lambda \rightarrow \frac{1 - \alpha}{1 - \alpha} - (1 - \alpha)\delta = 1 - (1 - \alpha)\delta . \quad (5.12)$$

$\Delta\lambda_i$ ist kleiner als oben, es ist aber keine analytische Aussage möglich. δ ist der Fehler in Gl.(4.51).

Nach Gl.(5.8) ist für ein *gewöhnliches* \underline{S} bereits $\lambda = \alpha$. Gleichzeitig sinkt für $\alpha \rightarrow 1$ die Streuung $\Delta\lambda_i$. Es ist also zu erwarten, daß die λ_i sich alle in der Umgebung von α aufhalten und mit steigendem α immer weniger \underline{S} angetroffen werden, bei denen es *einen* Platz i_0 mit $\lambda_{i_0} \leq 0$ gibt.

In der Tat gilt für einen *metastabilen* Zustand im Fall $\gamma < \alpha$: $\lambda \rightarrow \alpha(+)$, die λ_i sind folglich hier etwas rechts von α angesiedelt, und mit steigendem α ist der Mittelwert λ in unmittelbarer Nähe des Mittelwertes für einen gewöhnlichen Zustand. $f \rightarrow \ln 2$ ($\alpha \rightarrow 1$) wird so im Fall $\gamma < \alpha$ anschaulich klar.

Im Fall $\gamma = \alpha$ gilt jedoch ebenfalls $f \rightarrow \ln 2$ ($\alpha \rightarrow 1$). Die metastabilen Zustände erfüllen hier

$$\lambda_i > \gamma = \alpha \quad \forall i .$$

Gl.(5.12) zeigt, daß die dafür notwendige Bedingung

$$\lambda > \gamma = \alpha \quad (5.13)$$

auch im Fall $\alpha \rightarrow 1$ erfüllt ist.

Eine Argumentation nach obigem Muster ist hier jedoch nicht möglich, weil die λ der *gewöhnlichen* Zustände sich gerade an der durch die notwendige Bedingung (5.13) bestimmten Grenze $\lambda = \gamma = \alpha$ aufhalten. Anschaulichen Begründungen des Grenzverhaltens sind hier Grenzen gesetzt; die obigen Betrachtungen für $\gamma < \alpha$ haben sowieso den Nachteil, daß sie hauptsächlich mit dem mittleren Feld

$$\lambda(\underline{S}) = \frac{1}{N} \sum_{i=1}^N \lambda_i$$

arbeiten. Um Aussagen über metastabile Zustände machen zu können, kommt man aber nicht daran vorbei, *alle* einzelnen λ_i zu betrachten. Die λ_i wiederum sind durch die p Muster ξ_i^μ miteinander verknüpft. Dieser Tatsache wird letztendlich nur durch die Kalkulation in Kapitel 4 Rechnung getragen.

Bevor wir diesen Unterpunkt abschließen, weisen wir noch auf zwei verlockende Fehlinterpretationen für $f \rightarrow \ln 2$ im Fall $\gamma < \alpha$ hin:

1. $f \rightarrow \ln 2$ erinnert natürlich daran, daß $p = 1 \cdot N$ linear unabhängige Muster den ganzen Phasenraum aufspannen und dann jeder Zustand \underline{S} zum Muster wird. Gleichwohl ist zu beachten, daß der Fall $\alpha = 1$ nicht durch die Rechnung erfaßt wird, weil N Zufallsmuster nicht als linear unabhängig angenommen werden können, die Projektormatrix also nicht existiert.

Bei $\alpha < 1$ kommt der Hauptbeitrag zu f von den unkorrelierten metastabilen Zuständen, die Korrelationen der Ordnung $\mathcal{O}\left(\frac{1}{\sqrt{N}}\right)$ mit allen Mustern haben *und* Anteile im Orthogonalraum der Muster besitzen. Dies liegt zum einen an der Skalierungsvoraussetzung

$$a_\mu = \mathcal{O}\left(\frac{1}{\sqrt{N}}\right) \quad \forall \mu$$

für die Entwicklungskoeffizienten, die in Abschnitt 4.1.1 getroffen wurde. Zum anderen gilt nach Gl.(4.5) und Gl.(3.2)

$$\begin{aligned} \underline{S} &= \sum_{\nu=1}^p a_\nu \underline{\xi}^\nu + \underline{\delta S} \\ \implies \underline{S} \cdot \underline{S} &= \sum_{\mu,\nu} a_\mu a_\nu \underline{\xi}^\mu \cdot \underline{\xi}^\nu + \underline{\delta S} \cdot \underline{\delta S} \\ \implies \frac{1}{N} \underline{\delta S} \cdot \underline{\delta S} &= 1 - \sum_{\mu=1}^p a_\mu m_\mu = 1 - \frac{1}{N} \sum_{i=1}^N \lambda_i = 1 - \lambda . \end{aligned}$$

Da nach Gl.(5.10) und Gl.(5.12) stets $\lambda < 1$ für die betrachteten metastabilen Zustände gilt, haben sie Anteile im Orthogonalraum. Der Hauptbeitrag zu f kommt also *nicht* von Zuständen \underline{S} , die *reine* Linearkombinationen der Muster sind. Eine Interpretation der Art „immer mehr \underline{S} liegen vollständig im Raum der Muster, wenn $\alpha \rightarrow 1$ “ mag zwar stimmen, spielt aber für $\alpha < 1$ quantitativ keine Rolle.

2. Im Fall $\gamma = 0$ waren für $\alpha \geq \frac{1}{2}$ mit einem Muster auch Zustände stabil, die sich an einer Stelle i vom Muster unterscheiden (siehe Abschnitt 3.2). Es ist jedoch nicht möglich, sich von diesen Zuständen aus weiter durch den ganzen Phasenraum zu bewegen, indem man immer einzelne Spins umdreht. Es mag gelingen, auf diese Weise eine Reihe von metastabilen Zuständen zu erzeugen; ihre Zahl wird jedoch im Vergleich zur Zahl der unkorrelierten metastabilen Zustände keine Rolle spielen.

Abschließend wird noch auf Gl.(5.6) hingewiesen:

Setzen wir einen metastabilen Zustand \underline{S} mit $\lambda_i > 0 \quad \forall i$ voraus und betrachten die exakte Gleichung

$$\sum_{i=1}^N \lambda_i (\lambda_i - 1) = 0, \quad (5.14)$$

so folgt

1. $\lambda_i = 1 \quad \forall i$, d.h. \underline{S} ist vollständig im Raum der Muster
oder
2. Ist ein $\lambda_{i_0} < 1$, so muß es zumindest einen Bruder mit $\lambda_{j_0} > 1$ haben. Gibt es also Orthogonalanteile in \underline{S} , so sind viele $\lambda_i > 1$ (*größer* als lokale Energien von Mustern), um ihre schlechten Brüder auszugleichen. Außerdem folgt noch aus Gl.(5.6), daß für *jeden* Zustand

$$E = -\frac{1}{2} \sum_{i=1}^N \lambda_i = -\frac{1}{2} \sum_{i=1}^N \lambda_i^2 \leq 0$$

gilt.

5.1.2 Der Effekt der Stabilität γ

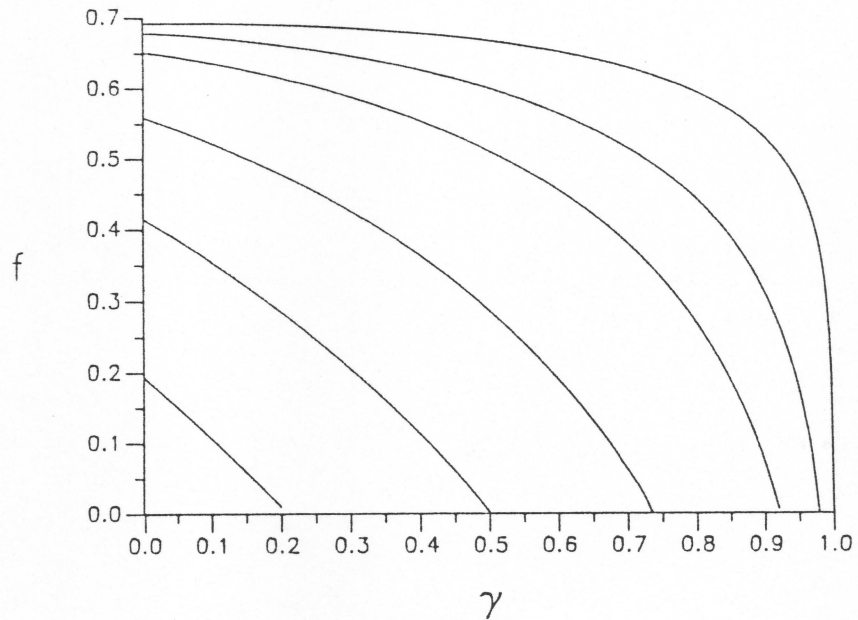


Abbildung 5.3: $f(\alpha, g = \frac{1}{2}, \gamma, \varepsilon_{max})$ in Abhängigkeit von γ für die Werte $\alpha = 0.1; 0.3; 0.5; 0.7; 0.8$ und 0.9 in aufsteigender Reihenfolge.

In Abb. 5.3 sind Kurven

$$\left(\gamma; f\left(\alpha, \gamma, g = \frac{1}{2}, \varepsilon_{max}\right) \right)$$

für verschiedene α -Werte gezeichnet. Wir betrachten hier natürlich nur $\gamma \geq 0$; es ist klar, daß

$$f \rightarrow \ln 2 \quad (\gamma \rightarrow -\infty).$$

Bei festem α ist f wie erwartet streng monoton fallend mit steigender Stabilität der metastabilen Zustände. Für jedes $\alpha < 1$ hat f eine Nullstelle $\gamma_\alpha < 1$. Oberhalb von γ_α ist $f < 0$ und der Mittelwert $\langle\langle \mathcal{N} \rangle\rangle$ ist dann beliebig klein. Nach der Jensenschen Ungleichung folgt weiterhin

$$\langle\langle h \rangle\rangle = \frac{1}{N} \langle\langle \ln \mathcal{N} \rangle\rangle \leq f < 0.$$

Da $N \langle\langle h \rangle\rangle$ einen typischen Wert für den Exponenten Nh von \mathcal{N} für einen festen Satz von Zufallsmustern darstellt, existieren in diesem Fall also keine

metastabilen Zustände. Wir müssen also wieder wie in Abschnitt 2.3.4 die Selbstmittelungseigenschaft *des Exponenten* von \mathcal{N} voraussetzen.

Auf der anderen Seite weiß man, daß zumindest p Muster mit Stabilität $\gamma = 1$ existieren und folglich im gesamten γ -Bereich $[0,1)$ als metastabile Zustände auftreten.

Dieser scheinbare Widerspruch löst sich auf, wenn man sich an die Voraussetzungen für die Rechnung erinnert. In die Mittelung in Abschnitt 4.1.1 gingen nur Zustände ein, die unkorreliert zu (fast) allen Mustern waren. Die in Abschnitt 4.1.1 gemachte Aussage, daß der Mittelwert nur Angaben über solche unkorrelierten Zustände macht, wird hier eindrucksvoll bestätigt. *Sobald* exponentiell viele metastabile Zustände auftreten, erwarten wir hingegen, daß die unkorrelierten Zustände (mit $g = \frac{1}{2}$) am häufigsten sind. Oberhalb von γ_α dürfen wir deshalb annehmen, daß kein exponentielles Verhalten mehr vorliegt und \mathcal{N} etwa polynomial in N geht.

Durch diese vorsichtige Formulierung lassen wir die Möglichkeit offen, daß noch weitere vereinzelte metastabile Zustände \underline{S} im γ -Bereich $(\gamma_\alpha, 1)$ existieren, die keine unkorrelierten Zustände sind und somit aus der Summe über die $\{S_i\}$, wie in Abschnitt 4.1.1 beschrieben, herausfallen.

Schließlich vergegenwärtigt man sich, daß die Kurven mit steigendem α kastenförmig werden. Dies liegt nach Abschnitt 5.1.1 am Grenzwert

$$f\left(\alpha, g = \frac{1}{2}, \gamma = \alpha, \varepsilon_{max}\right) \rightarrow \ln 2 \quad (\alpha \rightarrow 1) \quad :$$

Für $\alpha \rightarrow 1$ bildet sich die Ecke

$$(\gamma = \alpha; \ln 2)$$

heraus, und im Grenzfall $\alpha = 1$ liegt der Kasten

$$f(\gamma) = \begin{cases} \ln 2 & \text{für } \gamma < 1 \\ -\infty & \text{für } \gamma = 1 \end{cases}$$

vor.

Obwohl wie bereits erwähnt unsere Rechnung den Fall $\alpha = 1$ nicht erfaßt, habe ich hier angenommen, es lägen N linear unabhängige Muster vor. Für $\gamma < 1$ ist dann natürlich jeder Zustand metastabil, und für $\gamma = 1$ sind sogar die Muster instabil, so daß $f(\gamma = 1) = -\infty$.

5.2 Ergebnisse für $\langle\langle\mathcal{N}(\alpha, \mathbf{g}, \gamma, \varepsilon_{\max})\rangle\rangle$

In diesem Punkt wird f in Abhängigkeit vom Hamming-Abstand g in den Fällen $\gamma = 0$ und $\gamma = \alpha$ untersucht. Wir vergleichen auch mit den Kurven, die sich ergeben würden, wenn die metastabilen Zustände bezüglich g im Phasenraum genauso verteilt wären wie gewöhnliche Zustände:

Die Wahrscheinlichkeit, im Phasenraum einen (gewöhnlichen) Zustand mit Hamming-Abstand g zum ersten Muster zu finden, ist

$$P(g) = \frac{\binom{N}{Ng}}{2^N} \sim \exp [N (-g \ln g - (1-g) \ln (1-g) - \ln 2)] .$$

Bei α gibt es insgesamt nach Abschnitt 5.1 etwa

$$\langle\langle\mathcal{N}\rangle\rangle \propto \exp \left[N f \left(\alpha, g = \frac{1}{2}, \gamma, \varepsilon_{\max} \right) \right]$$

viele metastabile Zustände. Unter der obigen Annahme ist also die Zahl der metastabilen Zustände mit festem g gegeben durch

$$\langle\langle\mathcal{N}\rangle\rangle \cdot P(g) .$$

$$\begin{aligned} \implies f_{\text{Phasenraum}}(\alpha, g, \gamma, \varepsilon_{\max}) = \\ f \left(\alpha, g = \frac{1}{2}, \gamma, \varepsilon_{\max} \right) - g \ln g - (1-g) \ln (1-g) - \ln 2 . \end{aligned} \quad (5.15)$$

Die Gleichung liefert einen ersten Anhaltspunkt für das gesuchte $f(g)$. Es zeigt sich im folgenden, daß wie beim Hopfield-Modell

$$f_{\text{ph}}(g) < f(g)$$

gilt [Ga86]; in der Nähe der Muster findet man mehr metastabile Zustände als erwartet. Die Kurven $(g; f(g))$ weisen jedoch eine ebenso einfache Struktur auf wie die Kurven $(g; f_{\text{ph}}(g))$, sie sind streng monoton steigend mit g , $g \in (0, \frac{1}{2}]$.

5.2.1 Ergebnisse für $\gamma = 0$ und $\gamma = \alpha$.

In Abb. 5.4 sind Kurven $(g; f(g))$ für verschiedene α -Werte im Fall $\gamma = 0$ aufgetragen.

Abb. 5.5 dient zum Vergleich von f mit f_{ph} im Fall $\gamma = 0$. Für $\alpha < \frac{1}{2}$ hat f eine Nullstelle bezüglich g . Für kleine g ist $f < 0$ und der Mittelwert $\langle\langle \mathcal{N} \rangle\rangle$ wird beliebig klein. Analog zur Überlegung in Abschnitt 5.1.2 schließen wir wieder, daß dann keine metastabilen Zustände existieren können, die Hamming-Abstand g zum ersten Muster haben, aber keine Korrelation mit den anderen Mustern besitzen. Wir schließen trotzdem nicht aus, daß polynomial in N viele metastabile Zustände mit g vorkommen können, die aus der Mittelung in Abschnitt 4.1.1 herausgefallen sind.

Für $\alpha \geq \frac{1}{2}$ ist stets $f > 0$, in unmittelbarer Nähe der Muster liegen also stets exponentiell viele metastabile Zustände. Weitere Betrachtungen zum Grenzwert $g \rightarrow 0$ finden sich in Abschnitt 5.2.2.

Die Abbildungen 5.6 bzw. 5.7 bringen die zu den Abbildungen 5.4 bzw. 5.5 vergleichbaren Kurven im Fall $\gamma = \alpha$. Hier liegen die Nullstellen von f weiter rechts als im Fall $\gamma = 0$, sie treten für *alle* $\alpha < 1$ auf.

In den Abbildungen 5.8 und 5.9 werden die beiden Fälle $\gamma = 0$ und $\gamma = \alpha$ verglichen. Dabei wurde in Abb. 5.9 eine doppelt logarithmische Darstellung gewählt, um die Grenzwerte $g \rightarrow 0$ gut sehen zu können.

Abb. 5.10 schließlich stellt die Nullstellen

$$R_0 = 2g_0$$

von $f(\alpha, g, \gamma, \varepsilon_{\text{max}})$ für verschiedene α dar. Ist das System in einem Zustand \underline{S} mit $g < g_0$, also unterhalb der Kurve $(\alpha; R_0(\alpha))$, so erwarten wir, daß der dynamische $\tau = 0$ -Einzelspinflip-Prozeß ins erste Muster läuft, weil sich in der Nähe von \underline{S} so gut wie keine metastabilen Zustände mehr als Fallen für die Dynamik befinden. Wie schon oft erwähnt, machen wir keine Aussage über die parallele Dynamik, da dort die Zweierzyklen als Fallen für die Dynamik überwiegen. Die Wahl der Ordinate R_0 geschah, um einen Vergleich mit den von Kanter und Sompolinsky für die $\tau = 0$ -Einzelspinflip-Dynamik simulierten Attraktionsradien

$$R = 1 - m = 1 - (1 - 2g)$$

anstellen zu können [KaSo87]. Eine derartige Simulation für Paralleldynamik und weitere Literaturangaben findet man auch in [KrKo90].

Ich wiederhole die Definition von R aus [KaSo87]:

„Der Attraktionsradius R eines Musters ist definiert ... als der größte [doppelte] Hamming–Abstand, in dessen Bereich *fast* alle Zustände (aber nicht notwendigerweise alle) dem Muster zufließen. Genauer ausgedrückt: Ein Zustand, der zufällig unter der Zwangsbedingung, daß er einen Überlapp m mit dem Muster hat, ausgesucht wird, wird mit der Wahrscheinlichkeit $P(m)$ ins Muster laufen, wobei

$$\lim_{N \rightarrow \infty} P(m, N) = \begin{cases} 1 & , \quad m > 1 - R \\ < 1 & , \quad m < 1 - R \end{cases}$$

ist.“

Die simulierten R liegen weit über den theoretischen R_0 aus Abb. 5.10. Die metastabilen Zustände machen offenbar hier keine Aussage über die Einzelspinflip–Dynamik und liefern nur eine schlechte Abschätzung für R . Diese Abschätzung ist außerdem nicht rigoros, da bisher kein Beweis dafür gefunden wurde, daß g während der Dynamik (analog zur Lyapunov–Funktion E) monoton abnimmt. Es kann durchaus möglich sein, daß das System entgegen unserer obigen Erwartung zunächst größere g annimmt, obwohl sein Startzustand in der Nähe des Musters (vielleicht sogar bei $g_{\text{Start}} < g_0$) lag.

Die Diskrepanz zwischen R und R_0 stellt ein bisher noch ungelöstes *Problem* dar. Wir überlegen uns dazu folgendes:

1. Bei der seriellen Dynamik hängt die Kurve $(\alpha; R(\alpha))$ stark von der Reihenfolge der Spinflips ab. Kanter und Sompolinsky flippen zuerst die ausgewürfelten Spins

$$\{S_i, \quad mN < i \leq N\}$$

des Anfangszustandes und danach erst die zum ersten Muster parallelen

$$\{S_i, \quad 1 \leq i \leq mN\} .$$

Das dürfte R im Vergleich zu einer rein seriellen Dynamik erhöhen. Trotzdem erwarte ich, daß auch im letzten Fall die Diskrepanz zwischen R und R_0 bestehen bleibt.

2. Der Effekt ist wohl hauptsächlich darauf zurückzuführen, daß die Dynamik an den metastabilen Zuständen vorbeiläuft. Die Zahl der metastabilen Zustände ist ja im Vergleich zu den $\binom{N}{Ng}$ beim Hamming–Abstand

g existierenden gewöhnlichen Zuständen verschwindend gering. Offenbar gibt es auch genügend gewöhnliche Zustände *mit niedrigen Energien*, die dem dynamischen Prozeß einen Steg ins Muster bauen können.

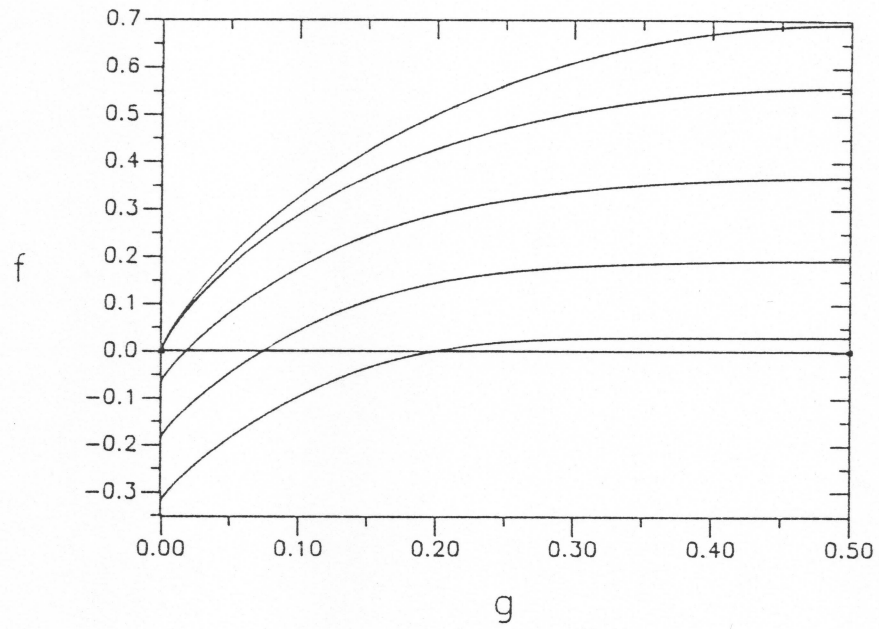


Abbildung 5.4: f in Abhängigkeit von g bei vorgegebenem $\gamma = 0$ für $\alpha = 0.01; 0.1; 0.25; 0.5$ und 0.9 in aufsteigender Reihenfolge.

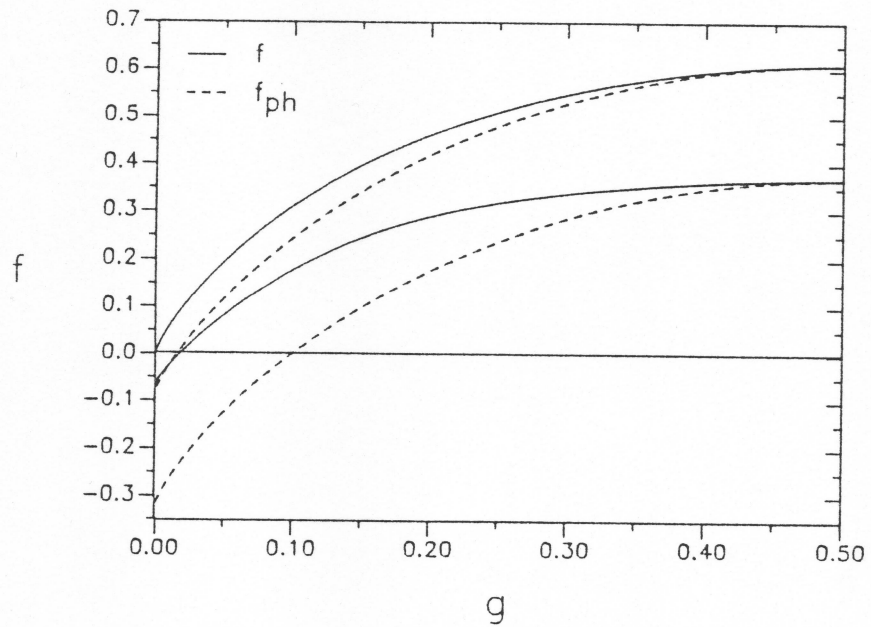


Abbildung 5.5: Vergleich von f mit f_{ph} für $\gamma = 0$ in den Fällen $\alpha = 0.25$ und $\alpha = 0.6$.

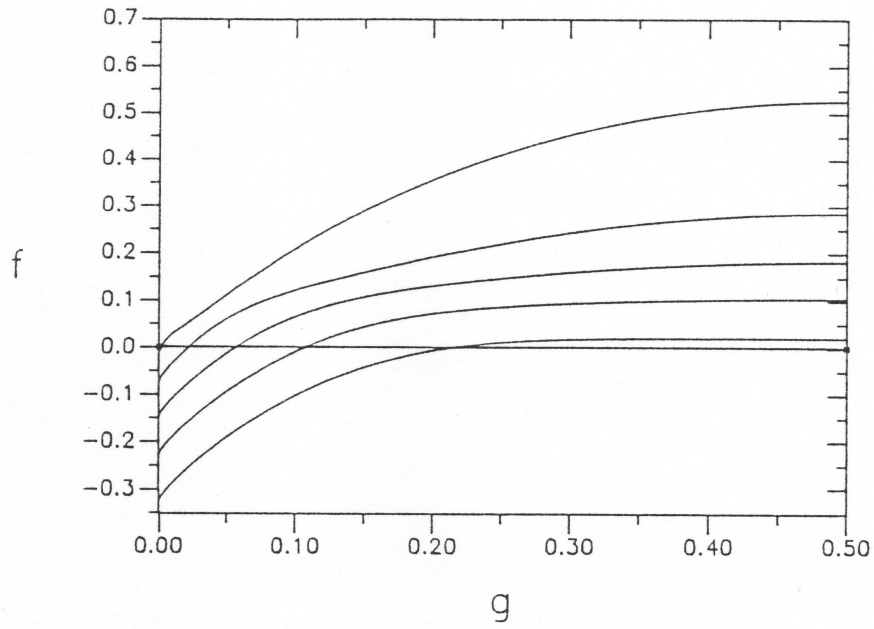


Abbildung 5.6: f in Abhängigkeit von g bei vorgegebenem $\gamma = \alpha$ für $\alpha = 0.01; 0.1; 0.25; 0.5$ und 0.9 in aufsteigender Reihenfolge.

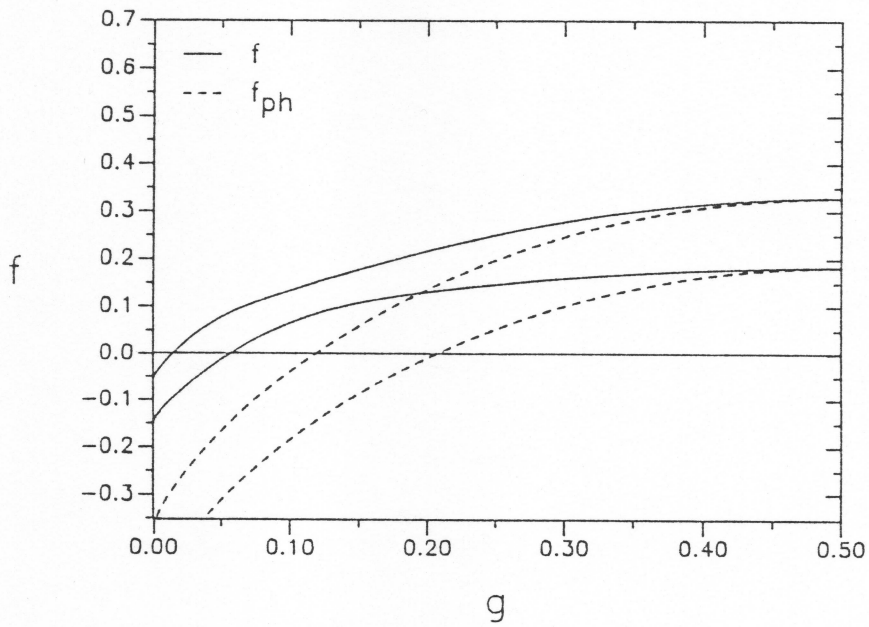


Abbildung 5.7: Vergleich von f mit f_{ph} für $\gamma = \alpha$ in den Fällen $\alpha = 0.25$ und $\alpha = 0.6$.

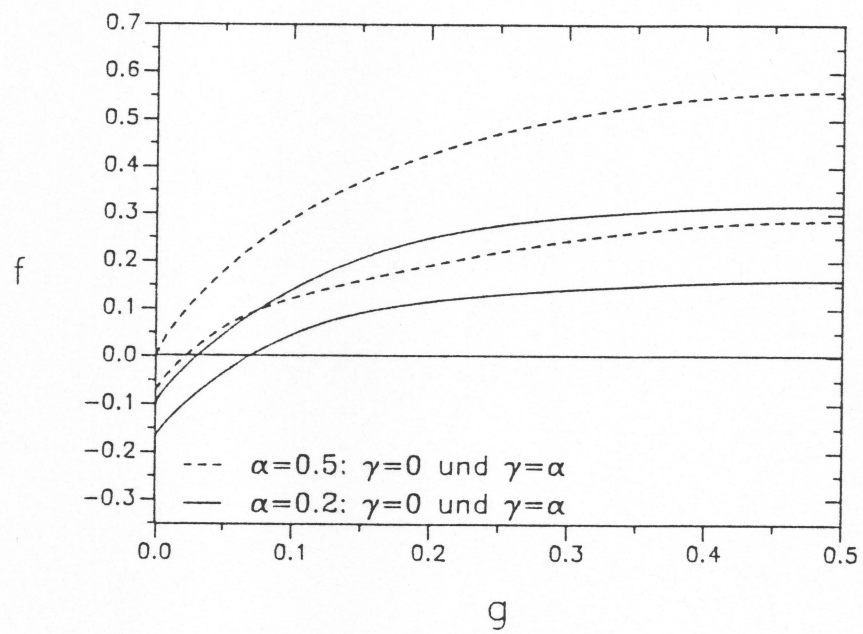


Abbildung 5.8: Ein Vergleich von $f(\gamma = 0)$ mit $f(\gamma = \alpha)$ in den Fällen $\alpha = 0.2$ und $\alpha = 0.5$.

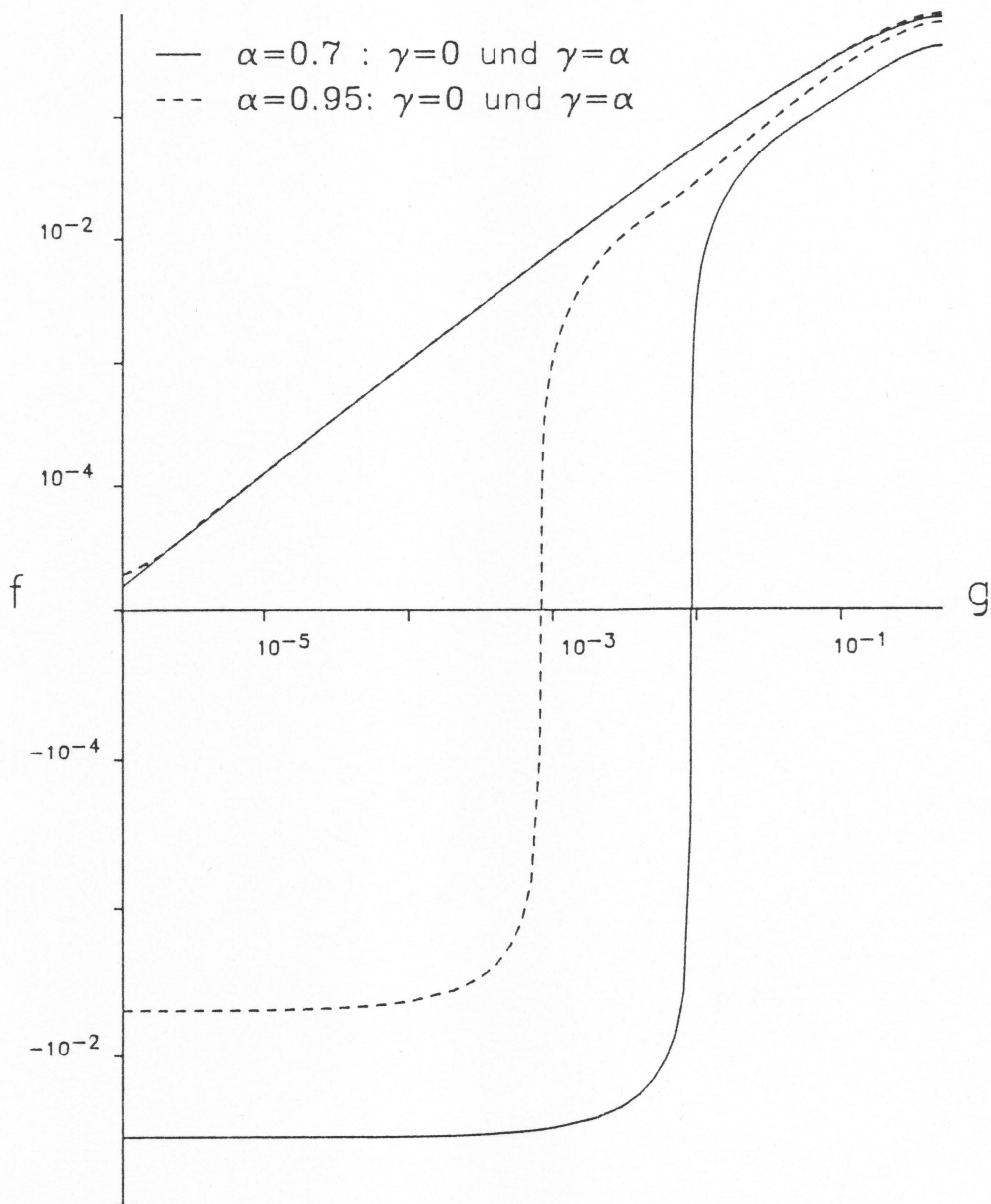


Abbildung 5.9: Ein Vergleich von $f(\gamma=0)$ mit $f(\gamma=\alpha)$ in den Fällen $\alpha=0.7$ und $\alpha=0.95$. Die beiden Kurven für $\gamma=0$ liegen fast aufeinander.

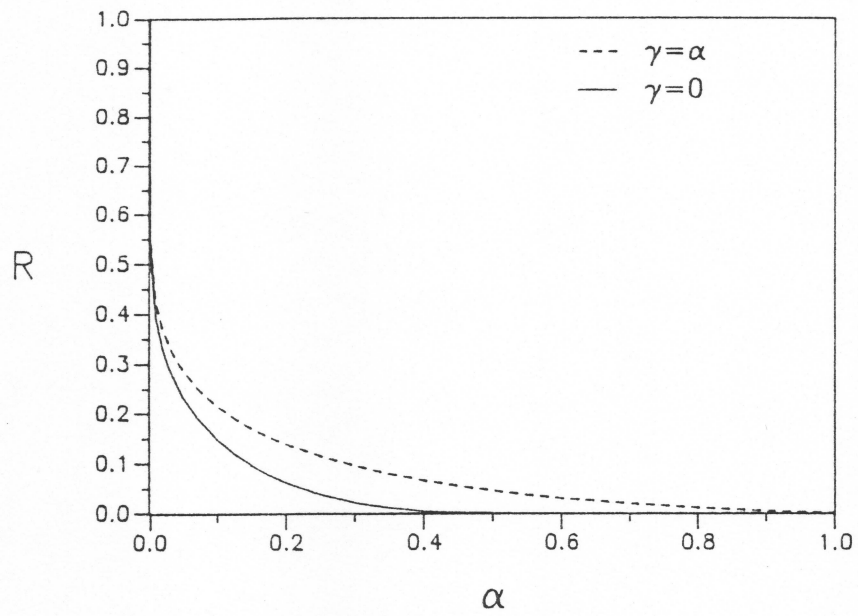


Abbildung 5.10: Nullstellen $R_0 = 2g_0$ von $f(g)$ für $\gamma = 0$ und $\gamma = \alpha$.

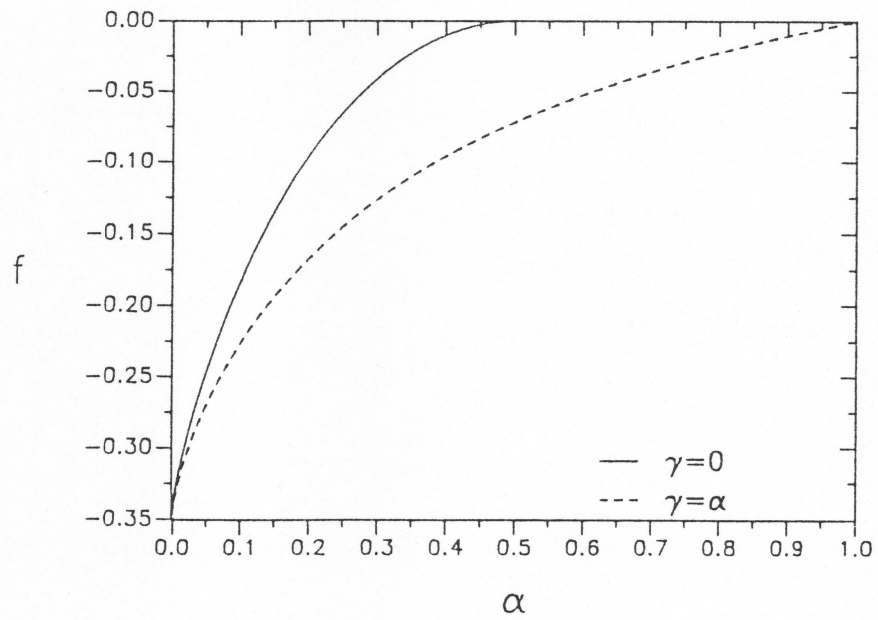


Abbildung 5.11: Der Grenzwert $f(g \rightarrow 0)$ in den Fällen $\gamma = 0$ und $\gamma = \alpha$.

5.2.2 Der Grenzwert von $f(\alpha, \mathbf{g}, \gamma, \varepsilon_{\max})$ für $\mathbf{g} \rightarrow \mathbf{0}$

Der Grenzwert $g \rightarrow 0$ ist auf besondere Weise mit der Systemgröße N verknüpft. Wenn ein sehr großes N vorgegeben ist, kann g trotzdem nicht beliebig klein werden. Die Zahl Ng der falschstehenden Spins muß nämlich eine makroskopische Größe bleiben, damit die Sattelpunktnäherung Gültigkeit behält. Aus diesem Grunde macht die Rechnung keine Aussage über Zustände, die nur an einzelnen Plätzen vom ersten Muster abweichen!

Wenn wir von der „Nähe zum ersten Muster“ reden, meinen wir trotzdem Zustände, die an makroskopisch vielen Punkten vom ersten Muster abweichen. Wir stellen uns den Grenzübergang $g \rightarrow 0$ also so vor, daß wir für jedes g die Größe N so groß wählen, daß Ng makroskopisch bleibt. Für jedes g tragen wir f ab und betrachten den Grenzwert $g \rightarrow 0$ in den Sattelpunktgleichungen (siehe Abschnitt 4.2.3).¹

Nach Gl.(4.30) gilt

$$\begin{aligned}
 f(\alpha, g, \gamma, \varepsilon_{\max}) \rightarrow & \\
 & -\frac{1-\alpha}{2} \ln(1-\alpha) - \frac{\alpha}{2} \ln \alpha - \frac{1}{2} \ln 2 + \\
 & + \frac{1-\alpha}{2} \ln(1-\gamma) + \frac{\alpha}{2} \ln(1+\gamma) \quad (g \rightarrow 0). \quad (5.16)
 \end{aligned}$$

Die Funktion ist für die beiden Parameterwerte $\gamma = 0$ und $\gamma = \alpha$ in Abb. 5.11 dargestellt.

$\gamma = \mathbf{0}$

Im Fall $\gamma = 0$ gilt Gl.(5.16) nur für $\alpha \in (0, \frac{1}{2}]$. Für $\alpha \in (\frac{1}{2}, 1)$ gilt dann Gl.(4.34):

$$\begin{aligned}
 f\left(\frac{1}{2} < \alpha < 1, g \rightarrow 0, \gamma = 0, \varepsilon_{\max}\right) = & \\
 -g \ln g - (1-g) \ln(1-g) + \mathcal{O}\left(\exp\left(-\frac{c}{g}\right)\right) \geq 0. & \quad (5.17)
 \end{aligned}$$

Ist also $\gamma = 0$ und $\alpha \geq \frac{1}{2}$, so gibt es in unmittelbarer Nähe des ersten Musters exponentiell viele metastabile Zustände. Wir können folglich annehmen, daß der dynamische Prozeß nicht ins erste Muster gelangt, sondern durch die

¹Ähnliche Überlegungen kann man auch für den Grenzwert $\alpha \rightarrow 0$ anstellen.

vielen metastabilen Zustände abgefangen wird. Dies ist eine Bestätigung der von Kanter und Sompolinsky gemachten Voraussage

$$\alpha_c = \frac{1}{2}$$

(siehe Abschnitt 3.2: Ist ein Muster für $\gamma = 0$, $\alpha \geq \frac{1}{2}$ vorgelegt, so ist auch ein Zustand \underline{S} metastabil, der nur an einem Platz i vom Muster abweicht $\implies \alpha_c = \frac{1}{2}$).

Für $\gamma = 0$ und $\alpha < \frac{1}{2}$ ist $f < 0$ für kleine g , und die Nullstelle g_0 von f liefert nach Abschnitt 5.2.1 eine (nicht ganz rigorose) Abschätzung für den Attraktionsradius. Streng genommen können wir also nur $\alpha_c \geq \frac{1}{2}$ nachweisen, denn wir haben nicht bewiesen, daß die gewöhnlichen Zustände dem dynamischen Prozeß im Fall $\alpha \geq \frac{1}{2}$ keinen „Steg“ ins Muster bauen können. In Abschnitt 5.2.1 hatten wir ja gesehen, daß dies prinzipiell möglich ist: Das System ist durchaus in der Lage, an exponentiell vielen metastabilen Zuständen vorbeizulaufen.

Die Betrachtung von Kanter und Sompolinsky zeigt schließlich auf, daß der Prozeß hier nicht ins Muster laufen kann und legt den kleinstmöglichen Wert $\alpha_c = \frac{1}{2}$ fest.

α_c für verschiedene γ

Wir gehen jetzt einen Schritt weiter und überlegen uns, daß bei vorgegebenen $\gamma = \gamma(\alpha)$ die Nullstelle der Kurve (α, f) mit f aus Gl.(5.16) das jeweilige kritische α liefern müßte. Wegen der strengen Monotonie von f bezüglich g ist dann nämlich

$$f(\alpha_c, g, \gamma, \varepsilon_{max}) > 0 \quad \forall g > 0 .$$

Nach Einsetzen sieht man, daß die Gleichung

$$f(\alpha_c, g \rightarrow 0, \gamma, \varepsilon_{max}) = 0$$

durch ein α_c gelöst wird, das wiederum der Bedingung

$$\gamma(\alpha_c) = 2\alpha_c - 1 \tag{5.18}$$

genügt.

Für $\gamma(\alpha) = 0$ wird somit $\alpha_c = \frac{1}{2}$ bestätigt. Ist $\gamma(\alpha) = \alpha$, so erhalten wir wie erwartet

$$\gamma = \alpha \implies \alpha_c = 1 . \quad (5.19)$$

Im Fall $\gamma = \alpha$, den wir naiv mit dem Projektornetzwerk ohne Hauptdiagonale in Verbindung bringen, haben also die Muster bei jedem α eine kleine Hülle, in der fast keine (d.h. nicht exponentiell viele) metastabilen Zustände liegen.

Unstetigkeit von f an der Stelle $g = 0$

Im Fall $g = 0$ gilt natürlich $\mathcal{N} = 1$ und somit $f = 0$. Allein daran sieht man schon, daß eine Sattelpunktintegration bei $g = 0$ nicht möglich ist, weil der bei einer Sattelpunktintegration auftretende Vorfaktor nicht gleich Eins (, sondern von $\mathcal{O}\left(\frac{1}{\sqrt{N}}\right)$) ist.

Um uns zu veranschaulichen, weshalb die Grenzwerte von $f(\alpha, g, \gamma, \varepsilon_{max})$ für $g \rightarrow 0$ negativ sein können, ziehen wir die Sattelpunktvariable

$$Z = \sum_{\mu=2}^p a_{\mu}^2$$

heran (siehe Gln. (5.3) bis (5.5)).

Für den Entwicklungskoeffizienten fürs erste Muster galt nach Gl.(C.12)

$$a_1 = 1 - 2g .$$

Zudem ist in Abschnitt 4.2.2 die Variable

$$V = \frac{Z}{4g(1-g)}$$

angegeben, deren Näherung für $g \rightarrow 0$ wir in Gl.(4.29) finden. Somit gilt

$$Z = \sum_{\mu=2}^p a_{\mu}^2 = 4g \cdot \frac{1+\gamma}{2(1-\alpha)} + \mathcal{O}(g^2) .$$

Wenn also für ein kleines g überhaupt ein metastabiler Zustand existiert, so muß in ihm das erste Muster stark vertreten sein. Außerdem müssen die anderen Muster und vermutlich auch der Orthogonalraum Anteile an ihm besitzen. V beschreibt, wie stark die Anteile der anderen Muster sind. Offenbar wird diese Forderung für kleinere g immer unerfüllbarer. Die Anteile der anderen Muster sollen zwar immer kleiner werden, aber nie verschwinden.

Bei $g = 0$ ist einfach $Z = 0$, es gilt $a_{\mu} = 0 \forall \mu \geq 2$ sowie $a_1 = 1$. Endlich sollen die anderen Muster überhaupt keine Anteile mehr besitzen. Diese Forderung ist viel leichter zu erfüllen und f macht einen Sprung.

5.3 Ergebnisse für $\langle\langle\mathcal{N}(\alpha, \mathbf{g} = \frac{1}{2}, \gamma = \alpha, \varepsilon)\rangle\rangle$

In diesem Abschnitt stelle ich als wichtigstes Ergebnis der vorliegenden Arbeit die Energieabhängigkeit der unkorrelierten metastabilen Zustände vor. Die in Abschnitt 4.4 aufgeführten Sattelpunktgleichungen wurden auf dem Rechner gelöst, und es wurden die Exponenten

$$f\left(\alpha, g = \frac{1}{2}, \gamma = \alpha, \varepsilon\right)$$

für verschiedene ε -Werte bestimmt.

Die in Abschnitt 4.1.1 durchgeführte Mittelung erfaßt wieder nur Zustände, die zu allen Mustern unkorreliert sind („unkorrelierte Zustände“). Weiterhin wird nur der Fall $\gamma = \alpha$ betrachtet, den wir mit dem Projektornetzwerk ohne Selbstkopplungen J_{ii} identifizieren möchten. Nur für dieses Modell existiert nämlich ein Hamiltonian für die serielle Glauber-Dynamik bei *endlicher* Temperatur τ (siehe Abschnitt 2.1). Kanter und Sompolinsky [KaSo87] haben für dieses Modell die Replica-Rechnung durchgeführt und die lokalen Minima der Freien Energie bestimmt. Dabei durften sie im Hamiltonian die Selbstkopplungen mitzählen, weil diese nur einen konstanten Beitrag lieferten:

$$\mathcal{H} = -\frac{1}{2} \sum_{i,j(\neq i)} J_{ij} S_i S_j - N \frac{\alpha}{2} = -\frac{1}{2} \sum_{i,j} J_{ij} S_i S_j .$$

\mathcal{H} war dann trotzdem ein Hamiltonian für die *serielle* Glauber-Dynamik *ohne* Selbstkopplungen.

Im Grenzwert $\tau \rightarrow 0$ erhielten sie lokale Minima der Energie für verschiedene Parameterbereiche. Uns interessiert natürlich ihr Ergebnis für die *lokalstabilen* Spinglaszustände. Diese sind besondere unkorrelierte metastabile Zustände und werden folglich von unserer Rechnung mit erfaßt.

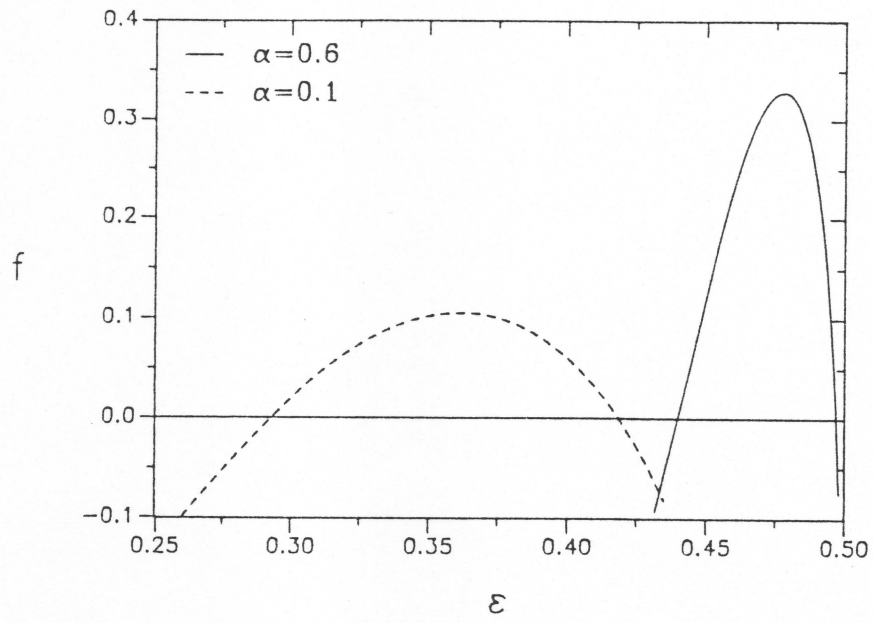


Abbildung 5.12: Spektren $(\varepsilon; f(\varepsilon))$ für $\alpha = 0.1; 0.6$

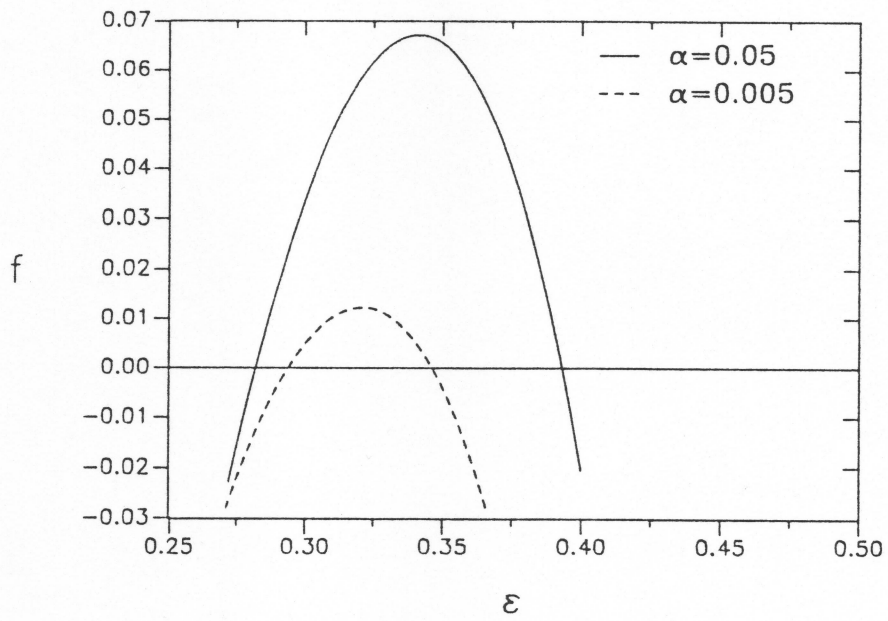


Abbildung 5.13: Spektren $(\varepsilon; f(\varepsilon))$ für $\alpha = 0.05; 0.005$

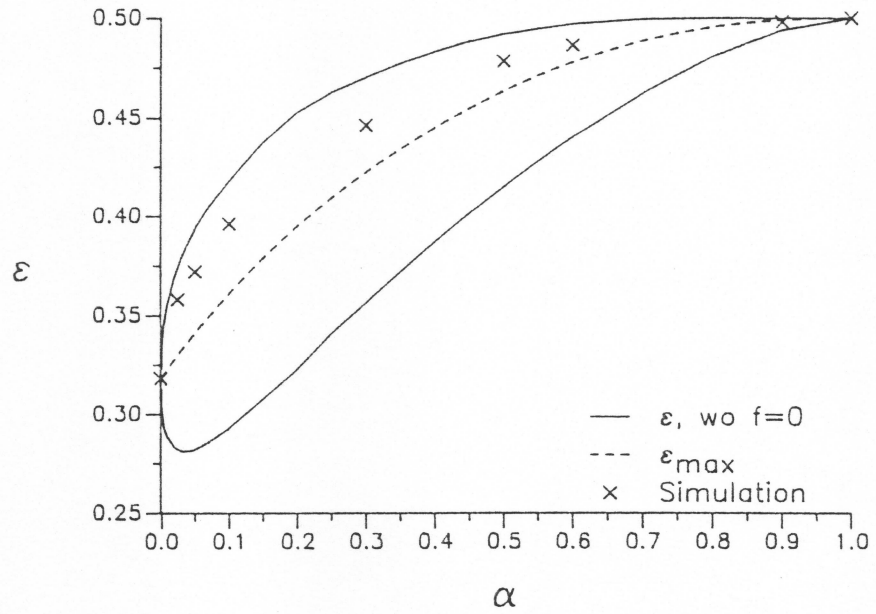


Abbildung 5.14: Die beiden Nullstellen und das Maximum im Spektrum als Funktion von α . Zusätzlich sind noch die Simulationsergebnisse von Kanter und Sompolinsky eingetragen.

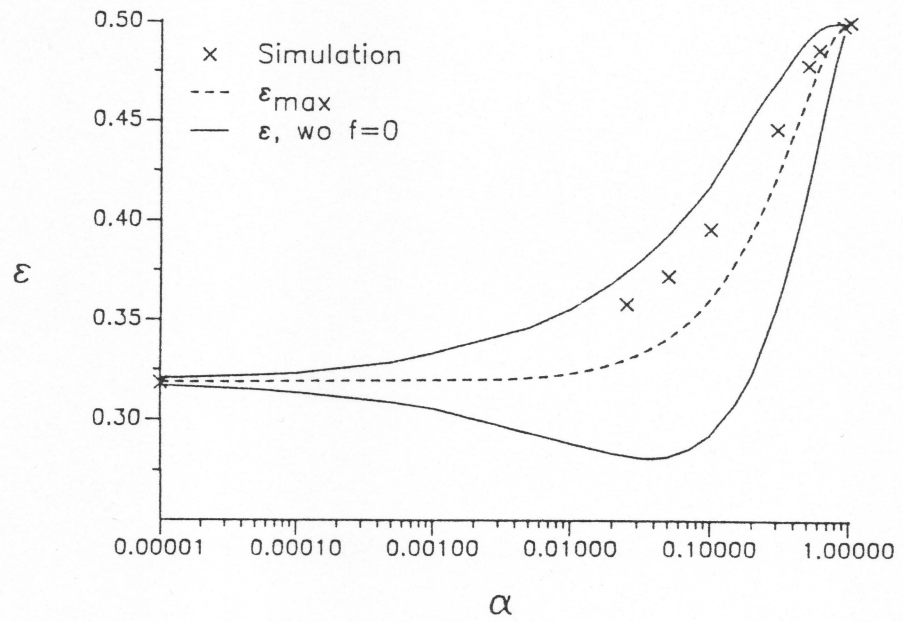


Abbildung 5.15: wie Abb. 5.14, logarithmische Darstellung.

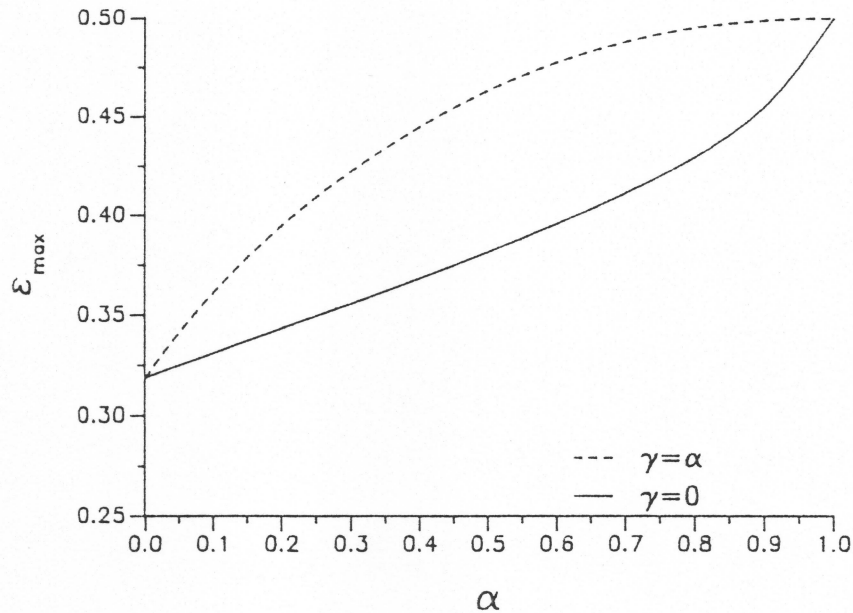


Abbildung 5.16: Die Kurven für ε_{max} in den Fällen $\gamma = 0$ und $\gamma = \alpha$.

In den Abbildungen 5.12 und 5.13 sind Kurven

$$\left(\varepsilon; f\left(\alpha, g = \frac{1}{2}, \gamma = \alpha, \varepsilon\right) \right)$$

für verschiedene α -Werte zu sehen. Eine solche Kurve stellt das Energieband dar, in dem (unkorrelierte) metastabile Zustände existieren, ich gebe ihr deshalb den anschaulichen Namen „Spektrum“. Anders als beim SK-Modell (siehe Abschnitt 2.3.4) gibt es hier nicht nur ein Spektrum. Für $\alpha \rightarrow 1$ werden die Spektren höher und schmaler, für $\alpha \rightarrow 0$ werden sie niedriger und ebenfalls schmaler.

In Abb. 5.14 und Abb. 5.15 sieht man die drei markanten Punkte in den einzelnen Spektren aufgetragen: Man kann die α -Abhängigkeit der linken und rechten Nullstellen sowie der Maxima ablesen. Die Kurve der Maxima aufzustellen bereitete dabei wenig Mühe, denn ε_{max} ist nach Gl.(4.59) durch die Sattelpunktvariable Z bestimmt,

$$\varepsilon_{max} = \frac{1}{2}(1 - \alpha)Z ,$$

die bereits aus der Lösung der Gleichungen in Abschnitt 4.3 bekannt ist. Den Funktionswert $f\left(\alpha, g = \frac{1}{2}, \gamma = \alpha, \varepsilon_{max}\right)$ haben wir ebenfalls schon in

Abschnitt 4.3 ausgerechnet und können ihn aus der obigen Abbildung 5.1 ablesen.

Nebenbei bemerkt bereitet dies alles auch im Fall $\gamma = 0$ wenig Mühe. In Abb. 5.16 sehen wir die beiden Kurven für ε_{max} im Vergleich.

Außerhalb der Bänder ist $f < 0$, es gilt sogar

$$f \rightarrow -\infty \quad \text{für} \quad \varepsilon \rightarrow \frac{1}{2} \quad \text{bzw.} \quad \varepsilon \rightarrow 0 .$$

Nach der Jensenschen Ungleichung können wir analog der Betrachtung in Abschnitt 5.1.2 wieder ausschließen, daß (unkorrelierte) metastabile Zustände in den Bereichen mit $f < 0$ existieren. Wir schließen wieder nicht aus, daß einige metastabile Zustände, die aus der Mittelung in Abschnitt 4.1.1 herausgefallen sind, in den für *unkorrelierte* metastabile Zustände verbotenen Bereichen leben können. Wir nehmen wieder an, daß ihre Zahl nicht exponentiell in N geht, da wir erwarten, daß unkorrelierte Zustände überwiegen, sobald exponentiell viele Zustände auftreten.

Insbesondere die von Kanter und Sompolinsky erwähnten „Mischzustände“ (siehe Abschnitt 3.2) können im verbotenen Bereich liegen und weit niedrigere Energien als die unkorrelierten metastabilen Zustände haben. Betrachten wir als Beispiel den Dreier-Mischzustand

$$S_i = \text{sign} \left(\xi_i^1 + \xi_i^2 + \xi_i^3 \right) \quad \forall i .$$

Dann gilt $m_\nu = \mathcal{O}\left(\frac{1}{\sqrt{N}}\right)$ für $\nu \geq 4$ und $m_1 = m_2 = m_3 = \frac{1}{2}$, das gleiche Verhalten zeigen die

$$\begin{aligned} a_\mu &= \sum_{\nu=1}^p C_{\mu\nu} m_\nu \sim m_\mu \\ \implies \varepsilon_M &= \frac{N}{2N} \sum_{\mu=1}^p a_\mu m_\mu \sim \frac{1}{2} \cdot 3 \cdot \frac{1}{4} = \frac{3}{8} = 0.375 . \end{aligned}$$

Dies gilt im α -Bereich ($0, \alpha_3 = 0.1$), denn aufgrund des Rauschens (in den a_ν, m_ν für $\nu > 3$) ist der Dreier-Mischzustand nur bis α_3 metastabil (siehe [KaSo87]). In Abb. 5.15 sieht man, daß $\varepsilon_M = \frac{3}{8}$ zumindest für $\alpha \leq 0.01$ über den oberen Nullstellen von f für die unkorrelierten Zustände liegt!

Bezüglich der unkorrelierten metastabilen Zustände selbst sind jedoch diese oberen Nullstellen exakte obere Schranken für die ε der Spinglaszustände, welche ja besondere unkorrelierte metastabile Zustände sind. Das ist nichts

anderes als die Wiederholung der Argumentation von Roberts für das SK-Spinglas (siehe Abschnitt 2.3.4).

Vergleichen wir also unsere Kurve der oberen Nullstellen mit der Kurve für die Energien der lokalstabilen Spinglaszustände von Kanter und Sompolinsky (siehe auch Abschnitt 3.2). Im Bereich *links* von

$$\alpha_g = 1 - \frac{2}{\pi},$$

für den nach Kanter und Sompolinsky die lokalstabilen Spinglaszustände noch existieren sollen, liegen von einem α an die Energien über den oberen Nullstellen aus Abb. 5.14. Für $\alpha \rightarrow \alpha_g (-)$ gilt sogar $\varepsilon \rightarrow \frac{1}{2}$. Das ist ein Widerspruch!

Die Replica-symmetrische Rechnung liefert also für $\tau \rightarrow 0$ falsche Ergebnisse. Dies liegt vermutlich an der Brechung der Replica-Symmetrie, die sich für $\tau \rightarrow 0$ stark auswirkt. Vermutlich ist in Wirklichkeit $\alpha_g = 1$, für alle α gibt es wohl Spinglaszustände.

Das haben Kanter und Sompolinsky bereits vermutet. Sie wollten ihr Ergebnis für α_g mit folgender Simulation testen:

Unkorrelierte Zustände (mit $g = \frac{1}{2}$) werden dem seriellen $\tau = 0$ -Glauber-Prozeß (ohne Selbstkopplungen) unterworfen. Die Zustände, die während des Prozesses durchlaufen werden, sind auch alle unkorrelierte Zustände, weil der Anfangszustand weit außerhalb des Attraktionsradius liegt. Die (mittlere) Energie der Endzustände wird gegen α aufgetragen.

Ein Blick auf Abb. 5.14 und Abb. 5.15 zeigt, daß die Simulationsergebnisse stets zwischen der Kurve der ε_{max} und der Kurve der oberen Nullstellen liegen. $\alpha_g = 1 - \frac{2}{\pi}$ ist in keiner Weise ausgezeichnet, so daß die Ergebnisse der Replica-Rechnung bereits aufgrund der Simulation verworfen werden können.

Für $\alpha \rightarrow 0$ und $\alpha \rightarrow 1$ laufen alle Kurven zusammen. Man kann die Grenzwerte also mit Hilfe der Kurve $(\alpha, \varepsilon_{max})$ bestimmen. Mit Gl.(4.44) und Gl.(4.51) erhält man für ε (!)

$$\begin{aligned} \varepsilon_{max} &= \frac{1}{2}(1 - \alpha)Z \\ \implies \varepsilon &\rightarrow \frac{1}{2} \cdot 1 \cdot \frac{2}{\pi} = \frac{1}{\pi} \quad (\alpha \rightarrow 0) \\ \varepsilon_{max} &= \frac{1}{2}(1 - \alpha)Z \end{aligned} \tag{5.20}$$

$$\implies \varepsilon \rightarrow \frac{1}{2} \cdot (1 - \alpha) \cdot \frac{1}{1 - \alpha} = \frac{1}{2} \quad (\alpha \rightarrow 1) . \quad (5.21)$$

Die vorliegende Arbeit liefert folglich eine Voraussage für die Kurve der Energien der Spinglaszustände. Eine Rechnung unter Berücksichtigung der Replica-Symmetrie-Brechung muß eine Kurve liefern, die zwischen der Kurve der ε_{max} und der Kurve der oberen Nullstellen liegt. Außen muß sie die Grenzwerte $\frac{1}{\pi}$ und $\frac{1}{2}$ annehmen. Die Kurve von Kanter und Sompolinsky liefert ebenfalls

$$\varepsilon \rightarrow \frac{1}{\pi} \quad (\alpha \rightarrow 0)$$

und gibt vermutlich für kleine α gute Werte.

Zuguterletzt weise ich noch auf eine verblüffende Analogie zum SK-Modell hin: R. D. Henkel und W. Kinzel haben fürs SK-Modell für verschiedene Arten der seriellen $\tau = 0$ -Glauber-Dynamik die mittleren Endenergien simuliert [HeKi87]. Ein Vergleich mit der Kurve von Roberts zeigte auch hier, daß der dynamische Prozeß über das Maximum im Spektrum hinausläuft und auf mittlerem Wege zwischen Maximum und Spinglas-Grundzustandsenergie stehenbleibt. Diese Analogie und die offenbar richtigen Grenzwerte (5.20) und (5.21) stellen wichtige Tests für die Rechnung in Kapitel 4 dar.

5.4 Vergleich des Projektor–Netzwerks mit dem Hopfield–Modell im Fall $\alpha \rightarrow 0$

Die vorliegende Arbeit liefert einen Einwand gegen die Erwartung:
„Für $\alpha \rightarrow 0$ wird das Projektor–Netzwerk zum Hopfield–Modell.“

Man wird zu dieser Erwartung veranlaßt, wenn man sich den kombinierten Überlapp eines Musters mit allen anderen (Zufalls–) Mustern ansieht [KaSo87]

$$\sum_{\nu=2}^p C_{1\nu} = \frac{1}{N} \sum_{\nu=2}^p \sum_{i=1}^N \xi_i^1 \xi_i^\nu = \mathcal{O}\left(\sqrt{\frac{p}{N}}\right) = \mathcal{O}(\sqrt{\alpha}) .$$

Man denkt sich also, daß für $\alpha \rightarrow 0$ die Nebendiagonalelemente in der Kopplungsmatrix ($C_{\mu\nu}$) keine Rolle mehr spielen, ($C_{\mu\nu}$) also so gut wie diagonal ist. Dann sind die Kopplungen des Projektornetzwerks

$$J_{ij} = \frac{1}{N} \sum_{\mu,\nu} \xi_i^\mu (C^{-1})_{\mu\nu} \xi_j^\nu$$

identisch mit den Kopplungen

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu$$

des Hopfield–Modells. Die Modelle sollten also in allen wichtigen Eigenschaften übereinstimmen. Testen wir diese Vermutung an der Zahl der metastabilen Zustände.

Zuerst beobachtet man im Fall $g = \frac{1}{2}$ das gleiche Verhalten für die Exponenten:

Hopfield–Modell (Gl.(2.34)):

$$f_H\left(\alpha, \gamma = \alpha, g = \frac{1}{2}, \varepsilon_{max}\right) = -\frac{\alpha}{2} \ln \alpha + \frac{\alpha}{2} \left(\ln \left(\frac{2}{\pi} \right) - 1 \right) + \mathcal{O}(\alpha^2) .$$

Projektornetzwerk (Gl.(4.46)):

$$f_P\left(\alpha, \gamma = \alpha, g = \frac{1}{2}, \varepsilon_{max}\right) = -\frac{\alpha}{2} \ln \alpha + \frac{\alpha}{2} \left(\ln \left(\frac{2}{\pi} \right) - 1 - \ln \left(1 - \frac{2}{\pi} \right) \right) + \mathcal{O}(\alpha^2) .$$

Die Maxima in den Energiespektren verhalten sich ebenfalls gleich in beiden Fällen:

$$\varepsilon_{max} \rightarrow \frac{1}{\pi} \quad (\alpha \rightarrow 0)$$

(siehe Gl.(2.37) und Gl.(5.20)).

Für die Abhängigkeit vom Hamming-Abstand erwartet man deshalb auch ein vergleichbares Verhalten der beiden Modelle. f_H sollte sich an f_P anschmiegen und erst für ein sehr kleines g wieder positiven Werten zustreben. Dies ist nicht der Fall, wie Abb. 5.17 aufzeigt. Auch für kleinere α sind die beiden Modelle unterschiedlich. Das gleiche gilt im Fall mit Selbstkopplungen ($\gamma = 0$).

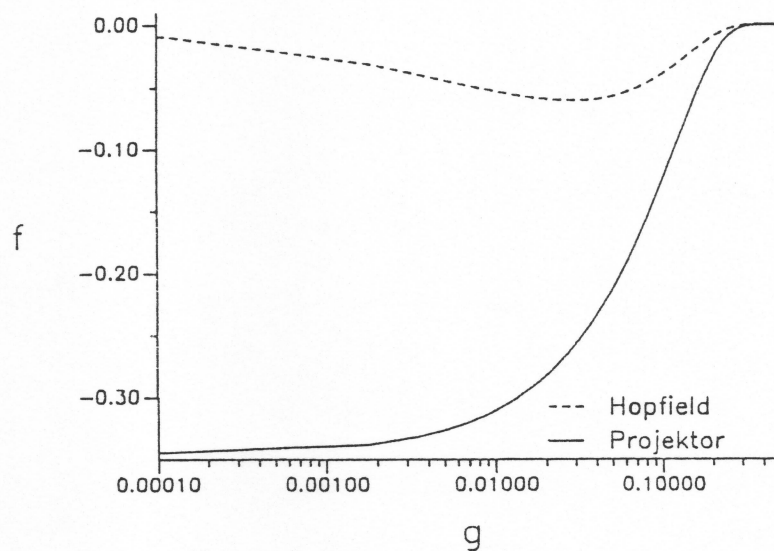


Abbildung 5.17: Vergleich von f_P mit f_H für $\alpha = 0.0001$ im Fall $\gamma = \alpha$.

Der prinzipielle Unterschied der beiden Modelle besteht darin, daß die Kopplungen des Hopfield-Modells eben doch keine exakten Projektorkopplungen sind. Die Projektormatrix legt nämlich strenge Zwangsbedingungen an, die beim Hopfield-Modell nicht auftreten. Als Beispiele seien genannt:

1. Die lokalen Energien der Muster sind für das Projektornetzwerk exakt 1, beim Hopfield-Modell genügen sie einer Gauß-Verteilung um 1 mit der Standardabweichung $\sqrt{\alpha}$.
2. Gl.(5.6) legt für $\gamma = 0$ die strenge Zwangsbedingung

$$\sum_{i=1}^N \lambda_i = \sum_{i=1}^N \lambda_i^2$$

für die lokalen Energien fest.

Kapitel 6

Zusammenfassung und Ausblick

Dieses Kapitel dient in seinem ersten Teil der Zusammenfassung der allerwichtigsten Ergebnisse dieser Arbeit. Dabei werden die umfangreichen Vorbereitungen in den Kapiteln 2 bis 4 vorausgesetzt, die getroffen wurden, um die Vorstellung und Interpretation der Ergebnisse in Kapitel 5 zu ermöglichen. Der zweite Teil des Kapitels gibt einen kurzen Ausblick auf die Rechnungen, die man im Zusammenhang mit der ursprünglichen Aufgabenstellung noch durchführen könnte.

6.1 Zusammenfassung

Diese Diplomarbeit befaßte sich mit der Berechnung der Zahl \mathcal{N} der metastabilen Zustände bei einem Neuronalen Netzwerk mit Projektorkopplungen. Die Kopplungen sind gegeben durch

$$J_{ij} = \frac{1}{N} \sum_{\mu, \nu} \xi_i^\mu (C^{-1})_{\mu\nu} \xi_j^\nu ,$$

wobei

$$C_{\mu\nu} = \frac{1}{N} \sum_{i=1}^N \xi_i^\mu \xi_i^\nu$$

die Korrelationsmatrix darstellt. In dem Netzwerk wurden $p = \alpha N$ Muster gespeichert. Die metastabilen Zustände wurden der Zwangsbedingung

$$\lambda_i = S_i \sum_{j=1}^N J_{ij} S_j > \gamma \quad \forall i$$

unterworfen, wobei γ die Stabilität genannt wird. Der Fall $\gamma = \alpha$ wurde dabei mit dem Projektornetzwerk ohne Selbstkopplungen in Verbindung gebracht.

Zusätzlich zur Stabilität betrachtete ich noch die Zwangsbedingungen Hamming-Abstand g und negative Energie pro Spin ε :

$$m_1 = \frac{1}{N} \sum_{i=1}^N \xi_i^1 S_i, \quad g = \frac{1 - m_1}{2}, \quad \varepsilon = \frac{1}{2N} \sum_{i,j} J_{ij} S_i S_j .$$

Aus vorherigen Betrachtungen für \mathcal{N} bei anderen Modellen (Spingläsern und Hopfield-Modell) heraus machten wir den Ansatz

$$\mathcal{N} = e^{Nh} ,$$

wobei h für große N als selbstmittelnd vorausgesetzt wurde, sofern zufällig ausgewürfelte Muster zugrunde gelegt worden waren. Die Jensensche Ungleichung besagt dann, daß der direkte Mittelwert von \mathcal{N} über die Muster eine obere Schranke für den typischen Wert

$$\langle\langle h \rangle\rangle = \frac{1}{N} \langle\langle \ln \mathcal{N} \rangle\rangle \leq \frac{1}{N} \ln \langle\langle \mathcal{N} \rangle\rangle \quad (6.1)$$

ist.

In der vorliegenden Arbeit wurde $\langle\langle \mathcal{N} \rangle\rangle$ berechnet. Für große N lieferte die Sattelpunktintegration das Ergebnis

$$\langle\langle \mathcal{N} \rangle\rangle \propto \exp(N f(\alpha, g, \gamma, \varepsilon)) .$$

Die Sattelpunktgleichungen zur Berechnung von f wurden gelöst. Die wichtigsten Ergebnisse in den einzelnen Parameterbereichen sind:

1. Abb. 5.1 für $f(\alpha, g = \frac{1}{2}, \gamma, \varepsilon_{max})$ in den Fällen $\gamma = 0$ und $\gamma = \alpha$. Für $\gamma = 0$ ist der Grenzwert $f \rightarrow \ln 2$ ($\alpha \rightarrow 1$) anschaulich klar, denn bei einem Projektornetzwerk, in dem $p = 1 \cdot N$ linear unabhängige Muster gespeichert sind, ist jeder Zustand metastabil.

Im Fall $\gamma = \alpha$ ist dasselbe Grenzverhalten überraschend, denn selbst die Muster sind im Fall $\gamma = 1$ instabil, f nimmt seinen Grenzwert im Fall $\alpha = 1$ also nicht an.

2. Die Abhängigkeit von f vom Hamming–Abstand g wurde in den Fällen $\gamma = 0$ und $\gamma = \alpha$ untersucht. $f(\alpha, g, \gamma, \varepsilon_{max})$ ist größer, als wenn die metastabilen Zustände gleichmäßig im Phasenraum verteilt wären. f ist für $g \in (0, \frac{1}{2}]$ streng monoton steigend mit g , wobei in der Regel gilt $f(g \rightarrow 0) < 0$. Dies besagt dann, daß in der Umgebung der Muster (fast) keine metastabilen Zustände liegen.

Das Projektornetzwerk macht also die Muster stabil und schafft um die Muster herum gleichzeitig ein „Vakuum“. Die Nullstellen von f bezüglich g kann man als untere Schranken für die Attraktionsradien bei der seriellen $\tau = 0$ –Glauber–Dynamik auffassen.

Aus dieser Überlegung folgen dann die maximalen Speicherkapazitäten $\alpha_c = \frac{1}{2}$ für den Fall $\gamma = 0$ und $\alpha_c = 1$ für den Fall $\gamma = \alpha$, den wir ja mit dem Fall ohne Hauptdiagonale identifizieren wollen.

3. Schließlich wurde noch die Energieabhängigkeit von

$$f\left(\alpha, g = \frac{1}{2}, \gamma = \alpha, \varepsilon\right)$$

untersucht. Es wurde nur $\gamma = \alpha$ betrachtet, weil nur für das Projektornetzwerk ohne Selbstkopplungen ein Hamiltonian der seriellen Glauber–Dynamik existiert. Meine Ergebnisse können somit verglichen werden mit dem $\tau = 0$ –Grenzwert der Replica–Rechnung von Kanter und Sompolinsky für die Freien Energien der lokalstabilen Spinglaszustände.

Bei festem α liefern die Kurven $(\varepsilon, f(\varepsilon))$ ein Band, in dem unkorrelierte metastabile Zustände auftreten können. Aus Simulationen der Endzustandsenergien beim SK–Spinglas ist bereits bekannt, daß der serielle $\tau = 0$ –Glauber–Prozeß zwischen Maximum und rechter Nullstelle in der Kurve stoppt. So verhält es sich auch hier. Die rechten Nullstellen sind wie beim Spinglas obere Schranken für die (negativen) Energien (pro Spin) der lokalstabilen Spinglaszustände.

Dadurch konnte das Ergebnis der Rechnung von Kanter und Sompolinsky für die Energien der Spinglaszustände widerlegt werden. Es gibt offenbar bei allen Werten von α lokalstabile Spinglaszustände.

Zusätzlich zu den obigen zentralen Ergebnissen wurden die folgenden zwei

Ergebnisse für das Neuronale Netzwerk mit Projektorkopplungen gewonnen, die, soviel ich weiß, bisher nicht bekannt waren:

- Die Wahrscheinlichkeitsverteilung der J_{ii} ist eine β -Verteilung mit Mittelwert α und Streuung

$$\sigma = \sqrt{\frac{\alpha(1-\alpha)}{\frac{N}{2} + 1}}.$$

- Die Projektormatrix zwingt den lokalen Energien die Zwangsbedingung

$$\sum_{i=1}^N \lambda_i = \sum_{i=1}^N \lambda_i^2$$

auf. Die Gesamtenergie kann deshalb auch für gewöhnliche Zustände nie positiv werden. Metastabile Zustände, die einen Anteil im Orthogonalraum der Muster haben, müssen einige λ_j mit $\lambda_j > 1$ haben.

Für das Hopfield-Modell berechnete ich ebenfalls die ε -Abhängigkeit von f . Die Grenzwerte

$$\varepsilon \rightarrow \frac{1}{\pi} \quad (\alpha \rightarrow 0)$$

und

$$\frac{\varepsilon_{max}}{\sqrt{\alpha}} \rightarrow 0.506 = \varepsilon_{\text{Spinglas}} \quad (\alpha \rightarrow \infty)$$

bestätigen die Richtigkeit des neuen Ergebnisses.

6.2 Ausblick

Obwohl in dieser Arbeit die interessantesten Parameterbereiche vorgestellt worden sind, in denen $f(\alpha, g, \gamma, \varepsilon)$ untersucht werden kann, gibt es weitere Kombinationen von Parametern, die man sich anschauen kann. So könnte man z.B. die g -Abhängigkeit von f bei verschiedenen vorgegebenen ε ausrechnen. Dazu müßte man jedoch wieder die sich aus Gl.(4.18) ergebenden komplizierten Sattelpunktgleichungen lösen.

Grundlegende Erkenntnisse würde ich mir von einer Replica-Rechnung versprechen, die zum Ziel hätte, den (gemittelten) typischen Exponenten

$$\langle\langle h \rangle\rangle = \frac{1}{N} \langle\langle \ln \mathcal{N} \rangle\rangle$$

zu bestimmen. Diese Rechnung wird besonders kompliziert, wenn man die Parameter g, γ und ε einbaut. Sie bietet aber die Möglichkeit festzustellen, wo f einen zu hohen Wert für $\langle\langle h \rangle\rangle$ veranschlagt. Die rechte Nullstelle der Kurve

$$(\varepsilon, \langle\langle h(\varepsilon) \rangle\rangle)$$

dürfte somit eine gute obere Schranke für das ε der Spinglaszustände liefern.

Der Vollständigkeit halber sei noch auf die (gemittelte) Zahl $\mathcal{N}(\tau)$ der Lösungen der TAP-Gleichungen verwiesen. Die Bestimmung von $\mathcal{N}(\tau)$ stellt eine Möglichkeit dar, die Temperatur in die Rechnung zu integrieren. Für $\tau \rightarrow 0$ verschwindet das Rauschen, und immer mehr Zustände werden stabil. Ich erwarte, daß der Grenzwert das gleiche Ergebnis $\langle\langle h \rangle\rangle$ wie die Replica-Rechnung liefert.

Die metastabilen Zustände sind vor allem für die *serielle* Dynamik von Bedeutung. Die *Paralleldynamik* hingegen hat im vorliegenden Fall außer metastabilen Zuständen auch die zahlenmäßig bei weitem überwiegenden Zweierzyklen als Endzustände. Die Berechnung der Zahl der Zweierzyklen nach Anhang D kann allerdings nur eine Ergänzung der dynamischen Betrachtungen in [HeOp90] darstellen, sie wird *nicht* dazu dienlich sein, die interessantesten Effekte in der remanenten Magnetisierung bei der Paralleldynamik zu erklären (siehe [HeOp90]).

Anhang A

Gauß–Integrale und die Φ –Funktion

A.1 Gauß–Integrale

Es werden zwei häufig verwendete Gauß–Integral–Formeln aufgeführt:

1.

$$\int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2}bz^2 + cz\right) \frac{dz}{\sqrt{2\pi}} = \frac{1}{\sqrt{b}} \exp\left(\frac{c^2}{2b}\right), \quad (\text{A.1})$$

dabei muß $b > 0$ sein. c kann komplex sein. Man nennt die Formel dann auch oft „Hubbard–Stratonovich–Identität“ [Hu59]. Sind b und c rein imaginär, so handelt es sich um das sogenannte „Fresnel–Integral“. Die Konvergenz ist dann ebenfalls gesichert.

2. Sei \mathbf{C} eine reelle symmetrische positiv definite Matrix und \vec{J} irgendein Vektor, der auch komplexe Komponenten haben kann. Dann gilt (siehe [NeOr88]):

$$\begin{aligned} & \left(\prod_{\nu=1}^p \int_{-\infty}^{+\infty} \frac{dx_\nu}{\sqrt{2\pi}} \right) \exp \left(-\frac{1}{2} \sum_{\mu,\nu} x_\mu C_{\mu\nu} x_\nu + \sum_{\nu=1}^p x_\nu J_\nu \right) \\ &= (\det \mathbf{C})^{-\frac{1}{2}} \exp \left(\frac{1}{2} \sum_{\mu,\nu} J_\mu (C^{-1})_{\mu\nu} J_\nu \right). \end{aligned}$$

In Vektorschreibweise lautet die Formel

$$\begin{aligned} & \left(\prod_{\nu=1}^p \int_{-\infty}^{+\infty} \frac{dx_\nu}{\sqrt{2\pi}} \right) \exp \left(-\frac{1}{2} \vec{x}^T \mathbf{C} \vec{x} + \vec{x}^T \vec{J} \right) \\ &= (\det \mathbf{C})^{-\frac{1}{2}} \exp \left(\frac{1}{2} \vec{J}^T \mathbf{C}^{-1} \vec{J} \right). \end{aligned} \quad (\text{A.2})$$

A.2 Die Φ -Funktion

Die Definition der Φ -Funktion ist

$$\Phi(t) := \int_{-\infty}^t \frac{d\lambda}{\sqrt{2\pi}} e^{-\frac{1}{2}\lambda^2}. \quad (\text{A.3})$$

Es gelten folgende Entwicklungen [AbSt65]

$$\Phi(t) = \frac{1}{2} \left(1 + \sqrt{\frac{2}{\pi}} t - \frac{1}{3\sqrt{2\pi}} t^3 + \mathcal{O}(t^5) \right) \quad (t \rightarrow 0) \quad (\text{A.4})$$

$$\Phi(t) = 1 - \frac{1}{t} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}t^2} \left(1 - \frac{1}{t^2} + \mathcal{O}\left(\frac{1}{t^4}\right) \right) \quad (t \rightarrow \infty) \quad (\text{A.5})$$

$$\Phi(t) = -\frac{1}{t} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}t^2} \left(1 - \frac{1}{t^2} + \mathcal{O}\left(\frac{1}{t^4}\right) \right) \quad (t \rightarrow -\infty) \quad (\text{A.6})$$

Anhang B

Zahl der metastabilen Zustände beim Hopfield–Modell

In diesem Anhang wird der direkte Mittelwert der Zahl der metastabilen Zustände $\mathcal{N}(\alpha, g, \varepsilon)$ beim Hopfield–Modell berechnet und insbesondere das in Abschnitt 2.4 geschilderte Ergebnis für die ε –Abhängigkeit hergeleitet. Analog Gl.(2.14) gilt

$$\begin{aligned} & \langle \langle \mathcal{N}(\alpha, g, E) \rangle \rangle \\ &= \sum'_{\{S_i\}} \left(\prod_{j=1}^N \Theta \left(S_j \sum_{k(\neq j)} J_{jk} S_k \right) \right) \cdot \delta \left(E + \frac{1}{2} \sum_{j,k(\neq j)} J_{jk} S_j S_k \right). \quad (\text{B.1}) \end{aligned}$$

Der Strich an der Summe deutet an, daß nur über Zustände mit vorgegebenem Hamming–Abstand g zum ersten Muster summiert wird. Die Darstellung (2.13) für die Θ –Funktion liefert mit $J_{ii} = \alpha$

$$\begin{aligned} & \mathcal{N}(\alpha, g, E) \\ &= \sum'_{\{S_i\}} \left(\prod_{j=1}^N \int_0^\infty d\lambda_j \right) \left(\prod_{j=1}^N \delta \left(\lambda_j - S_j \sum_{k=1}^N J_{jk} S_k + \alpha \right) \right) \cdot \\ & \quad \cdot \delta \left(E + \frac{1}{2} \sum_{j,k(\neq j)} J_{jk} S_j S_k \right) \end{aligned}$$

$$\begin{aligned}
&= \sum'_{\{S_i\}} \left(\prod_{j=1}^N \int_0^\infty d\lambda_j \int_{-i\infty}^{+i\infty} \frac{d\mu_j}{2\pi i} \right) \int_{-i\infty}^{+i\infty} \frac{dw}{2\pi i} \\
&\quad \exp \left[\sum_{j=1}^N \mu_j \left(\lambda_j - S_j \sum_{k=1}^N J_{jk} S_k + \alpha \right) + w \left(E + \frac{1}{2} \sum_{j=1}^N \lambda_j \right) \right]. \quad (\text{B.2})
\end{aligned}$$

Dabei wurde $\delta(\lambda_j - S_j \sum_{k(\neq j)} J_{jk} S_k)$ verwendet, um

$$\frac{1}{2} \sum_{j,k(\neq j)} J_{jk} S_j S_k \quad \text{durch} \quad \frac{1}{2} \sum_{j=1}^N \lambda_j$$

zu ersetzen. Im Gegensatz zur Rechnung im Anhang C fehlen in der Energie die Selbstkopplungsterme. Analog zu den Gl.(4.8) folgenden Betrachtungen in Abschnitt 4.1.1 wird jetzt die Spur über die $\{S_j\}$ ausgeführt. Wichtig ist dabei die Voraussetzung, daß die $\{S_j\}$ keine Korrelation zu den anderen Mustern ξ^ν , $\nu \geq 2$, besitzen. Mit den ξ_i^μ sind dann die $\sigma_j^\mu := S_j \xi_j^\mu$, $\mu \geq 2$, ebenfalls Zufallsvariablen und man erhält nach Umordnung der j -Indizes und Mittelung über die σ_j^μ

$$\begin{aligned}
\langle\langle \mathcal{N}(\alpha, g, E) \rangle\rangle &= \binom{N}{Ng} \left(\prod_{j=1}^N \int_0^\infty d\lambda_j \int_{-i\infty}^{+i\infty} \frac{d\mu_j}{2\pi i} \right) \int_{-i\infty}^{+i\infty} \frac{dw}{2\pi i} \\
&\quad \exp \left[wE + \sum_{j=1}^N \left(\mu_j (\lambda_j + \alpha) + \frac{1}{2} w \lambda_j \right) + \sum_{j=1}^{Ng} \mu_j (1 - 2g) \right. \\
&\quad \quad \left. + \sum_{j=Ng+1}^N \mu_j (2g - 1) \right] \cdot \\
&\quad \cdot \left\langle\left\langle \exp \left(\sum_{\nu=2}^p \sum_{j,k} \left(-\frac{1}{N} \mu_j \sigma_j^\nu \sigma_k^\nu \right) \right) \right\rangle\right\rangle \quad (\text{B.3})
\end{aligned}$$

Nach Ausführung der Mittelung über die σ_j^μ wird der sich ergebende cosh bis zu $\mathcal{O}\left(\frac{1}{N}\right)$ entwickelt und $\ln(1+x) \sim x$ sowie Gl.(A.1) verwendet:

$$\left\langle\left\langle \exp \left(\sum_{\nu=2}^p \sum_{j,k} \left(-\frac{1}{N} \mu_j \sigma_j^\nu \sigma_k^\nu \right) \right) \right\rangle\right\rangle = \left(\frac{1}{2^N} \sum_{\{\sigma_j\}} \exp \left(\sum_{j,k} -\frac{1}{N} \mu_j \sigma_j \sigma_k \right) \right)^{p-1}$$

mit

$$\begin{aligned}
& \frac{1}{2^N} \sum_{\{\sigma_j\}} \exp \left(\sum_{j,k} -\frac{1}{N} \mu_j \sigma_j \sigma_k \right) \\
&= \int_{-\infty}^{+\infty} dS \int_{-\infty}^{+\infty} \frac{ds}{2\pi i} \frac{1}{2^N} \sum_{\{\sigma_j\}} \exp \left[s \left(S - \frac{1}{\sqrt{N}} \sum_{k=1}^N \sigma_k \right) - \frac{1}{\sqrt{N}} \sum_{j=1}^N \mu_j \sigma_j S \right] \\
&= \int_{-\infty}^{+\infty} dS \int_{-\infty}^{+\infty} \frac{ds}{2\pi i} e^{sS} \prod_{j=1}^N \frac{2}{2} \cosh \frac{1}{\sqrt{N}} (\mu_j S + s) \\
&\sim \int_{-\infty}^{+\infty} dS \int_{-\infty}^{+\infty} \frac{ds}{2\pi i} e^{sS} \exp \left[\frac{1}{2N} \sum_{j=1}^N (\mu_j^2 S^2 + 2sS\mu_j + s^2) \right] \\
&= \int_{-\infty}^{+\infty} \frac{dS}{\sqrt{2\pi}} \exp \left[\frac{1}{2N} \sum_{j=1}^N \mu_j^2 S^2 + \frac{1}{2} S^2 \left(1 + \frac{1}{N} \sum_{j=1}^N \mu_j \right)^2 \right].
\end{aligned}$$

Jetzt werden die zwei Sattelpunktvariablen

$$U = \frac{1}{N} \sum_{j=1}^N \mu_j^2 \quad \text{und} \quad V = \frac{1}{N} \sum_{j=1}^N \mu_j$$

eingeführt. Das Integral über S wird ausgeführt und man erhält folgendes, nachdem man Substitutionen in den λ_i durchgeführt und $\varepsilon = -\frac{E}{N}$ gesetzt hat:

$$\begin{aligned}
\langle\langle \mathcal{N}(\alpha, g, \varepsilon) \rangle\rangle &= \\
& \binom{N}{Ng} \int_{-\infty}^{+\infty} dU \int_{-\infty}^{+\infty} N \frac{du}{2\pi i} \int_{-\infty}^{+\infty} \frac{dV}{i} \int_{-\infty}^{+\infty} N \frac{dv}{2\pi} \int_{-\infty}^{+\infty} \frac{dw}{2\pi i} \\
& \exp \left(N \left[-w\varepsilon + uU + vV - \frac{\alpha}{2} \ln \left((V+1)^2 - U \right) \right] \right) \\
& \left(\int_{\alpha+(1-2g)}^{\infty} d\lambda \int_{-\infty}^{+\infty} \frac{d\mu}{2\pi i} \exp \left[-u\mu^2 + \mu(\lambda - v) + \frac{1}{2} w(\lambda - \alpha + 2g - 1) \right] \right)^{Ng} \\
& \left(\int_{\alpha-(1-2g)}^{\infty} d\lambda \int_{-\infty}^{+\infty} \frac{d\mu}{2\pi i} \exp \left[-u\mu^2 + \mu(\lambda - v) + \frac{1}{2} w(\lambda - \alpha - (2g - 1)) \right] \right)^{N(1-g)}
\end{aligned}$$

Da sich die letzten beiden Faktoren nur im Vorzeichen von $m_1 = 1 - 2g$ unterscheiden, reicht es aus, den ersten von ihnen zu berechnen. Die μ -Integration wird mit Hilfe von Gl.(A.1) ausgeführt. Im verbleibenden λ -Integral wird anschließend

$$\lambda' = \frac{1}{\sqrt{-2u}} \left(\lambda - \left(v + \frac{1}{2}(-2u)w \right) \right)$$

substituiert. Wir erhalten

$$\begin{aligned} & \langle \langle \mathcal{N}(\alpha, g, \varepsilon) \rangle \rangle \\ &= N^2 \int_{-\infty}^{+\infty} dU \int_{-i\infty}^{+i\infty} \frac{du}{2\pi i} \int_{-i\infty}^{+i\infty} \frac{dV}{i} \int_{-\infty}^{+\infty} \frac{dv}{2\pi} \int_{-i\infty}^{+i\infty} \frac{dw}{2\pi i} \cdot \\ & \quad \cdot \exp(N f(\alpha, g, \varepsilon, u, U, v, V, w)) \end{aligned} \quad (\text{B.4})$$

mit

$$\begin{aligned} & f(\alpha, g, \varepsilon, u, U, v, V, w) \\ &= -w\varepsilon + uU + vV - \frac{\alpha}{2} \ln \left((V+1)^2 - U \right) + \\ & \quad + \frac{1}{2}w \left(-\alpha + (1-2g)^2 \right) + \frac{1}{2} \cdot \frac{1}{-2u} \left(v + \frac{1}{2}(-2u)w \right)^2 - \frac{1}{2} \cdot \frac{v^2}{-2u} + \\ & \quad + g \ln \Phi(t_1) + (1-g) \ln \Phi(t_2) + \frac{1}{N} \ln \binom{N}{Ng}, \end{aligned} \quad (\text{B.5})$$

dabei ist

$$t_1 = \frac{v + \frac{1}{2}(-2u)w - \alpha - (1-2g)}{\sqrt{-2u}}, \quad t_2 = \frac{v + \frac{1}{2}(-2u)w - \alpha + (1-2g)}{\sqrt{-2u}}.$$

Φ ist die Φ -Funktion (siehe Anhang A.2). Für große N faßt man Gl.(B.4) als Sattelpunktintegral auf. Dann ist

$$\langle \langle \mathcal{N}(\alpha, g, \varepsilon) \rangle \rangle \propto e^{N f(\alpha, g, \varepsilon)},$$

wobei $f(\alpha, g, \varepsilon)$ Sattelpunktwert über u, U, v, V, w von $f(\alpha, g, \varepsilon, u, U, v, V, w)$ aus Gl.(B.5) ist. Durch Ableiten nach U und V ersetzt man diese Größen zunächst am Sattelpunkt. Auflösen der beiden Gleichungen

$$\frac{\partial f}{\partial U} = 0, \quad \frac{\partial f}{\partial V} = 0$$

liefert

$$(V + 1)^2 - U = \frac{\alpha}{-2u} \quad (\text{B.6})$$

$$V = \frac{v}{-2u} - 1 \quad (\text{B.7})$$

$$U = \frac{v^2}{4u^2} + \frac{\alpha}{2u} . \quad (\text{B.8})$$

Um w aus dem Argument der Φ -Funktion zu beseitigen, substituiert man

$$v' = v + \frac{1}{2}(-2u)w . \quad (\text{B.9})$$

Nach Einsetzen ergibt sich unter Verwendung von Gl.(B.6) bis Gl.(B.8):

$$\begin{aligned} f(\alpha, g, \varepsilon, u, v', w) = & \\ & w \left(-\varepsilon - \left(\frac{\alpha}{2} - \frac{(1-2g)^2}{2} - \frac{1}{2}(-2u) \right) \right) + \\ & + \frac{\alpha}{2} - \frac{\alpha}{2} \ln \alpha - g \ln g - (1-g) \ln(1-g) + \frac{\alpha}{2} \ln(-2u) + \\ & + \frac{1}{2} \cdot \frac{(v')^2}{-2u} - v' + g \ln \Phi(t_1) + (1-g) \ln \Phi(t_2) . \end{aligned} \quad (\text{B.10})$$

Dabei wurde noch der asymptotische Ausdruck für $\frac{1}{N} \ln \left(\binom{N}{Ng} \right)$ für große N verwendet (Gl.(4.12)). Zudem gilt

$$t_1 = \frac{v' - \alpha - (1-2g)}{\sqrt{-2u}}; \quad t_2 = \frac{v' - \alpha + (1-2g)}{\sqrt{-2u}} .$$

Da die Jacobi-Matrix der Transformation (B.9) nichtsingulär ist, reicht es aus, den Sattelpunkt von f aus Gl.(B.10) bezüglich den Variablen u, v', w zu suchen. Ableiten nach w erleichtert die Arbeit nun wesentlich, denn man erhält sofort den Sattelpunktwert von u aus der Gleichung

$$\varepsilon = -\frac{E}{N} = \frac{(1-2g)^2}{2} - \frac{\alpha}{2} + \frac{1}{2}(-2u) \quad (\text{B.11})$$

Setzt man Gl.(B.11) wieder in Gl.(B.10) ein, so stellt man fest, daß der Sattelpunktwert von w für die weitere Betrachtung unerheblich ist. Es ist also nicht erforderlich, w aus der Gleichung

$$\frac{\partial f(\alpha, g, \varepsilon, u, v', w)}{\partial w} = 0$$

zu bestimmen. Mit den Substitutionen

$$-2u = \alpha a \quad \text{und} \quad v' - \alpha = -b\alpha$$

folgt das in Gl.(2.30) angegebene Endergebnis.

Der Grenzwert von f für $\alpha \rightarrow \infty$ wurde numerisch bestimmt, während Ergebnis (2.34) für $\alpha \rightarrow 0$ einer zu Abschnitt 4.3.2 analogen Betrachtung entstammt:

Am Maximum von $f(\alpha, g = \frac{1}{2}, \varepsilon)$ bezüglich ε ergibt sich die Sattelpunktgleichung:

$$r = \frac{1}{\sqrt{2\pi\alpha}} \cdot \frac{\Phi'\left(r\sqrt{\frac{\alpha}{a}}\right)}{\Phi\left(r\sqrt{\frac{\alpha}{a}}\right)} \cdot \frac{1}{\sqrt{a}} \quad (\text{B.12})$$

mit $b = -r$ und $a = \frac{1+r}{1-r}$,
wobei ε_{max} nach Gl.(2.32) aus a folgt.

Mit dem Ansatz

$$1 - r = r_1\alpha + r_2\alpha^2 + \mathcal{O}(\alpha^3)$$

ist dann

$$\sqrt{\frac{\alpha}{a}} r \rightarrow 0 \quad (\alpha \rightarrow 0)$$

erfüllt, und man kann Gl.(A.4) verwenden. Auf der rechten Seite von Gl.(B.12) steht dann auch eine Taylorreihe in α . Nach Koeffizientenvergleich erhält man

$$r = 1 - \pi\alpha - \left(-\frac{5}{2}\pi^2 + 2\pi\right)\alpha^2 + \mathcal{O}(\alpha^3) .$$

Daraus folgt

$$a = \frac{2}{\pi\alpha} + 4\left(1 - \frac{1}{\pi}\right) + \mathcal{O}(\alpha)$$

und

$$\varepsilon_{max} = -\frac{\alpha}{2}(1 - a) = \frac{1}{\pi} + \frac{\alpha}{2}\left(3 - \frac{4}{\pi}\right) + \mathcal{O}(\alpha^2) .$$

Anhang C

Nebenrechnungen zu Kapitel 4

C.1 Die Mittelung über die σ_j^μ

Vernachlässigung von Termen der Ordnung $\mathcal{O}\left(\frac{1}{N^2}\right)$ bei der Mittelung über den letzten Faktor von Gl.(4.9) liefert

$$\begin{aligned} & \left\langle \left\langle \exp \left[\frac{1}{\sqrt{N}} \sum_{j=1}^N \sum_{\mu=2}^p \sigma_j^\mu (-a_\mu \mu_j + l_\mu (\lambda_j - 1)) \right] \right\rangle \right\rangle \\ &= \frac{1}{2^{N(p-1)}} \sum_{\{\sigma_j^\mu\}} \exp[\dots] = \prod_{j=1}^N \prod_{\mu=2}^p \frac{2}{2} \cosh \left(\frac{1}{\sqrt{N}} (-a_\mu \mu_j + l_\mu (\lambda_j - 1)) \right) \\ &\sim \prod_{j=1}^N \prod_{\mu=2}^p \exp \left[\ln \left(1 + \frac{1}{2N} (-a_\mu \mu_j + l_\mu (\lambda_j - 1))^2 \right) \right] \\ &\sim \exp \left[\sum_{j=1}^N \sum_{\mu=2}^p \frac{1}{2N} (-a_\mu \mu_j + l_\mu (\lambda_j - 1))^2 \right] \\ &\sim \exp \left[\sum_{j=1}^N \frac{1}{2} (X(\lambda_j - 1)^2 - 2Y\mu_j(\lambda_j - 1) + Z\mu_j^2) \right], \end{aligned}$$

wobei X, Y, Z Abkürzungen für

$$X = \frac{1}{N} \sum_{\mu=2}^p l_\mu^2, \quad Y = \frac{1}{N} \sum_{\mu=2}^p a_\mu l_\mu, \quad Z = \frac{1}{N} \sum_{\mu=2}^p a_\mu^2$$

sind.

C.2 Berechnung des Gauß-Integrals zur Potenz $p - 1$

Wir wenden wiederholt Gl.(A.1) an und erhalten

$$\begin{aligned} I_{p-1} &= \int_{-\infty}^{+\infty} da \int_{-i\infty}^{+i\infty} \frac{dl}{2\pi i} \exp[-xl^2 - yla - za^2] \\ &= \frac{1}{\sqrt{2z}} \int_{-i\infty}^{+i\infty} \frac{dl}{\sqrt{2\pi i}} \exp\left[l^2\left(\frac{y^2}{4z} - x\right)\right] = \frac{1}{\sqrt{y^2 - 4xz}}. \end{aligned}$$

C.3 Berechnung des Gauß-Integrales zur Potenz Ng

Berechne das Integral zur Potenz Ng : I_g . Ich verwende wieder Gl.(A.1) und erhalte:

$$\begin{aligned} I_g &= \int_{\gamma}^{\infty} \frac{d\lambda}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{Z}} \cdot \\ &\quad \exp\left[-\frac{1}{2Z}(\lambda(1-Y) + a_1 + Y)^2 + \frac{1}{2}w\lambda + l_1(1-\lambda) + \frac{1}{2}X(1-\lambda)^2\right]. \end{aligned}$$

Ich führe die Abkürzung

$$K = (1-Y)^2 - XZ \tag{C.1}$$

ein und schreibe

$$\begin{aligned} I_g &= \int_{\gamma}^{\infty} \frac{d\lambda}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{Z}} \cdot \\ &\quad \exp\left[-\frac{1}{2} \cdot \frac{K}{Z} \left(\lambda - \frac{1}{K} \left(\frac{1}{2}wZ - l_1Z - XZ - (a_1 + Y)(1-Y)\right)\right)^2 + \right. \\ &\quad \left. + \frac{1}{2} \cdot \frac{K}{Z} \cdot \frac{1}{K^2} \left(\frac{1}{2}wZ - l_1Z - XZ - (a_1 + Y)(1-Y)\right)^2 \right. \\ &\quad \left. - \frac{1}{2Z}(a_1 + Y)^2 + l_1 + \frac{1}{2}X\right]. \end{aligned}$$

Nach der Substitution

$$\lambda' = \sqrt{\frac{K}{Z}} \left(\lambda - \frac{1}{K} \left(\frac{1}{2} w Z - l_1 Z - X Z - (a_1 + Y)(1 - Y) \right) \right)$$

und einigen Umformungen erhält man schließlich

$$\begin{aligned} I_g = & \exp \left[-\frac{1}{2} \ln K + \frac{1}{8} w^2 \cdot \frac{Z}{K} + \frac{1}{2} \cdot \frac{w}{K} (-l_1 Z - a_1(1 - Y) + K - 1 + Y) \right. \\ & \left. + \frac{1}{2K} (l_1^2 Z + 2l_1 a_1(1 - Y) + 2l_1(1 - Y) + a_1^2 X + 2a_1 X + X) \right] \cdot \\ & \cdot \Phi(t_1) , \end{aligned} \tag{C.2}$$

$$t_1 = \sqrt{\frac{K}{Z}} \left(1 - \gamma + \frac{1}{K} \left(\frac{1}{2} w Z - l_1 Z - a_1(1 - Y) + Y - 1 \right) \right) ,$$

$$K = (1 - Y)^2 - X Z ,$$

$$\Phi(t_1) = \int_{-\infty}^{t_1} \frac{d\lambda}{\sqrt{2\pi}} e^{-\frac{1}{2}\lambda^2} .$$

Das Integral zur Potenz $N(1 - g)$ entsteht aus Gl.(C.2), wenn man vor l_1 und a_1 die Vorzeichen wechselt.

C.4 Vereinfachung des Exponenten Nf am Sattelpunkt durch Ableiten

Für große N ist das Integral (4.14) durch den Sattelpunktwert von f bestimmt. Nach Ableiten nach x, y, z kann man diese drei Variablen folgendermaßen durch ihre konjugierten Brüder X, Y, Z ersetzen:

$$\begin{aligned} \frac{\partial f}{\partial x} &= X - \frac{\alpha}{2} \cdot \frac{-4z}{y^2 - 4xz} = 0 \\ \frac{\partial f}{\partial y} &= Y - \frac{\alpha}{2} \cdot \frac{2y}{y^2 - 4xz} = 0 \\ \frac{\partial f}{\partial z} &= Z - \frac{\alpha}{2} \cdot \frac{-4x}{y^2 - 4xz} = 0 . \end{aligned}$$

Dann gilt

$$xX + yY + zZ = \frac{\alpha}{2} \cdot \frac{1}{y^2 - 4xz} (-4xz + 2y^2 - 4xz) = \alpha \quad (\text{C.3})$$

und

$$\begin{aligned} Y^2 - XZ &= \frac{\alpha^2}{4} \cdot \frac{4y^2 - 16xz}{(y^2 - 4xz)^2} = \frac{\alpha^2}{y^2 - 4xz} \\ \implies y^2 - 4xz &= \frac{\alpha^2}{Y^2 - XZ}. \end{aligned} \quad (\text{C.4})$$

Gl.(C.3) und Gl.(C.4) ersetzen x, y, z in f . Man erhält so das allgemeine Endergebnis (4.18).

C.5 Das Ergebnis $\langle\langle \mathcal{N}(\alpha, \mathbf{g}, \gamma, \varepsilon_{\max}) \rangle\rangle$

Der Wert von f am Sattelpunkt bezüglich ε ist der Wert, den man erhalten hätte, wenn man die Energieabhängigkeit nicht berücksichtigt hätte. Formal folgt dann $w = 0$ aus $\frac{\partial f}{\partial \varepsilon} = 0$. Drei der verbleibenden fünf Variablen werden folgendermaßen transformiert:

$$K := (1 - Y)^2 - XZ \quad (\text{C.5})$$

$$r := \frac{(1 - \gamma)((1 - Y)^2 - XZ) + Y - 1}{\sqrt{(1 - Y)^2 - XZ}} = \frac{(1 - \gamma)K + Y - 1}{\sqrt{K}} \quad (\text{C.6})$$

$$s := \frac{-a_1(1 - Y) - l_1z}{\sqrt{(1 - Y)^2 - XZ}} = \frac{-a_1(1 - Y) - l_1z}{\sqrt{K}}, \quad (\text{C.7})$$

a_1 und Z bleiben gleich.

Die Jacobi-Determinante der obigen Transformation ist von Null verschieden, falls $Z > 0$ und $K > 0$ sind. Dies ist in den später berechneten Sattelpunkten der Fall. In Umgebung der Sattelpunkte ist die Transformation somit nichtsingulär. Die Umkehrtransformation ist

$$Y = r\sqrt{K} - (1 - \gamma)K + 1 \quad (\text{C.8})$$

$$X = \frac{1}{Z} \left((r\sqrt{K} - (1 - \gamma)K)^2 - K \right) \quad (\text{C.9})$$

$$l_1 = -\frac{1}{Z} \left(\sqrt{K}s - a_1 \left(r\sqrt{K} - (1-\gamma)K \right) \right). \quad (\text{C.10})$$

Setzt man diese drei Gleichungen und $w = 0$ in Gl.(4.18) ein, so ergibt sich nach einigen algebraischen Rechnungen

$$\begin{aligned} f(\alpha, g, \gamma, \varepsilon_{max}, K, r, Z, s, a_1) & \quad (\text{C.11}) \\ = & - (1-\alpha) \ln(1-\alpha) - \alpha \ln \alpha - g \ln g - (1-g) \ln(1-g) + \\ & + \frac{\alpha}{2} \ln \left(2r\sqrt{K} - (1-2\gamma)K + 1 \right) - \frac{1}{2} \ln K + \\ & + \frac{1}{2Z} \left(-2r\sqrt{K} + (1-2\gamma)K - 1 \right) - \frac{1}{2Z} a_1^2 + (2g-1) \frac{-a_1 - \sqrt{K}s}{Z} + \\ & + \frac{\gamma\sqrt{K}r}{Z} + (2g-1)\gamma\sqrt{K}\frac{s}{Z} + \frac{\gamma^2 K}{2Z} + \\ & + g \left[\ln \Phi \left(\frac{r+s}{\sqrt{Z}} \right) + \frac{(r+s)^2}{2Z} \right] + (1-g) \left[\ln \Phi \left(\frac{r-s}{\sqrt{Z}} \right) + \frac{(r-s)^2}{2Z} \right] \end{aligned}$$

Durch die Substitution wurden also die Ausdrücke in a_1 isoliert und die Argumente der Φ -Funktionen auf symmetrische Form gebracht.

$$\implies \frac{\partial f}{\partial a_1} = 0 \implies \frac{1}{Z} (2g-1+a_1) = 0 \implies a_1 = 1-2g \quad (\text{C.12})$$

Setzt man a_1 in Gl.(C.11) ein, so ergibt sich nach Umformungen

$$\langle\langle \mathcal{N}(\alpha, g, \gamma, \varepsilon_{max}) \rangle\rangle \propto e^{N f(\alpha, g, \gamma, \varepsilon_{max})},$$

wobei $f(\alpha, g, \gamma, \varepsilon_{max})$ der Sattelpunktwert über die Variablen K, r, Z, s von dem folgenden Ausdruck ist.

$$\begin{aligned} f(\alpha, g, \gamma, \varepsilon_{max}, K, r, Z, s) & \\ = & - (1-\alpha) \ln(1-\alpha) - \alpha \ln \alpha - g \ln g - (1-g) \ln(1-g) + \\ & + \frac{\alpha}{2} \ln \left(2r\sqrt{K} - (1-2\gamma)K + 1 \right) - \frac{1}{2} \ln K + \\ & + \frac{1}{2Z} \left(r^2 - 2r\sqrt{K}(1-\gamma) + (1-\gamma)^2 K - 1 + s^2 \right) + \\ & + \frac{(2g-1)^2}{2Z} + (2g-1) \frac{2rs - 2\sqrt{K}s(1-\gamma)}{2Z} + \\ & + g \ln \Phi \left(\frac{r+s}{\sqrt{Z}} \right) + (1-g) \ln \Phi \left(\frac{r-s}{\sqrt{Z}} \right). \quad (\text{C.13}) \end{aligned}$$

C.6 Herleitung der Sattelpunktgleichungen zur Berechnung von $\langle\langle \mathcal{N}(\alpha, \mathbf{g}, \gamma, \varepsilon_{\max}) \rangle\rangle$

Jetzt werden die vier partiellen Ableitungen von $f(\alpha, g, \gamma, \varepsilon_{\max}, K, r, Z, s)$ aus Gl.(C.13) gebildet. Man erhält

$$\frac{\partial f}{\partial K} = 0 \implies$$

$$0 = \frac{\alpha}{2} \cdot \frac{\frac{r}{\sqrt{K}} - (1 - 2\gamma)}{2r\sqrt{K} - (1 - 2\gamma)K + 1} - \frac{1}{2K} + \frac{1}{2Z} \left(-\frac{r}{\sqrt{K}}(1 - \gamma) + (1 - \gamma)^2 \right) + \frac{1}{2Z} \cdot (2g - 1) \left(-\frac{s}{\sqrt{K}}(1 - \gamma) \right). \quad (\text{C.14})$$

$$\frac{\partial f}{\partial r} = 0 \implies$$

$$0 = \frac{\alpha}{2} \cdot \frac{2\sqrt{K}}{2r\sqrt{K} - (1 - 2\gamma)K + 1} + \frac{1}{2Z} (2r - 2\sqrt{K}(1 - \gamma)) + \frac{1}{2Z} (2g - 1) \cdot 2s + \frac{g}{\sqrt{Z}} \cdot \frac{\Phi'(t_1)}{\Phi(t_1)} + \frac{1 - g}{\sqrt{Z}} \cdot \frac{\Phi'(t_2)}{\Phi(t_2)} \quad (\text{C.15})$$

$$\frac{\partial f}{\partial s} = 0 \implies$$

$$0 = \frac{s}{Z} + \frac{1}{2Z} \cdot (2g - 1) \cdot (2r - 2\sqrt{K}(1 - \gamma)) + g \cdot \frac{\Phi'(t_1)}{\Phi(t_1)} \cdot \frac{1}{\sqrt{Z}} + (1 - g) \cdot \frac{\Phi'(t_2)}{\Phi(t_2)} \cdot \left(-\frac{1}{\sqrt{Z}} \right) \quad (\text{C.16})$$

$$\frac{\partial f}{\partial Z} = 0 \implies$$

$$\begin{aligned}
0 = & -\frac{1}{2Z^2} \cdot \left(r^2 - 2r\sqrt{K}(1-\gamma) + (1-\gamma)^2K - 1 + s^2 \right) - \frac{1}{2Z^2} \cdot (2g-1)^2 + \\
& -\frac{1}{2Z^2} \cdot (2g-1) \left(2rs - 2\sqrt{K}s(1-\gamma) \right) + \\
& -\frac{\Phi'(t_1)}{\Phi(t_1)} \cdot \frac{g(r+s)}{2Z\sqrt{Z}} - \frac{\Phi'(t_2)}{\Phi(t_2)} \cdot \frac{(1-g)(r-s)}{2Z\sqrt{Z}} . \tag{C.17}
\end{aligned}$$

Ziel der folgenden Rechnung ist, r und K durch zwei verbleibende Sattelpunktvariablen zu ersetzen.

Die Gleichungen (C.18) := (C.15) + (C.16) und

(C.19) := (C.15) - (C.16) erhalten dann jeweils nur eine Phi-Funktion:

$$\begin{aligned}
0 = & \alpha \cdot \frac{\sqrt{K}}{2r\sqrt{K} - (1-2\gamma)K + 1} + \frac{1}{2Z} \left(2r + 2s - 2\sqrt{K}(1-\gamma) \right) + \\
& + \frac{2g-1}{2Z} \cdot \left(2r + 2s - 2\sqrt{K}(1-\gamma) \right) + \frac{2g}{\sqrt{Z}} \cdot \frac{\Phi'(t_1)}{\Phi(t_1)} \tag{C.18}
\end{aligned}$$

$$\begin{aligned}
0 = & \alpha \cdot \frac{\sqrt{K}}{2r\sqrt{K} - (1-2\gamma)K + 1} + \frac{1}{2Z} \left(2r - 2s - 2\sqrt{K}(1-\gamma) \right) + \\
& + \frac{2g-1}{2Z} \cdot \left(-2r + 2s + 2\sqrt{K}(1-\gamma) \right) + \frac{2(1-g)}{\sqrt{Z}} \cdot \frac{\Phi'(t_2)}{\Phi(t_2)} \tag{C.19}
\end{aligned}$$

enthalten dann jeweils nur eine Phi-Funktion. Man verrechnet mit Gl.(C.17) und erhält eine weitere Gleichung ohne Phi-Funktion:

$$(C.20) := Z \cdot \left((C.17) \cdot 2Z + (C.18) \cdot \frac{r+s}{2} + (C.19) \cdot \frac{r-s}{2} \right)$$

$$\begin{aligned}
0 = & Z\alpha \frac{r\sqrt{K}}{2r\sqrt{K} - (1-2\gamma)K + 1} + r\sqrt{K}(1-\gamma) + \\
& + s\sqrt{K}(1-\gamma)(2g-1) - (1-\gamma)^2K + 1 - (2g-1)^2 . \tag{C.20}
\end{aligned}$$

$$(C.21) := 2ZK \cdot (C.14) + (C.20)$$

erlaubt dann r durch K und Z auszudrücken

$$0 = Z\alpha \frac{2r\sqrt{K} - (1-2\gamma)K}{2r\sqrt{K} - (1-2\gamma)K + 1} - Z + 1 - (2g-1)^2 . \tag{C.21}$$

Die Gleichung wird nach der im folgenden definierten Hilfsvariablen m aufgelöst. m hängt *nur* von Z ab.

$$2r\sqrt{K} - (1 - 2\gamma)K = \frac{Z - 1 + (2g - 1)^2}{1 - Z(1 - \alpha) - (2g - 1)^2} =: m \quad (\text{C.22})$$

$m + 1$ hat dann folgende einfache Form

$$m + 1 = \frac{\alpha Z}{1 - Z(1 - \alpha) - (2g - 1)^2}. \quad (\text{C.23})$$

Jetzt muß man noch K in Abhängigkeit von Z und s ausrechnen. Man setzt $m = m(Z)$ und $r \cdot \sqrt{K}$ in die Gleichung (C.14) $\cdot 2ZK$ ein und erhält nach Umformen

$$K = \frac{Z + 1 - (2g - 1)^2 + (1 - \gamma)m + 2s\sqrt{K}(2g - 1)(1 - \gamma)}{(1 - 2\gamma) \cdot (Z(1 - \alpha) + (2g - 1)^2) + \gamma} \quad (\text{C.24})$$

Der hier beschrittene Weg vermeidet das Auflösen der Gl.(C.24) nach \sqrt{K} : Man ersetzt s mit Hilfe von

$$U = s\sqrt{K}. \quad (\text{C.25})$$

Dies ist möglich, weil m nicht von s , sondern nur von Z abhängt. U und Z sind dann die verbleibenden Sattelpunktvariablen. r und K sind Funktionen von U und Z . Aus den Gl.(C.18) und Gl.(C.19) gewinnt man im folgenden Iterationsgleichungen für Z bzw. für U .

(C.18) $\cdot \frac{Z}{\sqrt{K}}$ ergibt mit Hilfe von Gl.(C.23)

$$Z = \frac{1}{1 - \alpha} \cdot \left[1 - (2g - 1)^2 + 2g \left(\frac{m}{2K} - \frac{1}{2} + \frac{U}{K} \right) + 2g \sqrt{\frac{Z}{K}} \cdot \frac{\Phi'(t_1)}{\Phi(t_1)} \right]. \quad (\text{C.26})$$

Dabei ist

$$t_1 = \frac{r + \frac{U}{\sqrt{K}}}{\sqrt{Z}}.$$

(C.19) $\cdot \frac{Z\sqrt{K}}{2(1-g)}$ ergibt mit Hilfe von Gl.(C.23) und Gl.(C.25)

$$U = \frac{K}{2(1-g)} \left(1 - (2g-1)^2 - Z(1-\alpha) \right) + \frac{m-K}{2} + \sqrt{ZK} \cdot \frac{\Phi'(t_2)}{\Phi(t_2)}. \quad (\text{C.27})$$

Das Argument der Φ -Funktion lautet hier

$$t_2 = \frac{r - \frac{U}{\sqrt{K}}}{\sqrt{Z}}.$$

Bei der Lösung von Gl.(C.26) und Gl.(C.27) durch Iteration bemerkt man, daß Z mit g skaliert. Man führt deshalb die neue Variable V ein:

$$V := \frac{Z}{4g(1-g)}. \quad (\text{C.28})$$

Mit $4g(1-g) = 1 - (2g-1)^2$ folgt dann

$$m = \frac{V4g(1-g) - 4g(1-g)}{4g(1-g) - 4g(1-g)V(1-\alpha)} = \frac{V-1}{1-V(1-\alpha)}. \quad (\text{C.29})$$

Damit ist der letzte Schritt getan, und wir erhalten nach Einsetzen von V die Gleichungen in Abschnitt 4.2.2.

Zwei andere Iterationsgleichungen gewinnt man durch Umformen der Gleichungen (4.27) und (4.28):

$$V(1-\alpha) = \frac{m}{2K} + \frac{1}{2} + \frac{1}{2} \cdot \sqrt{\frac{Z}{K}} \left(\frac{\Phi'(t_1)}{\Phi(t_1)} + \frac{\Phi'(t_2)}{\Phi(t_2)} \right). \quad (\text{C.30})$$

$$U = (2g-1) \cdot K \cdot (1-V(1-\alpha)) - \frac{1}{2} \sqrt{ZK} \left(\frac{\Phi'(t_1)}{\Phi(t_1)} - \frac{\Phi'(t_2)}{\Phi(t_2)} \right). \quad (\text{C.31})$$

Bei der Lösung auf dem Rechner ist Gl.(C.31) für g in der Nähe von $\frac{1}{2}$ der Gl.(4.28) vorzuziehen. Für kleine g ist Gl.(4.28) hingegen vorteilhaft. Für $\alpha \rightarrow 1$ beobachtet man zudem, daß U und V mit $(1-\alpha)$ skalieren. Es ist deshalb in diesem Bereich ratsam, die Ausdrücke $U(1-\alpha)$ und $V(1-\alpha)$ zu iterieren. Für die theoretischen Untersuchungen werden später jedoch nur die beiden Iterationsgleichungen im Haupttext verwendet.

C.7 Näherungen für $\langle\langle\mathcal{N}(\alpha, \mathbf{g} = \frac{1}{2}, \gamma, \varepsilon_{\max})\rangle\rangle$

Wir führen den selbstkonsistenten Taylorreihenansatz an dieser Stelle für $f(\alpha, g = \frac{1}{2}, \gamma = \alpha, \varepsilon_{\max})$, also für den in Abschnitt 4.3.2 erwähnten Fall 2 durch. Das Vorgehen im Fall 1 ist ganz analog. Auch die Rechnung zu Abschnitt 4.2.3 ist ganz ähnlich und wird in dieser Arbeit nicht explizit aufgeführt.

Sei also $\gamma = \alpha, g = \frac{1}{2}, \varepsilon = \varepsilon_{\max}$. Betrachtet seien Gl.(4.35) bis Gl.(4.40). Man macht nun den Taylorreihen-Ansatz

$$Z = \frac{2}{\pi} + \sum_{n=1}^{\infty} Z_n \alpha^n = \frac{2}{\pi} + Z_1 \alpha + \mathcal{O}(\alpha^2) \quad (\text{C.32})$$

für $\alpha \rightarrow 0$. Nach Einsetzen dieses Ansatzes in Gl.(4.36) bis Gl.(4.39) erhält man mit $\gamma = \alpha$ für die Hilfsvariablen

$$m = -1 + \frac{1}{\frac{\pi}{2} - 1} \alpha + \mathcal{O}(\alpha^2) \quad (\text{C.33})$$

$$K = 1 + \alpha \left(3 + \frac{1}{1 - \frac{2}{\pi}} \right) + \mathcal{O}(\alpha^2) \quad (\text{C.34})$$

$$r = \frac{\alpha}{1 - \frac{2}{\pi}} + \mathcal{O}(\alpha^2) \quad (\text{C.35})$$

$$t = \sqrt{\frac{\pi}{2}} \cdot \frac{\alpha}{1 - \frac{2}{\pi}} + \mathcal{O}(\alpha^2) . \quad (\text{C.36})$$

Dabei wurden die Näherungen

$$\sqrt{1+x} = 1 + \frac{1}{2}x + \mathcal{O}(x^2)$$

und

$$\frac{1}{1-x} = 1 + x + \mathcal{O}(x^2)$$

für kleine x benutzt.

Es gilt also $t = \mathcal{O}(\alpha)$. Man kann folglich Gl.(A.4) für $\Phi(t \rightarrow 0)$ verwenden. Wir erhalten

$$\Phi(t) = \frac{1}{2} \left(1 + \sqrt{\frac{2}{\pi}} t + \mathcal{O}(t^3) \right) = \frac{1}{2} \left(1 + 1 \cdot \frac{\alpha}{1 - \frac{2}{\pi}} \right) + \mathcal{O}(\alpha^2) \quad (\text{C.37})$$

und

$$\sqrt{\frac{Z}{K}} = \sqrt{\frac{2}{\pi} \left(1 + \alpha \cdot \frac{\pi}{4} \left(Z_1 - \frac{2}{\pi} \left(3 + \frac{1}{1 - \frac{2}{\pi}} \right) \right) \right)} + \mathcal{O}(\alpha^2). \quad (\text{C.38})$$

Man setzt schließlich Gl.(C.32) bis Gl.(C.38) in die Sattelpunktgleichung (4.40) ein und erhält nach Entwicklung der Exponentialfunktion $\Phi'(t)$ bis zur Ordnung α^2 :

$$\begin{aligned} \left(\frac{2}{\pi} + Z_1 \alpha \right) (1 - \alpha) &= \frac{1}{\pi} \left(1 + \alpha \cdot \frac{\pi}{4} \left(Z_1 - \frac{2}{\pi} \left(3 + \frac{1}{1 - \frac{2}{\pi}} \right) \right) \right) \\ &\cdot 2 \left(1 - \frac{\alpha}{1 - \frac{2}{\pi}} \right) + \frac{\alpha}{1 - \frac{2}{\pi}} + \alpha + \mathcal{O}(\alpha^2) \end{aligned} \quad (\text{C.39})$$

Koeffizientenvergleich in α^0 liefert dann die Identität $\frac{2}{\pi} = \frac{2}{\pi}$.

Koeffizientenvergleich in α^1 liefert Z_1 , und es gilt

$$Z = \frac{2}{\pi} + \alpha \left(3 - \frac{2}{\pi} + \frac{1 - \frac{4}{\pi}}{1 - \frac{2}{\pi}} \right) + \mathcal{O}(\alpha^2). \quad (\text{C.40})$$

Wir entwickeln somit f aus Gl.(4.35) bis zu $\mathcal{O}(\alpha^2)$ und erhalten unter Verwendung von $\ln(1+x) = x + \mathcal{O}(x^2)$:

$$f\left(\alpha, g = \frac{1}{2}, \gamma = \alpha, \varepsilon_{max}\right) = -\frac{\alpha}{2} \ln \alpha + \frac{\alpha}{2} \left(\ln\left(\frac{2}{\pi}\right) - 1 - \ln\left(1 - \frac{2}{\pi}\right) \right) + \mathcal{O}(\alpha^2)$$

C.8 Berechnung von $\langle\langle \mathcal{N}(\alpha, \mathbf{g} = \frac{1}{2}, \gamma, \varepsilon) \rangle\rangle$

Wie in Abschnitt 4.4 beschrieben, erhalten wir das allgemeine Endergebnis (4.18) mit $g = \frac{1}{2}, a_1 = l_1 = 0$:

$$\langle\langle \mathcal{N}(\alpha, g = \frac{1}{2}, \gamma, \varepsilon) \rangle\rangle \propto \exp\left(Nf\left(\alpha, g = \frac{1}{2}, \gamma, \varepsilon\right)\right),$$

wobei f der Sattelpunktwert der folgenden Funktion über die Sattelpunktvariablen X, Y, Z, w ist:

$$\begin{aligned} f\left(\alpha, g = \frac{1}{2}, \gamma, \varepsilon, X, Y, Z, w\right) &= \\ &- (1 - \alpha) \ln(1 - \alpha) - \alpha \ln \alpha + \ln 2 + \ln \Phi(t) + \frac{1}{8} w^2 \cdot \frac{Z}{K} + \\ &+ w \cdot \frac{K + Y - 1}{2K} + \frac{X}{2K} - w\varepsilon - \frac{1}{2} \ln K + \frac{\alpha}{2} \ln(Y^2 - XZ). \end{aligned} \quad (\text{C.41})$$

Dabei ist K die Abkürzung $K = (1 - Y)^2 - XZ$, und t ist gegeben durch

$$t = \frac{(1 - \gamma)K + Y - 1 + \frac{1}{2}wZ}{\sqrt{ZK}}. \quad (\text{C.42})$$

Jetzt wird die Variablensubstitution

$$(X, Y, Z, w) \rightarrow (K, Y', Z, w)$$

durchgeführt, um w aus dem Argument t der Phi-Funktion zu entfernen (Methode 2 in Abschnitt 4.1.2).

$$\begin{aligned} K &= (1 - Y)^2 - XZ \\ Y' &= Y + \frac{1}{2}wZ \\ Z &= Z \\ w &= w \end{aligned}$$

Nach Einsetzen in Gl.(C.41) und Gl.(C.42) erhält man

$$\begin{aligned} f\left(\alpha, g = \frac{1}{2}, \gamma, \varepsilon, K, Y', Z, w\right) = \\ -(1 - \alpha) \ln(1 - \alpha) - \alpha \ln \alpha + \ln 2 + \ln \Phi(t) + w\left(\frac{1}{2} - \varepsilon\right) + \\ + \frac{\left((Y' - 1)^2 - K\right)}{2ZK} - \frac{1}{2} \ln K + \frac{\alpha}{2} \ln(K + 2Y' - 1 - wZ); \end{aligned} \quad (\text{C.43})$$

$$t = \frac{(1 - \gamma)K + Y' - 1}{\sqrt{ZK}}. \quad (\text{C.44})$$

Durch Ableiten nach w kann man jetzt w durch die anderen Variablen ersetzen (Methode 1 in Abschnitt 4.1.2):

$$\frac{\partial f}{\partial w} = 0 \implies w = \frac{K + 2Y' - 1}{Z} - \frac{\alpha}{1 - 2\varepsilon}. \quad (\text{C.45})$$

Nach Einsetzen erhält man

$$\begin{aligned} f\left(\alpha, g = \frac{1}{2}, \gamma, \varepsilon, K, Y', Z\right) = \\ -(1 - \alpha) \ln(1 - \alpha) - \alpha \ln \alpha + \ln 2 + \ln \Phi(t) + \frac{\frac{1}{2} - \varepsilon}{Z}(K + 2Y' - 1) + \\ - \frac{\alpha}{2} + \frac{\left((Y' - 1)^2 - K\right)}{2ZK} - \frac{1}{2} \ln K + \frac{\alpha}{2} \ln\left(\frac{\alpha Z}{1 - 2\varepsilon}\right). \end{aligned} \quad (\text{C.46})$$

Dieses Zwischenergebnis kann man leicht testen, wenn man den Fall $\varepsilon = \varepsilon_{max}$ betrachtet. Dort muß $w = 0$ sein (siehe Bemerkungen in Abschnitt 4.2.1) und man erhält aus Gl.(C.45)

$$\frac{K + 2Y' - 1}{Z} = \frac{\alpha}{1 - 2\varepsilon_{max}} .$$

Nach Einsetzen in Gl.(C.46) erhält man dann mit

$$r = \frac{(1 - \gamma)K + Y' - 1}{\sqrt{K}}$$

genau das Ergebnis (4.35) für $f(\alpha, g = \frac{1}{2}, \gamma, \varepsilon_{max})$.

Schließlich führen wir die Substitution

$$(K, Y', Z) \rightarrow (P, t, Z)$$

durch:

$$P = \frac{K}{Z}, \quad t = \frac{(1 - \gamma)K + Y' - 1}{\sqrt{KZ}}, \quad Z = Z . \quad (C.47)$$

Man erhält dann Gl.(4.54) mit P, t und Z als Sattelpunktvariablen. Die drei Sattelpunktgleichungen lauten

$$\frac{\partial f}{\partial Z} = 0 \implies \frac{\varepsilon}{Z^2} - \frac{1 - \alpha}{2Z} = 0 \quad (C.48)$$

$$\frac{\partial f}{\partial P} = 0 \implies t \frac{-2\varepsilon + \gamma}{2\sqrt{P}} + \frac{1}{2}(1 - \gamma)^2 - \left(\frac{1}{2} - \varepsilon\right) \cdot (1 - 2\gamma) - \frac{1}{2P} = 0 \quad (C.49)$$

$$\frac{\partial f}{\partial t} = 0 \implies \frac{\Phi'(t)}{\Phi(t)} + \sqrt{P}(-2\varepsilon + \gamma) + t = 0 . \quad (C.50)$$

Aus Gl.(C.48) ergibt sich sofort Gl.(4.55) für Z . Die Einführung der Iterationsvariablen

$$S = \frac{t\sqrt{P}}{2}(\gamma - 2\varepsilon) \quad (C.51)$$

vereinfacht noch die beiden verbleibenden Gleichungen. Aus Gl.(C.49) erhält man dann Gl.(4.56) für P . Gl.(4.57) für t folgt dann aus Gl.(C.51). Die Iterationsgleichung (4.58) ergibt sich schließlich aus Gl.(C.50).

Zum Schluß sei noch darauf hingewiesen, daß die Parameter γ und ε nicht unabhängig sind. Es gilt

$$\varepsilon = \frac{1}{2N} \sum_{i=1}^N \lambda_i > \frac{\gamma}{2},$$

also ist stets $2\varepsilon - \gamma > 0$. Setzt man nicht erlaubte Kombinationen von Parametern ε und γ ein, so haben die sich ergebenden Sattelpunktgleichungen keine Lösung. So erhält man z.B. für $2\varepsilon = \gamma$ die Gleichung

$$-t = \frac{\Phi'(t)}{\Phi(t)},$$

die nur für $t \rightarrow -\infty$ erfüllt ist.

Anhang D

Ein Berechnungsverfahren für die Zahl der Zweierzyklen

Berechnet werden soll die Zahl der Zweierzyklen für ein Spin-System mit Kopplungen J_{ij} . Das Spin-System soll der parallelen $\tau = 0$ -Glauber-Dynamik unterworfen werden (siehe Abschnitt 2.1.2), die alle Spins nach der Regel

$$S_i(t + \Delta t) = \text{sign}(h_i(t)) = \text{sign}\left(\sum_{j=1}^N J_{ij} S_j(t)\right) \quad \forall i$$

auffrischt. Ein Zweier-Zyklus gehorcht für alle Zeiten t der Bedingung

$$S_i(t + \Delta t) = S_i(t - \Delta t) \quad \forall i .$$

Zu einem beliebigen Zeitpunkt gilt somit

$$\lambda_i = S_i \sum_{j=1}^N J_{ij} \text{sign}\left(\sum_{k=1}^N J_{jk} S_k\right) > 0 \quad \forall i . \quad (\text{D.1})$$

Z. Li und A.V.M. Herz haben u.a. nachgewiesen [LiHe90], daß man das äquivalente Problem

$$\lambda_i = S_i \sum_{j=1}^N J_{ij} \sigma_j > 0 \quad \forall i \quad (\text{D.2})$$


lösen kann. Die Verbannung der sign-Funktion wird dabei durch die Einführung eines „großen Systems“ mit

$$S_i := S_i(t + \Delta t), \quad \sigma_i := S_i(t) = \text{sign}(h_i(t + \Delta t)) \quad \forall i$$

erkaufte. \underline{S} und $\underline{\sigma}$ sind unabhängig und setzen sich zum Zustandsvektor $(\underline{S}, \underline{\sigma})$ des großen Systems zusammen.

Für die Zahl der Zweierzyklen gilt schließlich

$$\mathcal{N} = \sum_{\{\underline{S}\}} \sum_{\{\underline{\sigma}\}} \prod_{i=1}^N \Theta \left(S_i \sum_{j=1}^N J_{ij} \sigma_j \right) \cdot \Theta \left(\sigma_i \sum_{j=1}^N J_{ij} S_j \right) \quad (\text{D.3})$$

Jetzt kann man wieder die δ -Funktions-Darstellung für die Θ -Funktion einführen und anschließend über die Muster mitteln. 

Literaturverzeichnis

- [AbSt65] M. Abramovitz und I. Stegun: *Handbook of Mathematical Functions*, National Bureau of Standards, Washington 1965.
- [Al72] A. Albert: *Regression and the Moore–Penrose pseudoinverse*, Academic Press, New York, 1972.
- [Am+85a] D. J. Amit, H. Gutfreund und H. Sompolinsky: *Spin–Glass Models of Neural Networks*, Phys.Rev.,**A32** (1985), 1007.
- [Am+85b] D. J. Amit, H. Gutfreund und H. Sompolinsky: *Storing Infinite Numbers of Patterns in a Spin–Glass Model of Neural Networks*, Phys.Rev.Lett.,**55** (1985), 1530.
- [Am+87] D. J. Amit, H. Gutfreund und H. Sompolinsky: *Statistical Mechanics of Neural Networks near Saturation*, Annals of Physics,**173** (1987), 30.
- [Am89] D. J. Amit: *Modelling brain function: the world of attractor neural networks*, Cambridge University Press, Cambridge, 1989.
- [AnBi89] J.K. Anlauf und M. Biehl: *The AdaTron: An Adaptive Perceptron Algorithm*, Europhys.Lett.,**10** (1989), 687.
- [AsMe76] N. W. Ashcroft und N. D. Mermin: *Solid State Physics*, W. B. Saunders, Philadelphia, 1976.
- [Be+90] K. W. Berryman *et al.*: *Convergence of an iterative neural network learning algorithm for linearly dependent patterns*, J.Phys.,**A23** (1990), L223.

- [BiYo86] K. Binder und A.P. Young: *Spin glasses: Experimental facts, theoretical concepts and open questions*, Rev.Mod.Phys.,**58** (1986), 801.
- [BrMo80] A.J. Bray und M.A. Moore: *Metastable states in spin glasses*, J.Phys.,**A22** (1980), L469.
- [BrMo81] A. J. Bray und M. A. Moore: *Metastable states in spin glasses with short-ranged interactions*, J.Phys.,**C14** (1981), 1313.
- [DHS90] E. Domany, J. L. van Hemmen und K. Schulten eds.: *Physics of Neural Networks*, Springer, Heidelberg, 1990.
- [DiOp87] S. Diederich und M. Opper: *Learning of Correlated Patterns in Spin-Glass Networks by Local Learning Rules*, Phys.Rev.Lett.,**58** (1987), 949.
- [DoGr76] C. Domb und M. S. Green, eds.: *Phase Transitions and Critical Phenomena*, Band 5b, Academic Press, London, 1976.
- [Eck+90] R. Eckmiller, G. Hartmann, G. Hauske, eds.: *Parallel Processing in Neural Systems and Computers*, North Holland, Amsterdam, 1990.
- [EdAn75] S.F. Edwards und P.W. Anderson: *Theory of spin glasses*, J.Phys.,**F5** (1975), 965.
- [EnHe84] A. C. D. van Enter und J. L. van Hemmen: *Statistical Mechanics Formalism for Spin-Glasses*, Phys.Rev.,**A29** (1984), 355.
- [Fe72] R. P. Feynman: *Statistical Mechanics*, W. A. Benjamin, Reading, 1972.
- [Fl87] R. Fletcher: *Practical Methods of Optimization*, Wiley & Sons, Chichester, 1987.
- [FoKo87] J. F. Fontanari und R. Köberle: *Information storage and retrieval in synchronous neural networks*, Phys.Rev.,**A36** (1987), 2475.
- [Ga86] E. Gardner: *Structure of metastable states in the Hopfield model*, J.Phys.,**A19** (1986), L1047.

- [Ga88] E. Gardner: *The Space of Interactions in Neural Network Models*, J.Phys., **A21** (1988), 257.
- [Ge90] T. Gezti: *Physical Models of Neural Networks*, World Scientific, Singapur 1990.
- [Gl63] R. J. Glauber: *Time-Dependent Statistics of the Ising Model*, J.Math.Phys., **4** (1963), 294.
- [Gu+88] H. Gutfreund, J. D. Reger und A. P. Young: *The nature of attractors in an asymmetric spin glass with deterministic dynamics*, J.Phys., **A21** (1988), 2775.
- [He49] D. O. Hebb: *The Organisation of Behavior*, Wiley, New York, 1949.
- [HeKi87] R. D. Henkel und W. Kinzel: *Metastable states of the SK model of spin glasses*, J.Phys., **A20** (1987), L727.
- [HeKu90] J. L. van Hemmen und R. Kühn: *Collective Phenomena in Neural Networks*, Heidelberg Preprint Nr. 553, 1990.
- [HeOp90] R. D. Henkel und M. Opper: *Parallel Dynamics of the Neural Network with Pseudoinverse Coupling Matrix*, Gießen Preprint, 1990.
- [Hi75] K. Hinderer: *Grundbegriffe der Wahrscheinlichkeitstheorie*, Springer, Berlin, 1975.
- [Ho82] J.J. Hopfield: *Neural networks and physical systems with emergent collective computational abilities*, Proc. Natl. Acad. Sci. USA, **79** (1982), 2554.
- [Hu59] J. Hubbard: *Calculation of Partition Functions*, Phys.Rev.Lett., **3** (1959), 77.
- [Huang] K. Huang: *Statistical Mechanics*, Wiley & Sons, New York, 1963 und 1987.
- [INNS89] International Neural Network Society (Hrsg.): *International Joint Conference on Neural Networks 1989*, IEEE TAB Neural Network Committee, 1989.

- [Is25] E. Ising: *Beitrag zur Theorie des Ferromagnetismus*, Z.Phys., **31** (1925), 253.
- [KaSo87] I. Kanter und H. Sompolinsky: *Associative recall of memory without errors*, Phys.Rev., **A35** (1987), 380.
- [KFA85] Kernforschungsanlage Jülich, Institut für Festkörperforschung (Hrsg.) : *Ferienkurs '85, Unordnungsphänomene in Festkörpern*, (Skriptum), KFA Jülich, 1985.
- [Ki85] W. Kinzel: *Learning and Pattern Recognition in Spin Glass Models*, Z.Phys., **B60** (1985), 205.
- [Ki Vorl] W. Kinzel: *Spingläser und Neuronale Netzwerke*, Vorlesung, Justus-Liebig-Universität Gießen, 1988.
- [Ko88] T. Kohonen: *Self-Organisation and Associative Memory*, Springer, Berlin, 1984 und 1988.
- [KrKo90] J. Krätzschar und G. A. Kohring: *Retrieval dynamics of neural networks constructed from local and nonlocal learning rules*, J.Phys.(Paris), **51** (1990), 223.
- [Li74] W.A. Little: *The Existence of Persistent States in the Brain*, Math.Biosc., **19** (1974), 101.
- [LiHe90] Z. Li und A. V. M. Herz: *Lyapunov functional for neural networks with delayed interactions*, Preprint.
- [Me+53] N. Metropolis *et al.*: *Equation of state calculations by fast computing machines*, J. Chem. Phys. **21** (1953) 1087.
- [Me+87] M. Mezard, G. Parisi und M.A. Virasoro: *Spin Glass Theory and beyond*, World Scientific, Singapur, 1987.
- [MiPa69] M.L. Minsky und S.A. Papert: *Perceptrons*, MIT Press, Cambridge, 1969 und 1988.
- [NeOr88] J. W. Negele und H. Orland: *Quantum Many-Particle-Systems*, Addison-Wesley, Redwood City, 1988.

- [Op88] M. Opper: *Learning times of neural networks: Exact solution for a Perceptron algorithm*, Phys.Rev.,**A38** (1988), 3824.
- [Pa80] G. Parisi: *A sequence of approximated solutions to the S–K model for spin glasses*, J.Phys.,**A13** (1980), L115.
- [Pe84] P. Peretto: *Collective properties of neural networks: A statistical physics approach*, Biol. Cybern., **50** (1984), 51.
- [Pe+85] L. Personnaz, I. Guyon, G. Dreyfus: *Information storage and retrieval in spin–glass neural networks*, J. Phys. (Paris) **46** (1985), L359
- [Pe+86] L. Personnaz, I. Guyon und G. Dreyfus: *Collective computational properties of neural networks: New learning mechanisms*, Phys.Rev.,**A34** (1986), 4217.
- [Reif] F. Reif: *Fundamentals of statistical and thermal physics*, McGraw–Hill, Singapur, 1965.
- [Ri+90] H. Ritter, T. Martinetz und K. Schulten: *Neuronale Netze*, Addison–Wesley, Bonn, 1990.
- [Ro81] S. A. Roberts: *Metastable states and innocent replica theory in an Ising spin glass*, J.Phys.,**C14** (1981), 3015.
- [Ru86] D.E. Rumelhart und J.L. McClelland: *Parallel distributed processing*, Bradford Books, Cambridge und London, 1986.
- [Schwarz] H. R. Schwarz: *Numerische Mathematik*, Teubner, Stuttgart, 1986.
- [SK75] D. Sherrington und S. Kirkpatrick: *Solvable Model of a spin glass*,Phys.Rev.Lett.,**35** (1975), 1792.
- [SK78] D. Sherrington und S. Kirkpatrick: *Infinite–ranged models of spin–glasses*, Phys.Rev.,**B17** (1978), 4384.
- [TaEd80] F. Tanaka und F. S. Edwards: *Analytic theory of ground state properties of a spin glass: I. Ising spin glass*, J.Phys.,**F10** (1980), 2769.

- [TAP77] D. J. Thouless, P. W. Anderson und R. G. Palmer: *Solution of 'Solvable model of a spin glass'*, Philosophical Magazine, **35** (1977) 3, 593.
- [To77] G. Toulouse: *Theory of the frustration effect in spin glasses: I.*, Comm. Phys., **2** (1977), 115.
- [VaTo77] J. Vannimenus und G. Toulouse: *Theory of the frustration effect: II. Ising spins on a square lattice*, J.Phys., **C10** (1977), L537.
- [Wa71] W. Walter: *Analysis 2*, Vorlesungsskript, Universität Karlsruhe, 1971.
- [ZuFa84] R. Zurmühl, S. Falk: *Matrizen und ihre Anwendungen*, Springer, Berlin, 1984.

Danksagung

An dieser Stelle möchte ich mich bei Prof. Dr. W. Kinzel für die interessante Themenstellung und die gute Betreuung bedanken. Mein besonderer Dank gilt Dr. J. K. Anlauf für seine ständige Diskussionsbereitschaft und seinen ansteckenden Enthusiasmus auf dem Gebiet der Neuronalen Netzwerke. Bei Dr. M. Opper bedanke ich mich ebenfalls für viele Ratschläge, die mir bei der Einarbeitung halfen und mich mit einigen Rechenverfahren vertraut machten.

P. Spitzner, R. D. Henkel und H. Eißfeller danke ich für ihre Unterstützung bei Hard- und Softwareproblemen. P. Spitzner und R. D. Henkel haben mir ebenfalls dankenswerterweise einige ihrer Arbeitsergebnisse zur Verfügung gestellt (Teil der Abbildung 5.17 bzw. Abbildung 5.2).

„Last but not least“ bedanke ich mich bei meinen Zimmergenossen A. Wendemuth und M. Biehl für ihre Diskussionsbereitschaft und ihre mentale Unterstützung.

Peter Kuhlmann



6332 Ehringshausen

Hiermit versichere ich, die vorliegende Arbeit selbständig und unter ausschließlicher Verwendung der angeführten Hilfsmittel und Quellen angefertigt zu haben.

Gießen, den