

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE FÍSICA

REDES DE NEURÔNIOS COM INTERAÇÕES
SINÁPTICAS HIERÁRQUICAS *

Marco Aurélio Pires Idiart

Tese realizada sob a orientação da
Prof. Dr. Alba G. R. de Theumann e
apresentada ao Instituto de Física da
UFRGS, em preenchimento parcial dos
requisitos para a obtenção do grau de
Doutor em Ciências.

Porto Alegre
1991

*Trabalho parcialmente financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e pela Coordenação do Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brasil.

Agradecimentos

Agradeço a Prof. Alba Theumann pela proposição do tema, por ter possibilitado meu aprimoramento como pesquisador e pela sua inestimável amizade.

Ao Prof. Walter Theumann pela introdução ao fascinante assunto de Redes de Neurônios.

Aos meus companheiros que ainda hoje me inspiram, Angela, Gilberto, Pedro e Zepka.

Aos meus amigos e colegas que de uma forma ou outra propiciaram o meu crescimento como físico e como homem.

Aos meus pais pela minha formação.

Abstract

In the first part we perform a detailed investigation of the storage properties of a model for neural networks that exhibits the same organization into clusters as Dyson's hierarchical model, for ferromagnetism, combined with Hebb's learning algorithm for p stored patterns.

In the case of finite p , we show that together with the original stored patterns or "ancestors" the system retrieves also a hierarchy of "descendants". The "descendants" differ from the "ancestor" in the signs of the cluster overlaps, and the number of this solutions increases exponentially with the cluster number, $n(\ell) \approx e^{0.45\ell}$, for large values of ℓ .

For an extensive number of stored patterns $p = \alpha N$, where N is the size of the network, we investigate the critical storage capacity of the model to both "ancestor" and "descendant" patterns. By using two different methods, a statistical mechanics formulation (mean-field) and a signal-to-noise analysis, we obtain a succession of critical storage capacities that are below the corresponding value for Hopfield's model. We use the ratio of this critical storage capacities to the same quantity as evaluated in Hopfield's model to compare the results in both methods. We present the phase diagram in the $\alpha - T$ plane for the particular case of two clusters and one descendant.

The second part is related with the study of the space of interactions in neural network models. In this case we search for the maximal critical storage capacity of models that present, in some sense, the same cluster structure that we discussed before. In order to take this structure in account we redefine the Gardner Program.

Concerning the storage of "ancestors" we obtain a value of α_c^{max} always lower than 2, which corresponds to the original case of models without structure. This result agrees with that we found before in the signal-to-noise analysis. On the other side we show in the special case of two clusters that this structure barely affects the storage of "ancestors" and "descendants" together.

Resumo

Na primeira etapa investigamos detalhadamente as propriedades de armazenamento de um modelo de redes de neurônios que apresenta uma organização em aglomerados semelhante àquela existente no Modelo Hierárquico de Dyson para ferromagnetismo. As memórias, neste modelo, são armazenadas através da Regra de Aprendizagem de Hebb, superposta à estrutura hierárquica.

No caso de um número finito de padrões, mostramos que, junto com os padrões originais ou “ancestrais”, o sistema é capaz de recuperar uma hierarquia de padrões “descendentes”. Estes padrões diferem dos “ancestrais” nos sinais relativos das magnetizações nos diferentes blocos, e o número destas soluções cresce exponencialmente com o número de blocos, $n(\ell) \approx e^{0.45\ell}$, para ℓ grande.

Para um número extensivo de padrões armazenados $p = \alpha N$, onde N é o tamanho do sistema, nós investigamos a capacidade crítica de armazenamento do modelo tanto para “ancestrais” como para “descendentes”. Usando dois métodos distintos, uma formulação de mecânica estatística de equilíbrio (campo médio) e uma análise de sinal-ruído, nós obtemos uma sucessão de capacidades de armazenamento que são sempre menores que o valor correspondente ao modelo de Hopfield. Usamos as razões entre estas capacidades e a do modelo de Hopfield para comparar os resultados dos dois métodos. Nós apresentamos o diagrama de fases no plano $\alpha - T$ para o caso especial de dois blocos e um único “descendente”.

A segunda parte é relacionada com o estudo do espaço de interações dos modelos de redes de neurônios. Neste caso nós buscamos determinar a máxima capacidade crítica de armazenamento para modelos que apresentem, de alguma forma, a mesma estrutura de blocos que discutimos antes. Para ter em conta esta estrutura foi preciso modificar o problema de Gardner.

Para o armazenamento de padrões “ancestrais”, obtivemos um valor de α_c^{max} sempre menor que 2, o qual corresponde ao caso de modelos sem estrutura predefinida. Este resultado combina com aquele que encontramos anteriormente na análise de sinal-ruído. Por outro lado, mostramos que no caso especial de dois blocos esta estrutura pouco afeta a armazenagem conjunta de “ancestrais” e “descendentes”.

Índice

1	Introdução	4
1.1	Bases Biológicas	4
1.2	Modelamento Matemático	6
1.2.1	Redes de Neurônios Dirigidas	7
1.2.2	Redes de Neurônios Atratoras	8
1.3	Modelo de Hopfield-Little	10
1.3.1	Principais Resultados do Modelo de Hopfield	12
1.4	Além do Modelo de Hopfield	17
1.4.1	Codificação. Geração de padrões.	18
1.4.2	Regras de Aprendizado para padrões hierárquicos	21
1.4.3	Modelos com Estrutura Espacial	27
2	Número finito de padrões.	30
2.1	O Modelo	31
2.2	Propriedades da Matriz $\underline{A}(\ell)$	36
2.3	Cálculo da energia livre	39
2.4	Soluções que envolvem um único padrão	42
2.4.1	Soluções onde \vec{m} é autovetor de \underline{A}	43
2.4.2	Soluções onde \vec{m} não é necessariamente um autovetor de \underline{A}	46
2.5	Solução com mistura de padrões	53
2.5.1	Soluções onde \vec{m} é autovetor de \underline{A}	53

<i>Índice</i>	2
2.5.2 Soluções com mistura de autovetores e memórias	54
3 Propriedades de saturação.	55
3.1 Campo Médio - Cálculo da Função de Energia Livre	57
3.1.1 Cálculo de $\langle \exp(\frac{\beta}{2} \sum_{i,j} \Lambda_{ij} \xi_i \xi_j) \rangle_\xi$	59
3.1.2 Cálculo do traço sobre as variáveis de spin em $\langle Z^n \rangle$	65
3.2 Teoria simétrica de réplicas	67
3.2.1 Soluções homogêneas	69
3.3 Diagrama de fases $\alpha \times T$ para as soluções homogêneas	71
3.3.1 T_G - Transição de fases Vidro de Spin \Leftrightarrow Paramagnetismo	74
3.3.2 Temperaturas críticas em $\alpha = 0$	75
3.3.3 Diagrama a $T = 0$ - Determinação do armazenamento crítico	77
3.3.4 Condição para a existência da curva T_C^2	81
3.4 Análise de Sinal \times ruído	84
4 Investigação do espaço de interações	90
4.1 O problema de Gardner	90
4.2 Método de Gardner para um sistema com hierarquia nas interações	93
4.2.1 Determinação de α_c no caso de armazenamento de padrões descortrelacionados	96
4.2.2 O Método de Gardner e a Armazenagem de Padrões Descendentes	103
5 Conclusões	113
5.1 Resultados	114
5.1.1 Redes Atratoras	114
5.1.2 Investigação do Espaço de Modelos	116
5.2 Comentários Finais	117
5.2.1 Possíveis Extensões	118

A	Estabilidade de Soluções	120
A.1	Cálculo específico de $\frac{\partial^2 \mathcal{F}_n}{\partial q_c^{\rho\rho} \partial q_d^{\gamma\delta}}$	123
A.2	Solução com simetria de réplicas	125
A.3	Estabilidade das Soluções Homogêneas	131
A.3.1	Flutuações Fortemente Assimétricas	135
A.3.2	Flutuações Levemente Assimétricas	137
A.3.3	Flutuações Simétricas	139
B	Cálculos de Gardner com descendentes	141
B.1	Cálculo das equações de ponto de sela	141
B.2	Cálculo das equações no limite $\epsilon_1, \epsilon_2 \rightarrow 0$	143
B.2.1	Cálculo de $\Delta(C_1)$	144
	Referências Bibliográficas	152

Capítulo 1

Introdução

1.1 Bases Biológicas

Todo o processamento que é feito no cérebro ocorre devido a presença de células especializadas chamadas “neurônios”. Morfologicamente um neurônio é composto de um corpo celular ramificado, em forma de estrela, onde está presente seu núcleo. Entre as ramificações ou “dendritos” existe uma principal, em geral muitas vezes maior que o corpo celular, chamada “axônio”. O axônio também apresenta ramificações cujos terminais se ligam aos dendritos de outros neurônios. Estas ligações são as “sinapses”.

Cada um dos neurônios atua como uma unidade processadora. Através de seus dendritos informações vindas de outras células, em forma de pulsos elétricos, são coletadas, processadas no corpo celular e posteriormente retransmitidas no axônio.

Um neurônio pode ser encontrado em dois estados: ou está em repouso ou está ativado emitindo sinais através de seu axônio. O estado de repouso caracteriza-se por uma diferença de potencial de aproximadamente -70 mV entre o interior e o exterior da célula. Tal polarização se encontra presente em qualquer célula e é resultado do transporte ativo, ou seja, contra o gradiente de concentração, de íons de sódio para fora da membrana celular [1]. Quando está ativo ou “disparando” o neurônio apresenta uma série de pulsos positivos que viajam pelo axônio em direção contrária ao corpo celular. Estes pulsos ou “potenciais de ação” são inversões locais da polarização da membrana que envolve

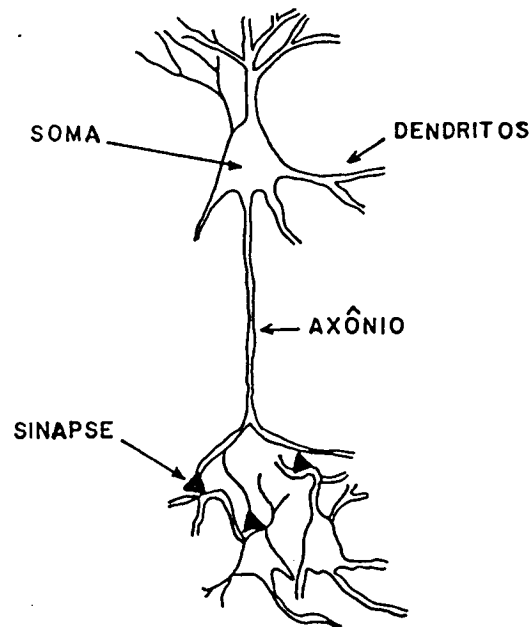


Figura 1.1: As três partes do neurônio: O corpo celular, o axônio e os dendritos. As ramificações do axônio terminam em sinapses (triângulos) por meio das quais as células se comunicam.

o axônio e podem atingir até +100 mV de amplitude, contada a partir do potencial de repouso, e uma duração da ordem de meio segundo.

A ativação de um neurônio, inicialmente em repouso, depende do potencial de despolarização causado pela soma dos sinais que chegam de outros neurônios durante um tempo característico. Em geral, observa-se que existe um valor limite do potencial de despolarização que um neurônio pode suportar sem ser ativado. Este limiar é em torno de -50 mV. A amplitude do potencial de ação, sua duração e velocidade de condução parecem independem da excitação que o neurônio sofre [1]. Já a frequência destes potenciais depende tanto do tipo do neurônio como da intensidade do potencial de despolarização sobre ele [2].

A informação de que um neurônio está ativo é passada aos demais através das sinapses. Estas ligações de caráter estritamente unidirecionais podem estimular o próximo neurônio a também se ativar ou então inibi-lo, procurando mantê-lo em repouso. No primeiro caso elas são ditas “excitatórias” e no outro “inibitórias”. Além disso uma sinapse pode ser forte ou fraca dependendo de sua eficiência em excitar ou inibir. A principal característica

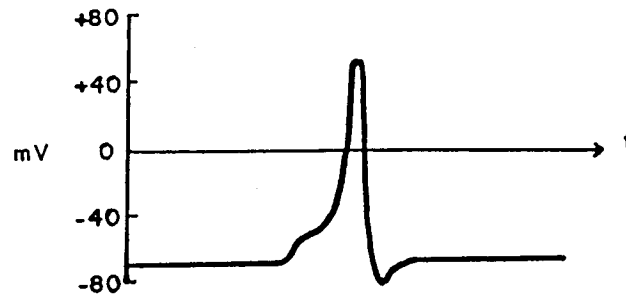


Figura 1.2: Diferença de potencial local entre o interior e o exterior da membrana citoplasmática de um neurônio durante a emissão de um “Potencial de Ação”

das ligações sinápticas, no entanto, é a sua capacidade de se modificar plasticamente. Ou seja, uma sinapse pode aumentar ou diminuir sua eficiência dependendo se os estados dos neurônios que liga coincidem ou não. Este é apontado como sendo o fenômeno chave para o processo de aprendizado no cérebro.

1.2 Modelamento Matemático

Na seção anterior procurou-se dar uma idéia dos principais mecanismos biológicos que, no entender da teoria de Redes de Neurônios, seriam responsáveis pela maioria das funções cognitivas desempenhadas pelo cérebro. Baseados nestes mecanismos, poderemos formular algumas hipóteses que serão muito importantes para a exata delimitação dos modelos matemáticos usados para estudar este sistema.

Primeira hipótese: Informação. O cérebro processa informações e a representação destas informações é feita através da configuração de estados dos seus neurônios em um dado intervalo de tempo. Além disso, considera-se que a variável de estado de um neurônio assuma somente dois valores: ativo ou inativo [3]. Com isso é descartada a possibilidade de haver informações codificadas em termos de da frequência de potenciais

de ação [2]. Define-se então uma correspondência biunívoca entre estados de sistema e informações. Para o caso de uma rede de N neurônios isto é representado pelo vetor N -dimensional

$$\vec{S} = (S_1, S_2, \dots, S_N) \quad ,$$

onde $S_i = \pm 1$.

Segunda Hipótese: Funcionamento. A rede deve possuir uma regra dinâmica que determine para um dado neurônio qual será o seu estado em um certo instante de tempo, baseado nas informações recebidas de outros neurônios ativados em um tempo anterior. A decisão de um neurônio disparar é função da integração temporal dos sinais recebidos de outros neurônios, devidamente ponderados pelos valores das eficiências das sinapses correspondentes e também de um valor limiar que pode depender do neurônio em questão. Isso pode ser representado na forma de um conjunto de N equações dinâmicas, uma para cada um dos neurônios:

$$S_i(t) = F\left(\sum_{j \neq i} J_{ij} S_j(t - \tau), \theta_i\right) \quad .$$

A quantidade J_{ij} representa o valor da eficiência sináptica na transmissão de um sinal do neurônio “i” para o neurônio “j”. Se for negativa, a sinapse é inibitória caso contrário será excitatória. E θ_i é o limiar do neurônio “i”.

Terceira Hipótese: Aprendizado. Uma vez que a dinâmica de uma rede neuronal está definida, os seus estados estacionários são única e exclusivamente determinados pela natureza e eficiência de cada uma das suas sinapses; e é na modificação destas que se dá o aprendizado da rede.

Dentro do panorama criado pelas hipóteses acima, podemos destacar dois grandes conjuntos de modelos: as redes de neurônios dirigidas e as redes de neurônios atratoras.

1.2.1 Redes de Neurônios Dirigidas

Uma rede dirigida é composta por uma série de camadas de unidades, ou neurônios, onde a informação percorre o sistema de uma camada a outra sempre em uma mesma

direção. Este tipo de computação é denominada computação por fluxo dinâmico [4]. Deste modo existe uma camada bem específica que serve de entrada e outra de saída. A principal propriedade destas redes é o fato que a informação de saída pode ser de natureza diversa daquela de entrada. Por exemplo, as unidades de entrada podem estar relacionadas a um estímulo visual enquanto na saída pode estar a decisão de mover ou não um músculo. O caso mais simples destes sistemas é o Perceptron [5].

Composto de um certo número de unidades de entrada o Perceptron possui apenas duas camadas sendo que a camada de saída é constituída de uma única unidade. A sua tarefa típica é o reconhecimento de padrões. Por exemplo, ao apresentarmos uma série de fotografias digitalizadas em forma de unidades acesas e apagadas podemos desejar que a rede responda com a unidade de saída acesa caso haja uma árvore na foto, ou apagada caso não. Durante o processo que chamamos “aprendizagem”, a rede deve ser capaz de responder corretamente ao que foi requisitado para um número definido de fotos chamadas “exemplos”. Tal objetivo é atingido via a modificação de suas eficiências sinápticas, determinada pela regra de aprendizagem particular da rede. Após a etapa de aprendizagem a rede deve poder “generalizar”. A generalização consiste na sua capacidade em reconhecer ou não árvores em outras fotos diferentes daquelas apresentadas como exemplos. A principal ferramenta usada para o estudo deste tipo de sistemas permanece sendo a simulação numérica. Questões teóricas de interesse neste caso são, por exemplo: Como determinar se realmente existe um conjunto de valores das eficiência sinápticas que possibilitam êxito na etapa de aprendizagem? Quantos são os exemplos necessários para implementar uma boa capacidade de generalização? [6]

1.2.2 Redes de Neurônios Atratoras

Em uma rede do tipo atratora não existem unidades de entrada ou unidades de saída, e cada um dos neurônios pode se ligar ou receber ligações de qualquer outro. O tipo de computação realizada por elas consiste em um processo onde a rede parte de uma configuração inicial, interpretada como estímulo externo, e evolui livremente sob o efeito

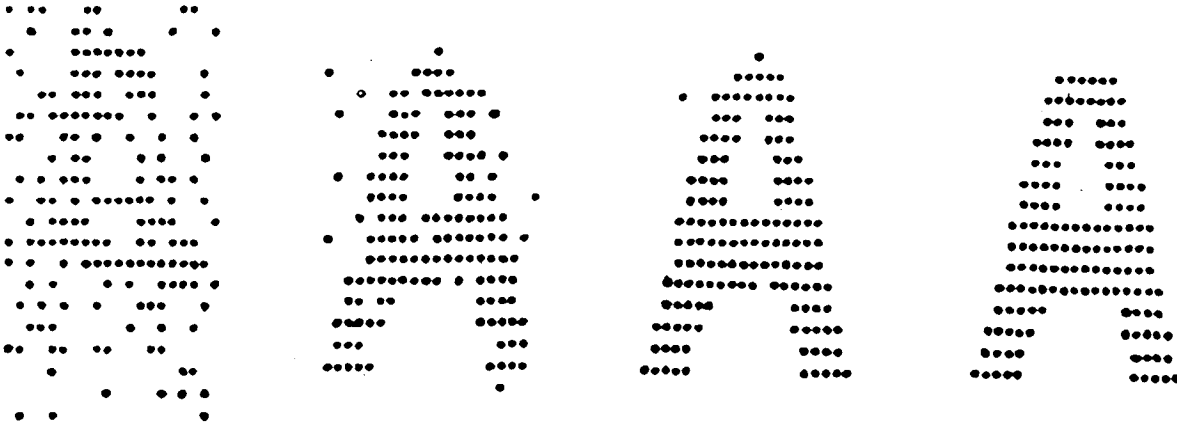


Figura 1.3: Simulação de uma rede atratora com 400 neurônios e 30 padrões armazenados, dos quais 29 são aleatórios e um deles é a letra A. A figura mostra 3 ciclos de atualização onde a rede recupera a memória A a partir de uma configuração inicial 30% diferente [7].

de sua regra dinâmica até atingir um estado estacionário. Este estado final tanto pode ser uma certa configuração como uma sequência oscilante delas.

No caso em que os estados estacionários são pontos fixos, ou seja configurações únicas, o sistema modela um processo de memória associativa. Os pontos fixos são as memórias e o processo de evolução dinâmica que vai desde uma configuração inicial qualquer, até um destes estados estacionários, é denominado “recuperação de memória”. Confira a figura 1.3.

O problema principal destes modelos é como escolher as ligações sinápticas de forma que se consiga armazenar o maior número possível de memórias.

Em geral o valor das eficiências sinápticas deve depender do conjunto de memórias que se deseja armazenar. Assim, se denotarmos o conjunto das p configurações ou padrões que queremos memorizar por: $\vec{\xi}^\mu$ com $\mu = 1, \dots, p$ e $\vec{\xi} = (\xi_1, \dots, \xi_N)$ com $\xi_i^\mu = \pm 1$, então a função

$$J_{ij} = J_{ij}(\{\xi_i^\mu\})$$

que determina as ligações sinápticas J_{ij} é chamada “Regra de Aprendizagem”. Uma regra de aprendizagem é dita local quando J_{ij} somente depender dos valores de ξ^μ nos sítios i

e j .

O modelo mais simples para uma destas redes é o Modelo de Hopfield que estudaremos a seguir.

1.3 Modelo de Hopfield-Little

Observando as hipóteses apresentadas anteriormente é imediato notar a semelhança entre estes modelos e os modelos para sistemas magnéticos. A analogia é simples. O estado de neurônio equivale a estado de spin de Ising, e as eficiências sinápticas seriam os acoplamentos magnéticos. O maior obstáculo a uma equivalência completa dos modelos é, no entanto, a não simetria entre as ligações sinápticas. Como as sinapses são unidirecionais e fisicamente independentes, não há como supor que a sinapse que liga o neurônio i ao j seja da mesma natureza e eficiência daquela que faz o caminho inverso, de j para i . Na verdade, a existência de uma nem ao menos garante a presença da outra. Já nos modelos de sistemas magnéticos com spins clássicos a simetria $J_{ij} = J_{ji}$ entre acoplamentos deve necessariamente existir. Este problema foi resolvido por John Hopfield que decidiu estudar as propriedades de uma rede de Neurônios com simetria de sinapses. No modelo de Hopfield [8] a lei dinâmica que determina o estado de cada neurônio em função dos demais é

$$S_i(t) = \text{sgn}\left(\sum_{j \neq i} J_{ij} S_j(t - \tau) - \theta_i\right) \quad , \quad (1.1)$$

onde S_i é o estado do neurônio i e J_{ij} é a eficiência sináptica entre neurônios que é dada pela regra de aprendizagem de Hebb [9],

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^P \xi_i^\mu \xi_j^\mu \quad . \quad (1.2)$$

O que a dinâmica definida acima faz é alinhar cada um dos neurônios com a grandeza $h_i = \sum_{j \neq i} J_{ij} S_j$, que é equivalente ao campo local no magnetismo. Assim, um estado estacionário desta dinâmica será aquele em que todos os neurônios estiverem alinhados com seus respectivos campos. O que Hopfield descobriu e' que, no caso de " J_{ij} " simétrico,

o processo de buscar estes estados é equivalente a minimizar uma certa função definida como a Função Energia da Rede

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

Esta energia resulta ser a mesma que escreveríamos para um sistema magnético. Observe que propositalmente já tem sido usada a notação de S (spin) para neurônio e J (acoplamento) para eficiência sináptica.

Apesar de inaceitável do ponto de vista biológico, a hipótese de simetria nas sinapses não se revelou tão absurda afinal. Além de tornar o problema tratável, ainda mostrou que as redes de neurônios poderiam ser comparadas com sistemas magnéticos que apresentassem desordem nas ligações, ou modelos de Vidro de Spin [10]; e estes modelos possuem precisamente, algumas das propriedades que se deseja para as redes. Um dos exemplos é a multiplicidade dos estados fundamentais. Nestes sistemas a aleatoriedade das ligações dá origem ao fenômeno da “frustração” [11], ou seja, a impossibilidade de satisfazer todas as ligações ferromagnéticas e antiferromagnéticas entre pares de spins. Desta forma, nem todos estes pares podem ficar na sua configuração de menor energia. O resultado disso a nível de sistema inteiro é a existência de vários estados de mínima energia.

Um exemplo clássico de frustração é a de um sistema hipotético contendo unicamente três spins, sendo dois deles acoplados antiferromagneticamente e as ligações restantes todas ferromagnéticas. Dentre seus 8 possíveis estados podemos destacar 6 de mínima energia contendo cada um deles um par frustrado. O número de estados de mínimo cresce com o número de spins da rede.

Existem dois modos de implementarmos a atualização dada pela equação (1.1). Ou usamos uma Dinâmica Assíncrona ou Monte Carlo, onde os neurônios são escolhidos ao acaso e atualizados cada um em um dado instante ou então a Dinâmica Síncrona, ou paralela, onde os spins se atualizam todos ao mesmo tempo. Existem razões para crer que a dinâmica em sistemas nervosos reais seja algo intermediário, nem totalmente síncrona nem totalmente assíncrona. A diferença fundamental entre as duas dinâmicas é a natureza de seus estados estacionários. Enquanto o processo Monte Carlo só pode

apresentar estados estacionários do tipo pontos fixos, a dinâmica paralela pode também possuir ciclos limites, onde o sistema ficaria oscilando entre duas configurações diferentes [12].

O Modelo de Hopfield é definido como sendo resultado da dinâmica Monte Carlo, enquanto que aquele com a dinâmica paralela é chamado Modelo de Little [13]. Neste trabalho vamos nos ocupar com sistemas que possuam dinâmica do tipo Monte Carlo por serem estes que possuem uma precisa analogia com sistemas magnéticos [14].

Um importante passo no estudo das redes de neurônios foi a introdução do conceito de temperatura [13]. Assim, ao invés de considerarmos a dinâmica rígida e determinista de (1.1), teríamos uma dinâmica estocástica governada pelas probabilidades:

$$S_i(t) = \begin{cases} +1 & \text{com prob. } 1/(1 + e^{-2h_i/T}) \\ -1 & \text{com prob. } 1/(1 + e^{2h_i/T}) \end{cases} . \quad (1.4)$$

Da mesma forma que a dinâmica anterior era um processo de busca de mínimos locais da função energia H , esta dinâmica procura mínimos de uma função energia livre definida tal qual a conhecemos. Cabe, no entanto, ressaltar que, assim como H não tem significação física em termos de uma grandeza medida em joules, a temperatura também não é algo que se possa medir em kelvin, é tão somente uma variável que indica o nível de ruído da rede. Geralmente introduzimos ruído para representar graus de liberdade não contemplados explicitamente no modelo; neste caso, estes graus de liberdade poderiam ser flutuações químicas no meio intercelular que influenciariam o comportamento das sinapses.

1.3.1 Principais Resultados do Modelo de Hopfield

Do ponto de vista de mecânica estatística de equilíbrio, a técnica que usaremos predominantemente neste trabalho, o tipo de informações que podemos extrair do modelo é, tendo fixada a função energia e a distribuição de probabilidades dos acoplamentos J_{ij} : quais são os seus estados estacionários, sua existência e estabilidade no diagrama de fases do sistema.

Os estados estacionários obtidos para um número não extensivo de padrões armazenados, ou seja $p/N \rightarrow 0$ no limite termodinâmico [15], sujeitos à distribuição de probabilidades

$$P(\xi_i^\mu) = \frac{1}{2} \delta(\xi_i^\mu - 1) + \frac{1}{2} \delta(\xi_i^\mu + 1) \quad , \quad (1.5)$$

e considerando como parâmetro de ordem o vetor $\vec{m} = (m^1, \dots, m^p)$, com suas componentes definidas como a superposição dos estados de equilíbrio do sistema com as memórias armazenadas, ou seja,

$$m^\mu = \frac{1}{N} \langle \sum_i \langle S_i \rangle_T \xi_i^\mu \rangle_\xi \quad , \quad (1.6)$$

são:

- Estados de recuperação : $\vec{m} = m (1, 0, \dots, 0)$
- Estados misturados simétricos : $\vec{m} = m_n (1, \dots, 1, 0, \dots, 0)$
- Estados misturados assimétricos : $\vec{m} = (m_1, \dots, m_1, \dots, m_k, \dots, m_k, 0, \dots, 0)$
- Estado paramagnético : $\vec{m} = 0$

Os estados de recuperação são em número de p estados e representam a situação em que os neurônios da rede estão alinhados com uma das memórias. Estes estados surgem continuamente a medida que baixamos a temperatura (ruído) a partir de $T = 1$, sendo estáveis e de mínima energia livre em todo o intervalo $0 < T < 1$.

Os estados misturados simétricos, onde o sistema recupera uma fração igual de n memórias, surgem também continuamente a partir de $T = 1$. Sendo, no entanto, os estados de n par sempre instáveis, enquanto que aqueles com n ímpar se estabilizam para uma temperatura $0 < T_n < 1$. A energia livre destes estados cresce com o número de memórias misturadas.

Os estados não simétricos, nos quais várias memórias são recuperadas em frações diferentes, aparecem somente abaixo de $T = 0.575$ e isto pode ocorrer de forma contínua ou descontínua. Em geral, aqueles que o fazem continuamente, logo que surgem são instáveis, se estabilizando a medida que a temperatura é diminuída. Estados que surgem descontínuamente, no entanto, são instáveis em todo o intervalo de temperatura.

Estados misturados são também denominados estados espúrios. A princípio, tais estados são indesejáveis pois atuam como atratores na dinâmica do sistema, atrapalhando a recuperação das memórias verdadeiras. No caso específico de uma máquina de recuperação de memória, o ideal seria então eliminar estes estados, o que pode ser tentado tanto por modificações na regra de aprendizagem [16] como no próprio modelo [17].

Agora, se observarmos com cuidado isto que foi chamado “máquina de memória”, veremos que seu funcionamento pouco tem a ver com aquele que conhecemos de um cérebro humano. Este não simplesmente congela informações mas as categoriza e cria conceitos. Essa categorização e criação de conceitos, no nosso formalismo nada mais é que a criação de novos estados estacionários a partir das memórias verdadeiras e, nesse sentido, estados espúrios podem vir a desempenhar um papel essencial [18] [19].

Os principais parâmetros externos do modelo de Hopfield são: o ruído T e o coeficiente de armazenamento $\alpha = \frac{p}{N}$ (considerado igual a zero na discussão anterior). Quando temos este último parâmetro diferente de zero, é imediato observar que a distribuição de probabilidades de termos um certo valor J_{ij} para a interação sináptica entre os sítios “i” e “j” é dada por

$$P(j_{ij}) = \sqrt{\frac{N}{2\pi\alpha}} e^{-N \frac{j_{ij}^2}{2\alpha}} .$$

Esta é uma distribuição Gaussiana que é idêntica àquela usada no modelo de Sherrington e Kirkpatrick [20] para o seu modelo de vidro de spin. Portanto, quando construirmos o diagrama de fases $\alpha \times T$ do modelo de Hopfield vamos encontrar, além da região paramagnética e a de recuperação, uma região de vidro de spin.

Estas soluções de vidro de spin também atuam como atratores na dinâmica do sistema, e sua principal característica é de não possuírem superposição finita com nenhuma das memórias armazenadas, ou seja, $\vec{m} = 0$. Apesar disso, tais estados diferem de um estado paramagnético devido a existência de uma correlação temporal entre os estados de neurônios, caracterizada pelo fato do parâmetro de Edward-Anderson, definido como

$$q(t_1, t_2) = \overline{S_i(t_1)S_i(t_2)}$$

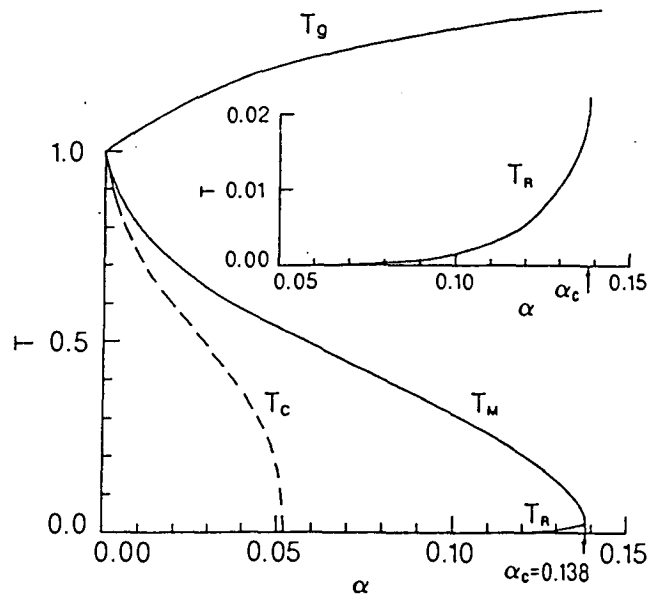


Figura 1.4: Diagrama de fases da solução de campo médio do Modelo de Hopfield [14].

ser diferente de zero.

Em redes de neurônios estamos sempre interessados em aspectos dinâmicos do problema. Neste sentido estados meta-estáveis são tão importantes quanto os estáveis, pois mínimos locais também são capazes de aprisionar dinamicamente o sistema. Por esta razão a região de recuperação é definida a partir do momento que surge a solução de recuperação, mesmo que esta não seja o mínimo global da energia livre.

Um diagrama completo para redes de neurônios deveria apresentar tanto estados de recuperação como estados espúrios. Isto na prática é inviável pela grande quantidade destes estados. No regime de recuperação eles são da ordem de $p^{\rho(T)}$, onde $\rho(T)$ é um polinômio em T , e, acima da linha crítica T_M (vide figura 1.4), crescem exponencialmente com N [21].

A figura 1.4 mostra o diagrama de fases da solução de campo médio do modelo de Hopfield sem os estados metaestáveis [14]. A linha T_M , que é a linha crítica $\alpha_c(T)$ a qual nos referimos antes, marca o aparecimento de forma descontínua da solução de recuperação, enquanto que a linha tracejada delimita a região onde esta solução é mínimo global da energia livre.

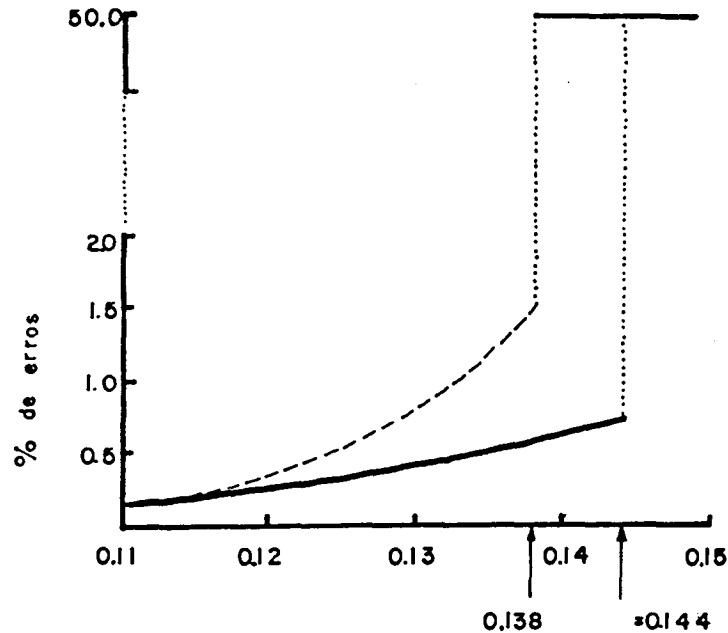


Figura 1.5: Percentual de sítios errados no caso de cálculo com simetria (linha tracejada) e com assimetria de réplicas (linha cheia) [22].

A $T = 0$ temos o maior valor de α_c , correspondente à máxima armazenagem. Este valor é $\alpha_c = 0.138$ se considerarmos simetria de réplicas na solução das equações de ponto de sela e aumenta para $\alpha_c = 0.144$ no caso de quebra desta simetria, concordando com valores obtidos por simulação numérica [8]. Tanto um como outro foram calculados permitindo uma fração de erro na recuperação, ou seja, nem todos os neurônios estão alinhados com a memória recuperada.

O armazenamento crítico para a situação de uma recuperação de 100% pode ser obtido pelo método de sinal \times ruído, que será descrito detalhadamente no capítulo 3. Neste caso a fração crítica de armazenamento vale

$$\alpha_c = \frac{1}{4 \log N}$$

sendo, portanto, zero no limite termodinâmico.

Muito se tem trabalhado no sentido de aprimorar este modelo em direção a modelos mais otimizados ou mais realísticos. Em particular, especial atenção tem sido dedicada a modelos de armazenagem e recuperação de memórias relacionadas de forma hierárquica. Isto significa que de acordo com suas superposições (overlaps) $\frac{1}{N} \sum_i \xi_i^\mu \xi_i^\nu$ estas podem ser dispostas em categorias seguindo uma estrutura de árvore hierárquica. A razão do

interesse nestes modelos vem do fato de que categorizações e hierarquias estarem presentes tanto nas funções cerebrais, que queremos reproduzir. como, aparentemente, na estrutura intrínseca do espaço de estados de um modelo de vidro de spin .

O objetivo deste trabalho de doutoramento é estudar as consequências nas propriedades de armazenamento e recuperação de memória de uma rede de neurônios, quando a sua arquitetura é tal que existe uma organização hierárquica entre neurônios e não entre memórias. Assim, como primeiro passo, vamos descrever alguns modelos que buscam aperfeiçoar o Modelo de Hopfield.

1.4 Além do Modelo de Hopfield

Um modelo de Redes de Neurônios serve a dois propósitos: primeiro, nos dar subsídios para entendermos como a mente humana funciona e em segundo lugar nos possibilita construir dispositivos que desempenhem tarefas semelhantes. Assim, para sabermos se um modelo de redes de neurônios é um bom modelo, em geral, devemos comparar seu comportamento e estrutura com o que conhecemos sobre o cérebro humano. Este confronto com o sistema real tende a ser frutífero na medida em que nos assegura uma direção a ser seguida no aprimoramento das capacidades destes modelos.

Nesta seção discutiremos duas características arquitetônicas que encontramos no cérebro humano e não são contempladas no modelo de Hopfield bem como alguns dos principais modelos criados para incluí-las.

A primeira delas está relacionada com as memórias ou os padrões que desejamos armazenar. O modelo de Hopfield foi construído de forma que seu melhor funcionamento se dê para padrões aproximadamente ortogonais, isto é ,

$$q^{\mu,\nu} = \frac{1}{N} \sum_i \xi_i^\mu \xi_i^\nu \approx 0 \quad . \quad (1.7)$$

Não é um absurdo considerar que, quando temos duas ideias próximas, ou então duas imagens semelhantes, os seus códigos neuronais sejam tamb' em semelhantes; em outras palavras, as representações destes dois conceitos, em termos de "+1" e "-1" tenham uma

superposição grande. Sendo assim, o armazenamento de padrões ortogonais (ou seja, de conceitos muito diferentes) é um evento muito raro no que a gente espera que seja a rotina normal de funcionamento de um cérebro [19].

Sabemos que qualquer informação nova que aprendemos é imediatamente catalogada em relação às demais já existentes. O tipo mais usual desta classificação é aquela que agrupa os padrões em categorias e subcategorias. Os modelos hierárquicos são uma classe de modelos construídos a partir de modificações convenientes da regra de Hebb, de forma que suportem, melhor que o Modelo de Hopfield, o armazenamento de informações que não são necessariamente ortogonais entre si. Os padrões armazenados nestes modelos são tais que suas superposições satisfazem relações ultramétricas, o que indica que eles podem ser classificados em uma árvore hierárquica [23].

A seguir descreverei algumas das propriedades destes modelos, sem, no entanto, descer a detalhes mais finos. A minha intenção, neste caso, é apenas dar uma idéia geral da área, pois o modelo que desenvolverei nos capítulos posteriores, apesar de conter como ingrediente principal a idéia de hierarquização, não está diretamente relacionado com estes modelos hierárquicos.

1.4.1 Codificação. Geração de padrões.

Os padrões que usamos em modelos de redes de neurônios atratoras são p conjuntos de N variáveis aleatórias. De acordo com a distribuição que gera estes padrões podemos classificá-los em três grupos : ortogonais, com atividade ou hierárquicos.

i) Padrões Ortogonais : O modo mais simples de construir padrões que sejam ortogonais no limite em que N é grande, é fazendo com que a probabilidade de cada sítio estar “aceso” (+1) seja igual a de estar “apagado” (-1), ou seja,

$$P(\xi_i^\mu) = \frac{1}{2} \delta(\xi_i^\mu - 1) + \frac{1}{2} \delta(\xi_i^\mu + 1) \quad . \quad (1.8)$$

Assim, se calcularmos a superposição de quaisquer dois destes padrões obteremos:

$$q^{\mu,\nu} = \frac{1}{N} \sum_i \xi_i^\mu \xi_i^\nu \simeq \langle \xi_i^\mu \xi_i^\nu \rangle = 0 \quad . \quad (1.9)$$

Os padrões ortogonais se encontram o mais longe possível um do outro no espaço de estados. A medida de distância em questão, é a distância de Hamming definida como $d_{\mu\nu} = 1 - q^{\mu\nu}$.

ii) Padrões com atividade : Estes padrões são aqueles para os quais existe uma tendência a ter mais neurônios em um estado do que no outro. Neste caso,

$$P(\xi_i^\mu) = \frac{1+a}{2} \delta(\xi_i^\mu - 1) + \frac{1-a}{2} \delta(\xi_i^\mu + 1) \quad . \quad (1.10)$$

E portanto, as superposições

$$q^{\mu,\nu} = \frac{1}{N} \sum_i \xi_i^\mu \xi_i^\nu \simeq \langle \xi_i^\mu \xi_i^\nu \rangle = a^2 \quad (1.11)$$

deixam de ser zero.

Os padrões com atividade se localizam em uma região mais restrita espaço de estados, e são equidistantes a uma distância menor $d = 1 - a^2$. Eles são tradicionalmente chamados de “padrões correlacionados” por possuírem $q \neq 0$. Esta nomenclatura, apesar de incorreta, é bastante difundida na área de redes de neurônios, de forma que será adotada em certas passagens deste trabalho. Cabe, como esclarecimento, ressaltar que correlação estatística está definida como

$$c^{\mu,\nu} = \langle \xi_i^\mu \xi_i^\nu \rangle - \langle \xi_i^\mu \rangle \langle \xi_i^\nu \rangle$$

e, neste sentido, nenhum dos dois grupos de padrões possui correlação.

iii) Padrões Hierárquicos. Os padrões hierárquicos são uma generalização dos dois tipos anteriores e são tais que se distribuem no espaço de estados de forma que a distância entre eles obedece uma certa hierarquia. Para gerarmos um conjunto de padrões com essa propriedade devemos usar um processo estatístico diferente dos anteriores. Um exemplo é o que descreveremos a seguir [18]:

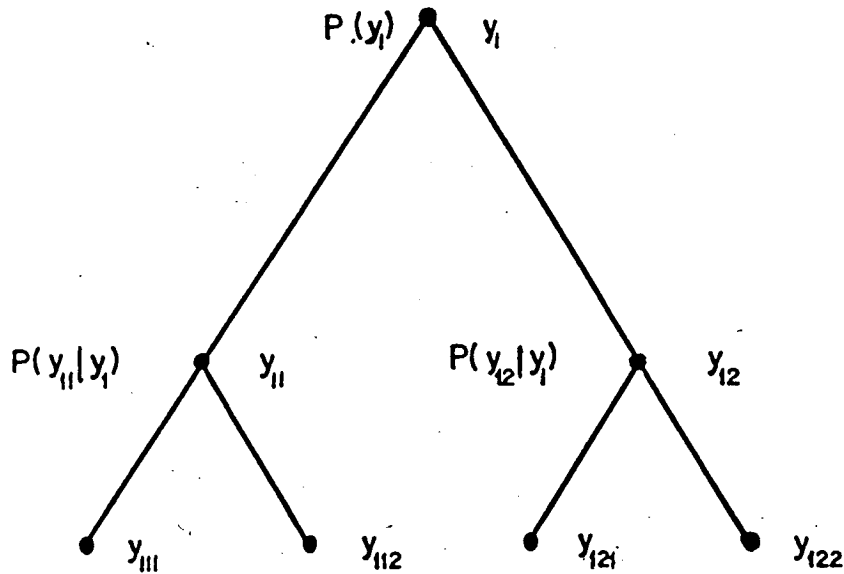


Figura 1.6: Procedimento estatístico para a geração de 4 padrões organizados em hierarquia.

Considere o problema de gerar 4 padrões tais que sejam divididos em duas categorias, cada uma delas com dois padrões. O objetivo é ter o que é mostrado na figura 1.6, onde os pontos inferiores, que representam as memórias, são tais que $d_{12} < d_{(13)} = d_{14}$.

O processo estatístico, para cada sítio, começa no topo da árvore. Lá a variável y_1 é gerada a partir da distribuição $P(y_1)$; a seguir, é a vez de gerarmos y_{11} que é condicionada a probabilidade $P(y_{11} | y_1)$, que dependerá do valor previamente obtido para y_1 . Para uma árvore genérica, este processo se repete até o extremo inferior da árvore ser atingido. No processo intermediário, cada variável $y_{l_1 \dots l_n}$ é sempre gerada por uma distribuição $P_l(y_{l_1 \dots l_n} | y_{l_1 \dots l_{n-1}})$ que dependerá do valor da variável $y^{(l-1)}$, o seu ancestral imediatamente anterior. A princípio, $y_{l_1 \dots l_n}$ pode ser contínua ou discreta, assumindo valores em qualquer intervalo real; essas variáveis não possuem, em geral, uma “significação física”, com exceção da última delas que está relacionada com os padrões. Estes padrões são definidos como

$$\xi^\mu = \text{sgn}(y_{l_1 \dots l_\kappa})$$

onde κ é a última geração da árvore. Obviamente existe na expressão acima um relação que une a numeração $\mu = 1, \dots, p$ às diádicas $\{l_1 \dots l_\kappa\}$. Por exemplo, no caso da figura

1.6 temos que 111 se refere à memória $\mu = 1$, 112 à memória $\mu = 2$, e assim por diante.

Calculando as superposições entre as memórias 1, 2 e 3 da figura 1.6 obtemos:

$$q^{12} \simeq \langle \xi^1 \xi^2 \rangle \quad (1.12)$$

$$q^{12} \simeq \int dy_1 P(y_1) \int dy_{11} P(y_{11} | y_1) \left(\int dy P(y_{111} | y_{11}) \text{sgn}(y_{111}) \right)^2$$

$$q^{13} \simeq \langle \xi^1 \xi^3 \rangle \quad (1.13)$$

$$q^{13} \simeq \int dy_1 P(y_1) \left(\int dy_{11} P(y_{11} | y_1) \int dy_{111} P(y_{111} | y_{11}) \text{sgn}(y_{111}) \right)^2$$

Observe que pela desigualdade triangular

$$\begin{aligned} q^{12} - q^{13} &= \langle \langle \langle \xi \rangle_3^2 \rangle_2 \rangle_1 - \langle \langle \langle \xi \rangle_3 \rangle_2^2 \rangle_1 \\ &= \langle \langle (\langle \xi \rangle_3 - \langle \langle \xi \rangle_3 \rangle_2)^2 \rangle_2 \rangle_1 \geq 0 \end{aligned} \quad (1.14)$$

onde os diversos $\langle \dots \rangle_i$ significam cada uma das integrais das expressões em 1.12.

Neste caso

$$P(\xi^\mu \xi^\nu) \neq P(\xi^\mu)P(\xi^\nu)$$

o que indica que as variáveis ξ^μ não são estatisticamente independentes e portanto a correlação (estatística) entre algumas delas pode ser não nula.

Existem muitos processos, uma vez definidas as distribuições P_k para cada geração, que podem resultar em um mesmo conjunto de $\{q^{11}, \dots, q^{1\kappa}\}$. O que os vai diferenciar são as correções de mais alta ordem. Com os dados disponíveis da neurobiologia não há nenhuma escolha natural para este processo [18].

1.4.2 Regras de Aprendizado para padrões hierárquicos

As principais propostas para o armazenamento de padrões hierárquicos são:

Modelo de Parga-Virasoro [18] Este modelo é baseado no conhecimento obtido da

estrutura de estados estacionários do modelo de Sherrington-Kirkpatrick. Do qual é

possível extrair a seguinte regra de aprendizagem.

$$J_{ij}^{(l)} = \frac{1}{N} \sum_{\mu_1, \dots, \mu_l} \frac{1}{q_l - q_{l-1}} (\xi_i^{\mu_1, \dots, \mu_l} - \xi_i^{\mu_1, \dots, \mu_{l-1}})(\xi_j^{\mu_1, \dots, \mu_l} - \xi_j^{\mu_1, \dots, \mu_{l-1}}) + J_{ij}^{(l-1)} \quad (1.15)$$

com

$$\xi^{\mu_1, \dots, \mu_l} = \text{sgn}(y^{\mu_1, \dots, \mu_l}) \quad (1.16)$$

$$q_l = \frac{1}{N} \sum_i \xi_i^{\mu_1 \dots \mu_{l+1} \dots \mu_\kappa} \xi_i^{\mu_1 \dots \mu_{l+1} \dots \mu'_\kappa} \quad (1.17)$$

Estamos usando agora os índices das diversas etapas do processo de geração para representar um padrão, de forma que o padrão $\xi_i^{\mu_1, \dots, \mu_l}$ é denominado “descendente” de $\xi_i^{\mu_1, \dots, \mu_{l-1}}$. O conjunto de descendentes de um mesmo ancestral se chama categoria.

Como os autores não especificaram o processo de geração dos padrões os resultados são em parte semiquantitativos. Considerando $m^{\mu\nu} = \frac{1}{N} \sum_i \xi_i^{\mu\nu} S_i$ os possíveis estados estacionários são, para o problema de duas hierarquias:

a) Soluções de recuperação $m^{\mu_0\nu_0} = 1$ $m^{\mu_0\nu} = q$, $m^{\mu\nu} = 0$. O sistema recupera a memória $\xi^{\mu_0\nu_0}$.

b) Soluções espúrias que misturam descendentes $m^{\mu_0\nu_i} \neq 0$ $i = 1, \dots, p$ $m^{\mu\nu} = 0$. O sistema recupera uma mistura de padrões que possuem mesmo ancestral ξ^{μ_0} , ou seja, confunde elementos de uma categoria mas ainda é capaz de diferenciar categorias.

c) Soluções espúrias que misturam ancestrais $m^{\mu_i\nu_j} \neq 0$ $j = 1, \dots, P'_\mu$ $i = 1, \dots, p'$. O sistema recupera uma mistura de padrões de várias categorias. Onde p' é o número de categorias confundidas e P'_μ o número de elementos confundidos na categoria μ .

As propriedades destes estados estacionários são:

i) Soluções do tipo a) são estáveis abaixo de uma certa temperatura crítica.

ii) O número de soluções espúrias cresce ao menos exponencialmente com P_μ (número de descendentes da memória μ).

iii) Soluções espúrias de baixas energias são essencialmente combinações lineares dos padrões originais.

iv) A bacia de atração das soluções espúrias cresce com P_μ , significando que estas bacias se localizam na região fronteira entre padrões originais.

v) É possível introduzir uma hierarquia entre erros, ou seja, fazer com que exista uma diferença entre o número de estados espúrios que misturam categorias e aqueles que misturam descendentes. O ideal seria achar um processo estatístico para geração de padrões para o qual o número de estados que confundissem categorias crescesse com P e não com o número total de padrões

$$P + \sum_{\mu=1}^P P_\mu \quad .$$

Modelo de Feigelman-Ioffe [24]

Como anteriormente a regra de aprendizagem é construída a partir de uma série de superposições de termos do tipo de Hebb. O primeiro armazena o ancestral e os subsequentes armazenam somente as diferenças entre ancestrais e descendentes nas diversas gerações:

$$J_{ij}^{(\ell)} = \frac{1}{N} \sum_{\mu_1, \dots, \mu_\ell} \frac{1}{\Delta_\ell} (\xi_i^{\mu_1, \dots, \mu_\ell} - a_\ell \xi_i^{\mu_1, \dots, \mu_{\ell-1}})(\xi_j^{\mu_1, \dots, \mu_\ell} - a_\ell \xi_j^{\mu_1, \dots, \mu_{\ell-1}}) + J_{ij}^{(\ell-1)} \quad . \quad (1.18)$$

Na primeira geração os padrões são ortogonais como em (1.9) e nas subsequentes

$$\xi^{\mu_1, \dots, \mu_\ell} = \xi^{\mu_1, \dots, \mu_{\ell-1}} \eta_{\mu_1, \dots, \mu_{\ell-1}} \quad (1.19)$$

com

$$P(\eta_{\mu_1, \dots, \mu_{\ell-1}}) = \frac{1 + a_\ell}{2} \delta(\eta_{\mu_1, \dots, \mu_{\ell-1}} - 1) + \frac{1 - a_\ell}{2} \delta(\eta_{\mu_1, \dots, \mu_{\ell-1}} + 1) \quad . \quad (1.20)$$

O parâmetro a_ℓ mede quanto os descendentes da ℓ -ésima geração são diferentes dos seus parentes, enquanto que Δ_ℓ dá um peso relativo entre ancestrais e descendentes. Este parâmetro pode variar dentro do intervalo

$$(1 - a_\ell^2) \leq \Delta_\ell < \frac{(1 - a_\ell^2)(1 + a_\ell)}{a}$$

fora do qual a estabilidade dos padrões é afetada. O limite inferior corresponde ao caso Parga-Virasoro com a distribuição especificada acima em (1.19) e (1.20).

Os resultados para duas hierarquias são:

a) Quando $\Delta = 1 - a^2$, existe degenerescência entre parentes e descendentes. A capacidade de armazenamento crítica é $\alpha_c \simeq 0.14$ onde α se refere ao número total de padrões armazenados

$$\alpha N = P + \sum_{\mu=1}^P P_\mu \quad .$$

b) Se $\Delta > 1 - a^2$ a degenerescência desaparece e os ancestrais tornam-se estados de menor energia.

c) A capacidade total de armazenamento permanece a mesma, mas os ancestrais e descendentes surgem para diferentes valores de α . Pode-se determinar duas curvas $\alpha_1(\Delta, T)$ e $\alpha_2(\Delta, T)$ com $\alpha_1 > \alpha_2$ tais que abaixo da primeira curva surgem os ancestrais e da segunda os descendentes.

Modelo de Gutfreund [25]

Nos modelos anteriores se observa que tanto ancestrais como descendentes são possíveis estados estacionários da rede. Em termos mais exatos tanto os protótipos de classe como os indivíduos desta classe são memórias recuperáveis pela mesma rede de neurônios. No modelo de Gutfreund, no entanto, temos tantas redes como hierarquias sendo que cada uma delas é responsável pelo armazenamento dos padrões daquela geração. Assim a regra de aprendizado na rede correspondente a geração ℓ fica sendo

$$J_{ij}^{(\ell)} = \frac{1}{N} \sum_{\mu_1, \dots, \mu_\ell} (\xi_i^{\mu_1, \dots, \mu_\ell} - a_\ell \xi_i^{\mu_1, \dots, \mu_{\ell-1}})(\xi_j^{\mu_1, \dots, \mu_\ell} - a_\ell \xi_j^{\mu_1, \dots, \mu_{\ell-1}}) \quad (1.21)$$

onde $\{\xi_i^{\mu_1, \dots, \mu_\ell}\}$ são os padrões armazenados nesta rede e $\{\xi_i^{\mu_1, \dots, \mu_{\ell-1}}\}$ são os ancestrais armazenados na rede anterior. A geração de descendentes é feita com no modelo de Feigelman-Ioffe.

A análise do modelo para duas gerações tem como resultados :

a) As propriedades da rede dependem somente do número total de protótipos. Portanto, a armazenagem é insensível ao agrupamento dos padrões.

b) A capacidade de armazenamento varia com a correlação mais forte aproximadamente com $\alpha_{Hopfield}(1 - a^2)$. Este resultado foi primeiramente obtido na análise de armazenamento de padrões com atividade, ou simplesmente correlacionados [26].

c) Este modelo apresenta uma dinâmica bem particular no qual cada rede de uma certa geração passa informações de forma unidirecional para a geração seguinte. Esta comunicação se faz de forma que, a rede descendente tem sua dinâmica, vinculada à região no espaço de fase onde se encontra o estado final da rede anterior.

Modelo de Dotsenko

Antes de apresentar diretamente o modelo de Dotsenko [21] é conveniente estudarmos algumas das propriedades das memórias hierárquicas.

Considere dois padrões ξ^μ , ξ^ν que possuem uma superposição $q^{\mu\nu} = q_\ell$. A nomenclatura significa que, em termos da árvore hierárquica, eles possuem o mesmo ancestral na ℓ -ésima geração. Se representarmos um padrão qualquer como

$$\xi^\mu = \xi^\mu(y_\mu^{(1)}, \dots, y_\mu^{(\kappa)})$$

onde $y_\mu^{(\ell)}$ com $\ell = 1, \dots, \kappa$ são as variáveis aleatórias que participam diretamente da geração de ξ^μ , vemos que os dois padrões μ e ν devem possuir o mesmo conjunto $y_\mu^{(1)}, \dots, y_\mu^{(\ell)}$. Supondo que cada uma destas variáveis possa assumir S_κ valores diferentes,

podemos dividir os N sítios da rede em $\prod_{k=1}^{\ell} S_k$ grupos de tal modo que os sítios pertencentes a cada um deles possua a mesma ℓ -upla $y_{\mu,\sigma_1}^{(1)} \cdots y_{\mu,\sigma_\ell}^{(\ell)}$, onde $y_{\mu,\sigma_\ell}^{(\ell)}$ com $\sigma_k = 1, \dots, S_k$ são cada um dos possíveis valores que $y_{\mu}^{(k)}$ pode ter.

Para um destes conjuntos de sítios temos então

$$M_{\sigma}^{\mu} = \frac{1}{|C_{\sigma}|} \sum_{i \in C_{\sigma}} \xi_i^{\mu} \simeq \langle \cdots \langle \xi^{\mu} \rangle_{\kappa} \cdots \rangle_{\ell+1} = f(y_{\mu,\sigma_1}^{(1)}, \dots, y_{\mu,\sigma_\ell}^{(\ell)}) \quad (1.22)$$

$$M_{\sigma}^{\nu} = \frac{1}{|C_{\sigma}|} \sum_{i \in C_{\sigma}} \xi_i^{\nu} \simeq \langle \cdots \langle \xi^{\nu} \rangle_{\kappa} \cdots \rangle_{\ell+1} = f(y_{\nu,\sigma_1}^{(1)}, \dots, y_{\nu,\sigma_\ell}^{(\ell)}) \quad (1.23)$$

e portanto

$$M_{\sigma}^{\mu} = M_{\sigma}^{\nu} \quad (1.24)$$

onde $|C_{\sigma}|$ é o número de sítios pertencentes a partição C_{σ} com $\sigma = 1, \dots, \prod_{k=1}^{\ell} S_k$. Observe que, igualar a soma sobre sítios à média configuracional sobre algumas das gerações somente é possível, já que a ℓ -upla $y_{\mu,\sigma_1}^{(1)} \cdots y_{\mu,\sigma_\ell}^{(\ell)}$ é constante para estes sítios.

Vemos portanto que, a categorização hierárquica está associada com a existência de vários níveis de informação representados pelas magnetizações parciais de blocos ou partições da rede. Quanto mais recente for o ancestral comum entre dois padrões, menores serão os blocos onde estes terão a mesma magnetização e, então, mais parecidos eles serão.

O modelo de Dotsenko considera que estas magnetizações também podem carregar informações, e propõe uma estrutura de camadas de redes de neurônios onde, em cada uma delas, as unidades fundamentais de informação (bits) são as magnetizações parciais de partições da rede da camada anterior. A rede maior, na primeira camada, possui N neurônios divididos em N_1 grupos cada um contendo $|C^{(1)}|$ neurônios. A segunda camada terá N_1 neurônios cada um deles tendo seu estado determinado por

$$S_{\sigma}^{(2)} = \text{sgn} \left\{ \frac{1}{|C^{(1)}|} \sum_{i \in C_{\sigma}^{(1)}} S_i^{(1)} \right\} ,$$

e assim por diante. A regra de aprendizagem usada em cada etapa é a regra de Hebb aplicada às magnetizações parciais. O autor sugeriu que desta forma se possa guardar um número exponencial de padrões; isso, no entanto, parece não se confirmar na prática.

1.4.3 Modelos com Estrutura Espacial

Uma outra classe das variações do modelo de Hopfield, que guardam certa semelhança com o modelo que introduziremos nos capítulos seguintes, é a dos modelos que apresentam estrutura espacial.

As principais motivações deste estudo são:

- Do ponto de vista tecnológico, sistemas completamente conectados ou sistemas cujas conexões são aleatórias [27] são pouco interessantes. Isso se deve ao fato de serem difíceis de implementação prática. Um dispositivo modesto, com 1000 neurônios, teria em torno de meio milhão de ligações e já representaria um complicado problema de fiação.

Outra vantagem que teria uma rede com conectividade mais local seria ter um tempo menor de comunicação entre suas unidades.

- Do ponto de vista biológico, é sabido que a topologia das ligações entre neurônios é muito mais local do que de longo alcance, apesar de ligações deste tipo também ocorrerem [28]. Em diversas regiões do sistema nervoso, inclusive no cortex cerebral humano que pode ser considerada a mais complexa delas, observa-se que os neurônios se agrupam em estruturas de subcircuitos que são recorrente em escala. Isso significa que podemos dividi-las em vários níveis. No primeiro deles teríamos uma grande quantidade de pequenos circuitos constituídos de um certo número de neurônios, no segundo um número menor de circuitos similares cujas unidades básicas agora seriam circuitos de primeiro nível e etc. Isto é muito semelhante a proposta de Dotsenko para as suas diversas camadas [29] [30].

Quanto a conectividade, as redes biológicas também estão longe de serem de curto alcance, isto é bom; pois, em geral, problemas deste tipo são matematicamente intratáveis.

Em mamíferos terrestres estima-se [31] que volume da fiação no cortex escala com o

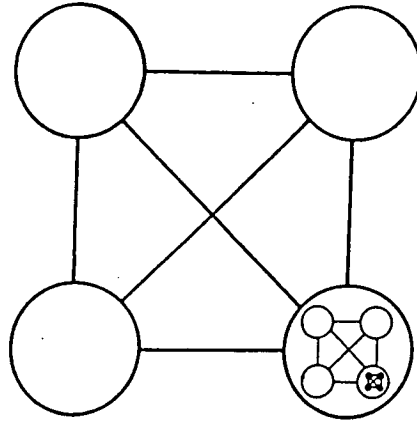


Figura 1.7: *Figura esquemática da organização de subcircuitos no córtex.*

número de neurônios como $N^{4/3}$. Comparando esta estimativa com os resultados limites N^1 , para estrito curto alcance, e N^2 , para uma rede completamente conectada, vemos que, apesar de não ser de curto alcance, a conectividade real ainda está mais próxima deste do que do outro extremo. Os números típicos para o córtex humano são $\sim 10^{10}$ neurônios com em média $\sim 10^4$ ligações sinápticas em cada um.

Os modelos que incorporam estruturas espaciais na matriz sináptica em geral possuem como estados estacionários estados relacionados entre si hierarquicamente. Na maioria dos casos são tratados usando técnicas de dinâmica de sistemas estocásticos ou então simulação numérica. Alguns dos principais deles são:

Modelo Assimétrico Diluído [27]

A regra de aprendizagem é definida como :

$$J_{ij} = a_{ij} \frac{1}{N} \sum_{\mu=1}^P \xi_i^{\mu} \xi_j^{\mu} \quad . \quad (1.25)$$

Onde a_{ij} é uma variável aleatória que pode assumir os valores 0 ou 1 dependendo se a ligação dirigida $i \rightarrow j$ entre os neurônios i e j existe ou não. A existência da assimetria significa que $a_{ij} \neq a_{ji}$. A distribuição para estas variáveis é :

$$P(a_{ij}) = a \delta(a_{ij} - 1) + (1 - a) \delta(a_{ij}) \quad . \quad (1.26)$$

A capacidade crítica de armazenamento para este modelo é $\alpha_c = \frac{2}{\pi}$ obtida para uma extrema diluição caracterizada por $a < \log N/N$.

Modelo de Canning-Gardner [28]

Este modelo possui a mesma regra de aprendizagem do anterior só que desta vez os a_{ij} não são aleatórios nem assimétricos. De forma que muitos resultados podem ser obtidos com o uso da mecânica estatística de equilíbrio. A matriz \underline{a} define a arquitetura do sistema, como antes, cada um de seus elementos pode ser 0 ou 1.

Os dois parâmetros importantes do problema são a dimensão d , determinada por \underline{a} e a razão de conectividade w definida como :

$$w = \frac{1}{N} \sum_1 a_{ij} \quad (1.27)$$

Os resultados básicos deste modelo são que o armazenamento crítico cresce na medida em que diminuímos w ou aumentamos d . O valor máximo de α_c é $\frac{2}{\pi}$ e corresponde as escolhas de $w = 0$ e $d = \infty$. Neste limite recuperamos o modelo diluído apresentado anteriormente.

Capítulo 2

Número finito de padrões.

Nas duas seções finais do capítulo anterior discutimos a importância de propriedades como estrutura espacial e correlação hierárquica entre memórias em uma rede de neurônios atratora. Acredita-se que a introdução correta destes dois ingredientes resultem em um considerável avanço no desempenho destes sistemas. As razões para esta crença são basicamente duas:

- É fato bem documentado [30] [32] que na região do córtex de mamíferos, existe uma estrutura bem definida para ligações e disposição espacial de neurônios. Para estes sistemas, a hipótese de conexão entre todos os neurônio é uma extrapolação grosseira.

- O estudo da quebra de simetria de réplicas no modelo de Sherrington-Kirkpatrick mostra que seus estados estacionários apresentam naturalmente uma estrutura hierárquica de superposições [23]. O número destes estados cresce exponencialmente com o tamanho do sistema. A maioria deles é meta-estável mas podem ser considerados como estáveis, devido a existência de barreiras de ordem $o(N)$ [33] separando-os dos verdadeiros estados de mínima energia. Se supõe que tais características sejam comuns a quaisquer sistemas que apresentem desordem e frustração, e em particular a redes de neurônios. (Durante a defesa deste trabalho fiquei sabendo que existe uma polêmica em torno deste ponto. O Dr. Daniel Fisher, de Harvard, advoga que sistemas de Ising, com ligações desordenadas entre primeiros vizinhos, possuem somente um mínimo na energia livre. Ele tem essa sua idéia apoiada por resultados de simulações numéricas nestes sistemas. Esse fato, no

entanto, não vem a enfraquecer o argumento acima, já que estamos trabalhando com Redes de Neurônios, que são sistemas com interações de longo alcance e que possuem, comprovadamente, um grande número de estados fundamentais.)

Neste capítulo introduziremos um modelo alternativo apresenta, de uma só vez, estas duas características. Não se tem a pretensão de resolver o problema completamente, nem ao menos acreditamos que a estrutura espacial proposta seja a descrição mais fiel da realidade. Cremos, sim, que este modelo, devido a sua tratabilidade, possa permitir um conhecimento maior de como a estrutura de ligações pode influenciar as grandezas que determinam o desempenho dos modelos de memória associativa. Apesar de nossa arquitetura ser de alguma forma bem particular, muitos dos resultados aqui derivados podem ser estendidos para outros arranjos.

2.1 O Modelo

O modelo que propomos é resultado da combinação da estrutura hierárquica de ligações do modelo Ferromagnético de Dyson com a regra de aprendizagem de Hebb.

No modelo de Dyson [35] os N spins da rede estão dispostos nos extremos de uma árvore hierárquica onde cada ancestral tem somente dois descendentes na próxima geração. Assim, para um número N de galhos finais, deveremos ter $\kappa = \ln N / \ln 2$ gerações. Todas as interações entre spins são ferromagnéticas e sua intensidade depende da distância entre eles na árvore hierárquica.

A aplicação da estrutura de interações deste modelo a uma rede de neurônios se dá da seguinte forma: Considere uma rede de N neurônios representados pelas variáveis de Ising $S_i = \pm 1, i = 1, \dots, N$ e um conjunto de padrões ortogonais $\xi_i^\mu, \mu = 1, \dots, p$.

i) No primeiro nível ($k=0$), que corresponde a r -ésima geração, a rede é dividida em $\ell = 2^r$ blocos de $N_0 = \frac{N}{\ell}$ neurônios e definimos as variáveis de bloco

$$S_{a_0}^{\mu(0)} = \sum_{i \in a_0} \xi_i^\mu S_i, \quad a_0 = 1, 2, 3 \dots \ell \quad (2.1)$$

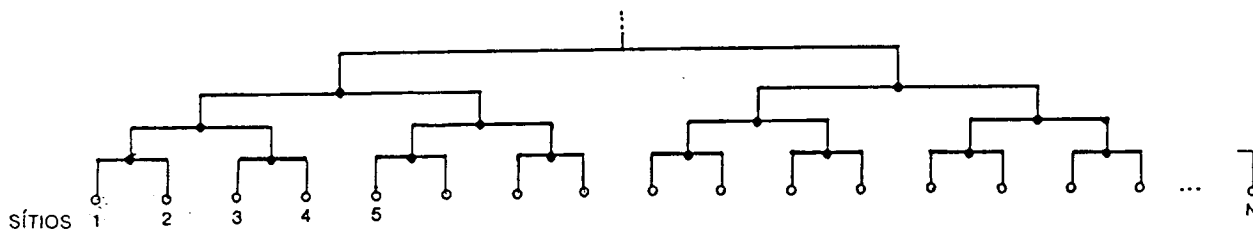


Figura 2.1: Hierarquia de interações entre sítios no Modelo Ferromagnético de Dyson. A intensidade da ligação ferromagnética entre dois spins é dada por uma soma onde cada termo corresponde a contribuição de um de seus ancestrais comuns (pontos escuros). Essa contribuição é ε^{r-k} onde k é o nível em que se encontra o ancestral e $r+1$ é o número total de níveis.

junto com a energia de interação dentro de cada bloco

$$\mathcal{H}(0) = -\frac{J}{N} \varepsilon^r \sum_{a_0=1}^{\ell} \sum_{\mu} [S_{a_0}^{\mu(0)}]^2 \quad (2.2)$$

onde ε é positivo e arbitrário.

ii) No segundo nível ($k=1$), geração $r-1$, cada dois blocos consecutivos são unidos em um bloco maior de $N_1 = 2N_0$ sítios, continuando este processo atinge-se o nível k com $\ell/2^k$ blocos com $N_k = N_0 2^k$ sítios cada. As variáveis de bloco para as diversas gerações são definidas recursivamente por :

$$S_{a_k}^{\mu(k)} = S_{2a_k-1}^{\mu(k-1)} + S_{2a_k}^{\mu(k-1)} = \sum_{i \in a_k} \xi_i^{\mu} S_i \quad (2.3)$$

$$a_k = 1, 2, \dots, \ell/2^k, \quad k = 0, 1, 2, \dots, r$$

e para cada partição associamos a energia:

$$\mathcal{H}(k) = -\frac{J}{N} \varepsilon^{r-k} \sum_{a_k=1}^{2^{r-k}} \sum_{\mu} [S_{a_k}^{\mu(k)}]^2 \quad (2.4)$$

O processo acaba quando $k = r$ e todos os sítios estão contidos em um único bloco do tamanho da rede. A energia total da rede é obtida somando as diversas energias de blocos

$$\mathcal{H} = \sum_{k=0}^r \mathcal{H}(k) \quad . \quad (2.5)$$

Se escrevermos a energia de interação de cada nível em termos das variáveis de bloco do primeiro nível obtemos

$$\begin{aligned} \mathcal{H}(0) &= -\frac{J}{N} \sum_{\mu} \varepsilon^r \{ (S_1^{\mu(0)})^2 + (S_2^{\mu(0)})^2 + \dots + (S_{\ell}^{\mu(0)})^2 \} \\ \mathcal{H}(1) &= -\frac{J}{N} \sum_{\mu} \varepsilon^{r-1} \{ (S_1^{\mu(0)} + S_2^{\mu(0)})^2 + (S_3^{\mu(0)} + S_4^{\mu(0)})^2 + \dots + (S_{\ell-1}^{\mu(0)} + S_{\ell}^{\mu(0)})^2 \} \\ &\dots \end{aligned}$$

até finalmente

$$\mathcal{H}(r) = -\frac{J}{N} \sum_{\mu} (S_1^{\mu(0)} + S_2^{\mu(0)} + S_3^{\mu(0)} + \dots + S_{\ell}^{\mu(0)})^2 \quad . \quad (2.6)$$

Assim podemos escrever o Hamiltoniano (2.5) totalmente em termos destes pequenos blocos como

$$\mathcal{H} = -\frac{N}{\ell} \sum_{a,b} \frac{A_{ab}(\ell)}{\ell} \sum_{\mu} S_a^{\mu} S_b^{\mu} \quad (2.7)$$

os índices a, b são índices de blocos variando de 1 a ℓ e definimos $S_a^{\mu} \equiv S_a^{\mu(0)}/N_0$, ou seja,

$$S_a^{\mu} = \frac{1}{N_0} \sum_i^{(a)} S_i \xi_i^{\mu} \quad (2.8)$$

onde $\sum_i^{(a)}$ indica que a soma em i , índice de neurônio, é restrito ao bloco a .

Os elementos A_{ab} são resultado da soma dos coeficientes correspondentes a cada um dos ancestrais comuns a sítios pertencentes aos blocos "a" e "b". Estes elementos definem uma matriz $\ell \times \ell$ determinada a partir da seguinte relação de recorrência:

$$\underline{\underline{A}}(\ell) = J \begin{bmatrix} \varepsilon \underline{\underline{A}}(\frac{\ell}{2}) + \underline{\underline{U}}(\frac{\ell}{2}) & \underline{\underline{U}}(\frac{\ell}{2}) \\ \underline{\underline{U}}(\frac{\ell}{2}) & \varepsilon \underline{\underline{A}}(\frac{\ell}{2}) + \underline{\underline{U}}(\frac{\ell}{2}) \end{bmatrix} \quad . \quad (2.9)$$

Onde $A(1) = 1$, $U_{ab} = 1$ para qualquer ℓ e J define uma normalização para $\underline{A}(\ell)$. Esta normalização será omitida daqui para frente sendo somente lembrada quando necessário.

Para facilitar a interpretação da função energia (2.7) como sendo o hamiltoniano de um sistema de redes de neurônios é conveniente explicitarmos as suas variáveis de sítios colocando-a na forma

$$\mathcal{H} = -\frac{1}{2} \sum_{a,b} A_{ab} \sum_i^{(a)} \sum_j^{(b)} J_{ij} S_i S_j \quad (2.10)$$

onde J_{ij} é a regra de aprendizagem de Hebb

$$J_{ij}^H = \frac{1}{N} \sum_{\mu=1}^P \xi_i^\mu \xi_j^\mu . \quad (2.11)$$

A regra de aprendizagem em um modelo de rede de neurônios indica como a eficiência sináptica varia com a memorização de uma certa palavra, ou padrão. Cada vez que uma nova configuração $\{\xi^{\mu_0}\}$ é aprendida o sistema passa por uma transformação de suas sinapses

$$J_{ij} \rightarrow J'_{ij}$$

Esta regra é local se cada elemento J_{ij} depende exclusivamente do que ocorre nos sítios “i” e “j”, e é dita aditiva se a transformação se dá como

$$J_{ij} \rightarrow J_{ij} + \Delta J_{ij}^{\mu_0}$$

onde $\Delta J_{ij}^{\mu_0}$ além de depender dos detalhes do padrão novo, isto é os valores $\xi_i^{\mu_0}$ e $\xi_j^{\mu_0}$, pode também depender separadamente dos sítios, como no caso de um modelo com estrutura espacial, ou ainda da ordem de apresentação dos padrões [36], ou seja, o valor $\mu = \mu_0$.

Essa dependência de ΔJ_{ij} com detalhes específicos dos padrões faz que esta não seja uma boa quantidade para descrever propriedades do sistema. Assim definimos uma nova quantidade, mais conveniente, denominada “intensidade de aquisição de padrões”, dada por

$$K(\mu_0) = \langle (\Delta J_{ij}^{\mu_0})^2 \rangle_{\xi} - \langle \Delta J_{ij}^{\mu_0} \rangle_{\xi}^2 , \quad (2.12)$$

onde $\langle \rangle_{\xi}$ indica a média sobre as variáveis ξ . Se considerarmos o modelo de Hopfield, este valor é

$$K(\mu) = \langle (\Delta \mathcal{J}_{ij}^{\mu})^2 \rangle_{\xi} = \langle (\Delta J_{ij}^H)^2 \rangle_{\xi} = \frac{1}{N^2} \quad (2.13)$$

ou seja, independente de i e j e μ . O fato de não haver dependência em μ nos indica que todas as memórias são equivalentes, são igualmente bem aprendidas ou esquecidas pelo sistema. A ausência de dependência em sítios indica falta de estrutura espacial, ou seja todas as ligações entre sítios são igualmente importantes. Isso também se passa com a maioria dos modelos que armazenam padrões hierarquicamente correlacionados, discutidos no capítulo um.

No caso do nosso modelo, definido por (2.10), temos como em um caso típico de modelo com estrutura espacial:

$$\Delta \mathcal{J}_{ij} = A_{ab} \Delta J_{ij}^H \quad . \quad (2.14)$$

Com $i \in a$ e $j \in b$, e portanto

$$K(\mu) = \langle \Delta(\mathcal{J}_{ij})^2 \rangle_{\xi} = A_{ab}^2 \left(\frac{1}{N} \right)^2 \quad . \quad (2.15)$$

Ou seja, existe uma dependência explícita com a estrutura de blocos através da matriz \underline{A} . Isso não caracteriza exatamente uma estrutura microscópica para rede pois estamos trabalhando com blocos grandes.

Aqui uma importante observação deve ser feita relativamente ao tamanho destes blocos. O modelo original de Dyson [35] foi proposto para simular as propriedades de uma cadeia unidimensional de spins com interações ferromagnéticas que decaíam com a distância como

$$\mathcal{J}_{ij} \sim (i - j)^{-\sigma},$$

onde σ é um parâmetro positivo real, que determina o alcance da interação. Esta equivalência somente ocorre quando consideramos, como foi dito anteriormente, um número infinito de gerações e portanto um número infinito de blocos, cada um com um número finito de spins (ou neurônios). No nosso caso, no entanto, é fundamental

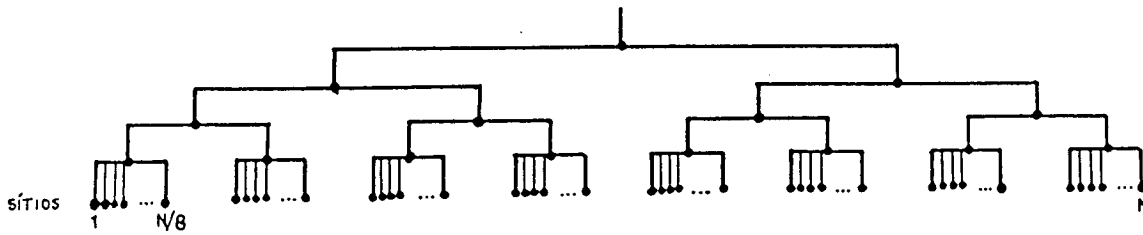


Figura 2.2: Adaptação da estrutura de ligações do modelo de Dyson para o caso de uma rede neuronal. O processo hierárquico é truncado de modo que tenhamos um número finito de gerações, e portanto o penúltimo ancestral fique com um número extensivo de descendentes. A figura mostra o caso especial onde o sistema se divide em 8 blocos.

que tenhamos blocos macroscópicos, isto é, com um número infinito de sítios, para que possamos usar o método desenvolvido por Amit et al. [14], onde a solução exata do problema é obtida diretamente da solução de ponto de sela (campo médio). Assim, o que estamos guardando do modelo de Dyson é somente a estrutura hierárquica de arranjo dos sítios, truncada em um certo número finito de gerações. Sendo assim, não podemos atribuir nenhuma dimensão finita a esta rede que, como as outras redes de neurônios que apresentam solução exata, continua sendo um sistema ∞ -dimensional.

2.2 Propriedades da Matriz $\underline{\underline{A}}(\ell)$

ULTRAMETRICIDADE - Se definirmos a distância entre dois elementos localizados nos extremos de uma árvore hierárquica, como a da figura 2.2, como sendo função da geração do ancestral comum mais jovem que eles possuem, de forma que quanto mais antigo tal

Um modo de construirmos uma destas bases é através da relação de recorrência

$$\bar{v}^{2\eta-1}(\ell) = \begin{bmatrix} \bar{v}^\eta(\frac{\ell}{2}) \\ \bar{v}^\eta(\frac{\ell}{2}) \end{bmatrix}, \quad \bar{v}^{2\eta}(\ell) = \begin{bmatrix} \bar{v}^\eta(\frac{\ell}{2}) \\ -\bar{v}^\eta(\frac{\ell}{2}) \end{bmatrix}, \quad (2.19)$$

$$\eta = 1, 2, \dots, \frac{\ell}{2}$$

que tem como “condição inicial” $\bar{v}^1(1) = 1$.

Para o caso da matriz (2.9) os autovetores gerados por este processo são

$$\begin{aligned} \bar{v}^1 &= (+, +, +, +, +, +, +, +) \\ \bar{v}^2 &= (+, +, +, +, -, -, -, -) \\ \bar{v}^3 &= (+, +, -, -, +, +, -, -) \\ \bar{v}^4 &= (+, +, -, -, -, -, +, +) \\ \bar{v}^5 &= (+, -, +, -, +, -, +, -) \\ \bar{v}^6 &= (+, -, +, -, -, +, -, +) \\ \bar{v}^7 &= (+, -, -, +, +, -, -, +) \\ \bar{v}^8 &= (+, -, -, +, -, +, +, -) \end{aligned} \quad (2.20)$$

onde $+, - \equiv +1, -1$.

Para calcularmos as relações de recorrência correspondentes aos autovalores devemos observar que a matriz $\underline{U}(\ell)$ é diagonalizada pelos mesmos autovetores de $\underline{A}(\ell)$. Isto se deve ao fato que de \underline{U} ser um caso especial de \underline{A} ($\varepsilon = 0$). Os seus autovalores são $\lambda = \ell$ com multiplicidade 1 e $\lambda' = 0$ com multiplicidade $\ell - 1$. Assim, para uma dada dimensão haverá somente um vetor com autovalor não nulo para esta matriz. Este vetor é

$$\bar{v}^1(\ell) = (+, +, \dots, +)$$

para qualquer outro \bar{v}^η , com $\eta \neq 1$, teremos

$$\underline{U} \cdot \bar{v}^\eta = 0 \quad .$$

Deste modo, obtemos as seguintes equações de recorrência para os autovalores:

$$\begin{aligned}\lambda_1(\ell) &= \ell + \varepsilon \lambda_1\left(\frac{\ell}{2}\right) \\ \lambda_2(\ell) &= \varepsilon \lambda_1\left(\frac{\ell}{2}\right) \\ \lambda_{2\eta-1}(\ell) &= \lambda_{2\eta}(\ell) = \varepsilon \lambda_\eta\left(\frac{\ell}{2}\right), \quad \eta \geq 2\end{aligned}\tag{2.21}$$

Podemos ver que, para uma dada dimensão $\ell > 2$ sempre haverá degenerescência destes autovalores. O número de autovalores diferentes pode ser facilmente calculado de (2.21) e é $r + 1$ onde $\ell = 2^r$, e se ordenam como:

$$\lambda_1 > \lambda_2 > \lambda_3 = \lambda_4 > \dots > \lambda_{\frac{\ell}{2}+1} = \lambda_{\frac{\ell}{2}+2} = \dots = \lambda_\ell\tag{2.22}$$

O fato de haver uma degenerescência indica que os vetores que saem de (2.19) não são os únicos possíveis autovetores de \underline{A} . Tome como exemplo o vetor

$$\bar{w}(\ell) = \begin{bmatrix} \bar{v}^\eta\left(\frac{\ell}{2}\right) \\ \bar{v}^\vartheta\left(\frac{\ell}{2}\right) \end{bmatrix}\tag{2.23}$$

onde \bar{v}^η , \bar{v}^ϑ e \bar{w} possuem o mesmo autovalor $\lambda = \lambda_\eta = \lambda_\vartheta$.

2.3 Cálculo da energia livre

Para um número finito de padrões armazenados nós seguimos o procedimento usual.

O nosso objetivo é calcular a energia livre através de

$$f = -\frac{1}{\beta N} \langle \ln Z \rangle_\xi, \tag{2.24}$$

os colchetes $\langle \dots \rangle_\xi$ significam a média configuracional tomada sobre os padrões $\{\xi_i^\mu\}$ e β é o inverso da temperatura. A função de partição é dada por

$$Z = Tr_{\{s_i\}} \exp \left(\frac{\beta}{2N} \sum_{a,b} A_{ab} \sum_{i,j}^{(a),(b)} \sum_{\mu}^p \xi_i^\mu \xi_j^\mu S_i S_j \right)\tag{2.25}$$

Usando a transformação de Hubbard-Stratonovich [39]

$$\begin{aligned} & \exp \left\{ \frac{\beta N}{2} \sum_{a,b} \frac{A_{ab}}{\ell^2} \left(\sum_i^{(a)} \frac{\ell}{N} \xi_i^\mu S_i \right) \left(\sum_j^{(b)} \frac{\ell}{N} \xi_j^\mu S_j \right) \right\} = \\ & \sqrt{|\underline{A}|} \left(\frac{\sqrt{\pi} \beta N}{\ell} \right)^\ell \int \prod_a dm_a^\mu \times \\ & \exp \left\{ -\beta N \left(\frac{1}{2} \sum_{a,b} \frac{A_{ab}}{\ell^2} m_a^\mu m_b^\mu + \sum_{a,b} \frac{A_{ab}}{\ell^2} \left(\sum_i^{(a)} \frac{\ell}{N} \xi_i^\mu S_i \right) m_b^\mu \right) \right\} \end{aligned} \quad (2.26)$$

podemos linearizar a dependência nas variáveis S_i , e ao mesmo tempo introduzir o parâmetro do ordem de bloco [40]

$$m_a^\mu = \frac{\ell}{N} \sum_i^{(a)} \langle \xi_i^\mu \langle S_i \rangle_T \rangle_\xi$$

A linearização simplifica a aplicação do traço resultando que

$$\begin{aligned} Z &= \left[\sqrt{|\underline{A}|} \left(\frac{\sqrt{\pi} \beta N}{\ell} \right)^\ell \right]^p \int \left(\prod_{a,\mu} dm_a^\mu \right) \times \\ & \exp \left\{ -\beta N_0 \frac{1}{2} \sum_{a,b} \frac{A_{ab}}{\ell} \bar{m}_a \bar{m}_b \right\} \times \\ & \exp \left\{ -N_0 \sum_a \frac{1}{N_0} \sum_i^{(a)} \ln(2 \cosh(\beta \sum_b \frac{A_{ab}}{\ell} \bar{m}_b \cdot \vec{\xi}_i)) \right\} \end{aligned} \quad (2.27)$$

Note que podemos escrever a soma sobre sítios dentro do expoente como

$$\frac{1}{N} \sum_i g(\vec{\xi}_i) = \frac{1}{N} \sum_{\{\vec{\xi}\}} \eta(\vec{\xi}) g(\vec{\xi}) \quad ,$$

onde a segunda soma é sobre todas as possíveis configurações que podemos ter de ξ^μ , $\mu = 1, \dots, p$, e a função $\eta(\vec{\xi})$ indica o número de vezes que uma dada destas configurações aparece na soma em i . A probabilidade de ocorrer uma destas configurações é

$$P(\vec{\xi}) = \frac{1}{2^p} \quad ,$$

e a probabilidade da configuração escolhida ocorrer n vezes, ou seja $\eta(\vec{\xi}) = n$ é

$$P(\vec{\xi}, n) = \frac{N!}{(N-n)!n!} \left(\frac{1}{2^p} \right)^n \left(1 - \frac{1}{2^p} \right)^{N-n} \quad .$$

O valor médio e a dispersão de n são

$$\begin{aligned} \langle n \rangle &= \frac{N}{2^p} \\ \langle n^2 \rangle - \langle n \rangle^2 &= \frac{N}{2^p} \left(1 - \frac{1}{2^p}\right) \end{aligned} \quad ,$$

de modo que no limite de $N \rightarrow \infty$ temos

$$\frac{\sqrt{\langle n^2 \rangle - \langle n \rangle^2}}{\langle n \rangle} \rightarrow 0 \quad .$$

Isto indica que neste limite esta distribuição é fortemente centrada em torno de seu valor médio o que nos permite escrever

$$\frac{1}{N} \sum_i g(\vec{\xi}_i) \simeq \frac{1}{2^p} \sum_{\{\vec{\xi}\}} g(\vec{\xi}) = \langle g(\vec{\xi}) \rangle_\xi \quad .$$

Nesta situação as variáveis ξ_i^μ são ditas serem auto-mediáveis, pois sua soma em sítios é equivalente a média configuracional. Observe que a condição para que isto ocorra é que

$$\lim_{N \rightarrow \infty} \frac{2^p}{N} = 0 \quad ,$$

que é justamente o caso se p é finito.

Assim na equação (2.27) podemos substituir

$$\frac{1}{N_0} \sum_i^{(a)} \ln(2 \cosh(\beta \sum_b \frac{A_{ab}}{\ell} \vec{m}_b \cdot \vec{\xi}_i))$$

por

$$\langle \ln(2 \cosh(\beta \sum_b \frac{A_{ab}}{\ell} \vec{m}_b \cdot \vec{\xi}_i)) \rangle_\xi \quad ,$$

de modo que a função de partição Z passa a ser independente de $\{\xi_i^\mu\}$ e, portanto, o sinal de média em (2.24) desnecessário. No limite em que $N_0 \rightarrow \infty$ a função Z pode ser substituída por seu valor no ponto de sela e, então, a energia livre fica:

$$\begin{aligned} f &= \frac{1}{2} \frac{1}{\ell} \sum_{a,b} \frac{A_{ab}}{\ell} \vec{m}_a \cdot \vec{m}_b \\ &\quad - \frac{1}{\beta} \frac{1}{\ell} \sum_a \left\langle \ln \cosh \left(\frac{\beta}{\ell} \sum_b A_{ab} \vec{\xi}_i \cdot \vec{m}_b \right) \right\rangle_\xi \quad . \end{aligned} \quad (2.28)$$

As quantidades m_a^μ com $\mu = 1, \dots, p$ e $a = 1, \dots, \ell$ são os parâmetros de ordem do sistema, representando o produto escalar do estado estacionário do sistema com um dado padrão μ , dentro do bloco a , ou seja ,

$$m_a^\mu = \frac{1}{N_0} \sum_{i \in a} \xi_i^\mu \langle \sigma_i \rangle_T = \langle \xi_i^\mu \langle \sigma_i \rangle_T \rangle_\xi, \quad (2.29)$$

onde foi usada a notação $\langle \dots \rangle_T$ para indicar a média térmica.

Os vários estados do sistema são determinados pelas soluções das equações para $\{m_a^\mu\}$ no ponto de sela

$$m_a^\mu = \left\langle \xi_i^\mu \tanh \left(\frac{\beta}{\ell} \sum_b A_{ab} \xi_i \vec{m}_b \right) \right\rangle_\xi, \quad (2.30)$$

e a estabilidade destas soluções pelos autovalores da matriz $p\ell$ -dimensional de elementos

$$M_{ab}^{\mu\nu} = \frac{\partial^2 f}{\partial m_a^\mu \partial m_b^\nu}, \quad (2.31)$$

ou seja,

$$M_{ab}^{\mu\nu} = \left(\frac{A_{ab}}{\ell^2} - \frac{\beta}{\ell^3} \sum_c A_{ac} A_{cb} \right) \delta^{\mu\nu} + \frac{\beta}{\ell^3} \sum_c A_{ac} A_{cb} \left\langle \xi_i^\mu \xi_i^\nu \tanh^2 \left[\frac{\beta}{\ell} \sum_d A_{cd} \xi_i \vec{m}_d \right] \right\rangle_\xi \quad (2.32)$$

2.4 Soluções que envolvem um único padrão

Considerando $m_a^\mu = m_a \delta^{1\mu}$, obteremos para (2.28) , (2.30) e (2.32).

$$f = \frac{1}{2\ell} \sum_{a,b} \frac{A_{ab}}{\ell} m_a m_b - \frac{1}{\beta\ell} \sum_a \ln 2 \cosh \left(\frac{\beta}{\ell} \sum_b A_{ab} m_b \right) \quad (2.33)$$

$$m_a = \tanh \left[\frac{\beta}{\ell} \sum_b A_{ab} m_b \right] \quad (2.34)$$

$$M_{ab}^{\mu\nu} = \left(\frac{A_{ab}}{\ell^2} - \frac{\beta}{\ell^3} \sum_c A_{ac} A_{cb} (1 - m_c^2) \right) \delta^{\mu\nu} \quad (2.35)$$

Dentro deste tipo de soluções, podemos considerar um caso mais simples ainda, que é quando m_a é proporcional a um autovetor de \underline{A} .

2.4.1 Soluções onde \bar{m} é autovetor de \underline{A} .

No caso em que \bar{m} é autovalor de \underline{A} , ou seja,

$$m_a^\gamma = m_\gamma v_a^\gamma \quad , \quad (2.36)$$

onde

$$\underline{A} \cdot \bar{v}^\gamma = \lambda_\gamma \bar{v}^\gamma \quad ,$$

substituindo (2.36) em (2.34), obtemos o seguinte conjunto de equações :

$$m_\gamma v_a^\gamma = \tanh \left[\beta \frac{\lambda_\gamma}{\ell} m_\gamma v_a^\gamma \right] \quad , \quad (2.37)$$

com $a = 1, \dots, \ell$.

Este sistema tem solução se $|v_a^\gamma| = \text{constante}$ para todo a . Neste caso,

$$m_\gamma = \tanh \left[\beta \frac{\lambda_\gamma}{\ell} m_\gamma \right] \quad (2.38)$$

dá como m_γ varia com a temperatura.

A condição de estabilidade deste tipo de solução vem de diagonalizar

$$M_{ab} = \frac{A_{ab}}{\ell^2} - \beta \sum_c \frac{A_{ac} A_{cb}}{\ell^3} (1 - m_\gamma^2 (v_a^\gamma)^2) \quad . \quad (2.39)$$

Considerando $|v_a^\gamma| = 1$, obtemos como elementos diagonais

$$M_\sigma = \frac{\lambda_\sigma}{\ell^2} - \frac{\beta}{\ell} \left(\frac{\lambda_\sigma}{\ell} \right)^2 (1 - m_\gamma^2) \quad , \quad \sigma = 1, \dots, \ell. \quad (2.40)$$

Como sabemos, a condição para que a solução seja estável é $M_\sigma > 0$ para todo σ , e, visto que os valores de M_σ estão ordenados, (de (2.22) $M_1 \leq M_2 \leq \dots \leq M_\ell$) a condição suficiente fica $M_1 > 0$, ou seja,

$$1 - \beta \frac{\lambda_1}{\ell} (1 - m_\gamma^2) > 0 \quad . \quad (2.41)$$

A temperatura crítica T^* para cada uma das soluções $\gamma = 1, \dots, \ell$ é dada pela solução conjunta das duas equações

$$\begin{cases} m_s^* = \tanh \left(\frac{T_s}{T_s^*} m_s^* \right) \\ \frac{T_s}{T_s^*} (1 - m_s^{*2}) = 1 \end{cases} \quad , \quad (2.42)$$

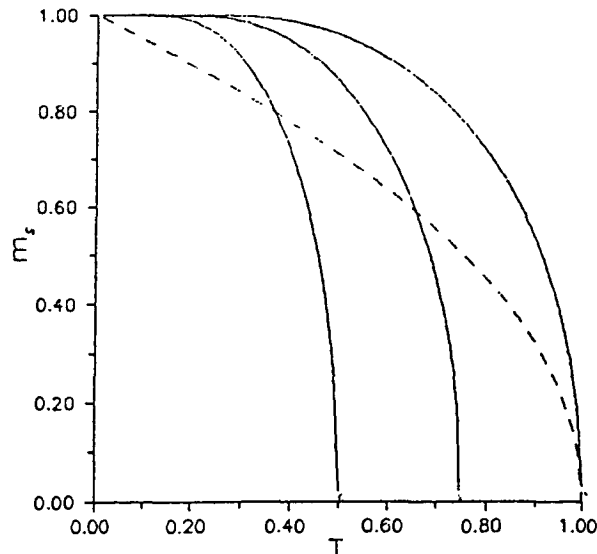


Figura 2.3: O gráfico mostra os valores de m_s como função da temperatura para 3 possíveis valores do autovalor λ_s . A curva pontilhada indica a condição de estabilidade, abaixo da qual estas soluções são instáveis.

onde, para termos em conta a degenerescência dos autovetores de \underline{A} , introduzimos a notação

$$\begin{aligned} T_1 &= \frac{\lambda_1}{\ell} \\ T_s &= \frac{\lambda_s}{\ell}, \quad 2^{s-2} \leq \gamma < 2^{s-1}, \quad s = 2 \dots r+1. \end{aligned} \quad (2.43)$$

Na figura 2.3 mostramos a solução gráfica destas equações para alguns valores de s .

Note que, com excessão de $T_1^* = T_1$, todas as outras temperaturas críticas satisfazem $T_s^* < T_s$. Isso indica que, fora a solução de recuperação ($\gamma = 1$), todas as soluções descendentes surgem descontinuamente quando diminuimos a temperatura a partir de $T = T_1$, além disso, as temperaturas estão ordenadas como

$$T_1 > T_2^* > T_3^* > \dots > T_{r+1}^* ; \quad (2.44)$$

Estes resultados estão presentes em um outro tipo de modelo, que é aquele que envolve armazenamento de padrões com diferentes intensidades [41].

A qualquer temperatura finita, estas soluções são mínimos locais da energia livre dada

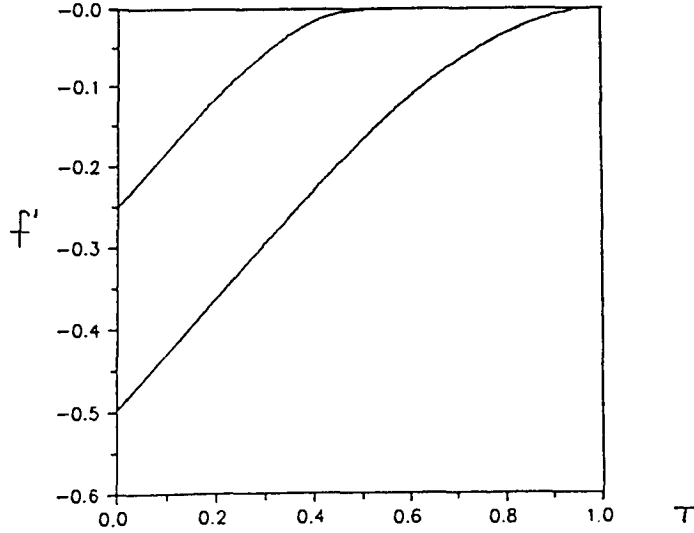


Figura 2.4: Energia livre f'_s para dois valores (de cima para baixo) $\lambda_s/\ell = 0.5, 1$.

por

$$f_s = T_s \left(\frac{m_s^2}{2} - \frac{T}{T_s} \ln \cosh\left(\frac{T_s}{T} m_s\right) \right) - T \ln 2 \quad , \quad (2.45)$$

que pode ser escrita em uma forma de escala como

$$f'_s = T_s g\left(\frac{T_s}{T}\right) \quad , \quad (2.46)$$

onde $f'_s = f_s + T \ln 2$ e $g(T)$ é uma função monotonicamente crescente.

Na figura 2.4 apresentamos um gráfico de $f' \times T$ para duas soluções s^1 e s^2 com $T_{s^1} > T_{s^2}$.

Devido a relação de escala entre elas as curvas não se cruzarão jamais. O resultado disso é um novo ordenamento, só que desta vez envolvendo os valores da energia livre das diversas soluções $f_1 < f_2 < \dots < f_{r+1}$.

Os resultados à temperatura zero são facilmente obtidos de (2.33) ,

$$\frac{E}{N} = -\frac{1}{2\ell} \sum_a \left| \sum_b \frac{A_{ab}}{\ell} m_b \right| \quad , \quad (2.47)$$

onde esta soma satisfaz a desigualdade

$$\frac{1}{\ell} \sum_a \left| \sum_b A_{ab} m_b \right| \leq \left[\sum_a \left(\sum_b A_{ab} m_b \right)^2 \right]^{1/2} \quad . \quad (2.48)$$

O sinal de igual se aplica somente ao caso que \bar{m} é um autovalor de $\underline{\underline{A}}$, e portanto

$$\left| \sum_b A_{ab} m_b \right| = |\lambda_\gamma v_a^\gamma| = \lambda_\gamma, \quad (2.49)$$

de onde se vê que $E_1 < E_2 < \dots < E_{r+1}$.

Nós concluímos então, que as verdadeiras soluções de recuperação, que passaremos a chamar de ancestrais, com $\gamma = 1$ e $v_a^1 = 1$, para todos os valores $T < T_1$ são os mínimos da energia livre. Existem outras solução para $\gamma \geq 2$ que formam uma “ família de descendentes ” que diferem dos padrões originais no sinal relativo das magnetizações de blocos, como mostramos na figura 2.5. Tais estados são mínimos locais da energia livre e possíveis atratores no processo de relaxação.

2.4.2 Soluções onde \bar{m} não é necessariamente um autovetor de $\underline{\underline{A}}$

Considere novamente o sistema de ℓ equações que determinam as soluções de ponto de sela para o caso de recuperação de uma única memória, ou seja, de (2.34)

$$\left\{ m_a = \tanh \left[\frac{\beta}{\ell} \sum_b A_{ab}(\ell) m_b \right] \quad a = 1, \dots, \ell \quad , \quad (2.50) \right.$$

onde voltamos a colocar o argumento (ℓ) na matriz $\underline{\underline{A}}$ para indicar sua dimensão. Como antes, as soluções deste sistema são os vetores $\bar{m} = (m_1, \dots, m_\ell)$. Usando a expressão que relaciona $\underline{\underline{A}}(\ell)$ com matrizes menores, de dimensão $\ell/2$, podemos reescrever o sistema acima como dois conjuntos de $\ell/2$ equações:

$$\left\{ \begin{array}{l} m_a = \tanh \left(\frac{\beta}{\ell} \sum_{b=1}^{\ell/2} (\epsilon A_{ab}(\ell/2) + U_{ab}(\ell/2)) m_b + \sum_{b=\ell/2+1}^{\ell} U_{ab}(\ell/2) m_b \right) \\ a \leq \frac{\ell}{2} \\ \\ m_a = \tanh \left(\frac{\beta}{\ell} \sum_{b=1}^{\ell/2} U_{ab}(\ell/2) m_b + \sum_{b=\ell/2+1}^{\ell} (\epsilon A_{ab}(\ell/2) + U_{ab}(\ell/2)) m_b \right) \\ \frac{\ell}{2} < a \leq \ell \end{array} \right.$$

Estes conjuntos estão acoplados devido a presença dos termos

$$\sum_{b=\ell/2+1}^{\ell} U_{ab}(\ell/2) m_b$$

no primeiro deles, e

$$\sum_{b=1}^{\ell/2} U_{ab}(\ell/2) m_b$$

no segundo.

Excetuando-se as soluções proporcionais aos dois autovetores de autovalores máximo de $\underline{\underline{A}}(\ell)$, que são

$$\bar{m}^1(\ell) = m_1 \begin{pmatrix} + \\ + \\ + \\ \vdots \\ + \\ + \end{pmatrix} \quad \bar{m}^2(\ell) = m_2 \begin{pmatrix} + \\ \vdots \\ + \\ - \\ \vdots \\ - \end{pmatrix} , \quad (2.51)$$

todas as demais soluções de (2.34) podem ser escrita como

$$\bar{m}(\ell) = \begin{pmatrix} \bar{m}'(\frac{\ell}{2}) \\ \bar{m}''(\frac{\ell}{2}) \end{pmatrix} . \quad (2.52)$$

Neste caso, tanto \bar{m}' como \bar{m}'' são tais que

$$\underline{\underline{U}}(\ell/2)\bar{m}(\ell/2) = 0 , \quad (2.53)$$

e, portanto, desacoplam os dois sistemas de (2.51).

Com isso estabelecemos uma relação entre o número de soluções do sistema de ℓ blocos e aquele de um sistema menor de $\ell/2$ blocos. Até agora não se considerou as soluções simétricas em sinal (lembre-se que devido a forma da Regra de Hebb também armazenamos os padrões $\{-\xi^\mu\}$), mas caso forem contadas também, obteremos para o número de soluções

$$N(\ell) = 4 + (N(\ell/2) - 2)^2 , \quad (2.54)$$

com a condição inicial $N(1) = 2$. Ou seja, o número de soluções para o caso ℓ -dimensional é igual a 4, que corresponde as soluções (2.51) e seus simétricos, mais todas as possíveis

combinações de duas soluções do sistema $\frac{\ell}{2}$ -dimensional que sejam autovetores de autovalor nulo de $\underline{U}(\ell/2)$. Como existem duas destas que não satisfazem este requerimento se subtrai 2.

A análise numérica da relação (2.54) mostra que para $\ell \gg 1$ esta pode ser bem representada pelo seguinte comportamento exponencial

$$N(\ell) = e^{0.45\ell} < 2^\ell \quad (2.55)$$

A comparação com 2^ℓ é importante pois este valor corresponde ao caso trivial onde teríamos ℓ sistemas completamente independentes.

O número de soluções aqui calculadas englobam as soluções de recuperação, soluções que são autovetores de $\underline{A}(\ell)$ e outras. Justamente estas outras nos interessa discutir nesta seção, e para isso vamos tomar um exemplo. Considere a seguinte escolha para o vetor \bar{m} :

$$m_c^{\gamma\gamma'} = \begin{cases} m_\gamma v_c^\gamma(\frac{\ell}{2}), & 1 \leq c < \frac{\ell}{2} \\ m_{\gamma'} v_c^{\gamma'}(\frac{\ell}{2}), & \frac{\ell}{2} \leq c \leq \ell \end{cases} \quad (2.56)$$

Esta forma desacopla sistema de ℓ equações de (2.34), transformando-o em um sistema de somente duas equações

$$\begin{cases} m_\gamma = \tanh \left[\beta \frac{e\lambda_\gamma}{\ell} \left(\frac{\ell}{2} \right) m_\gamma \right] , \\ m_{\gamma'} = \tanh \left[\beta \frac{e\lambda_{\gamma'}}{\ell} \left(\frac{\ell}{2} \right) m_{\gamma'} \right] . \end{cases}$$

Nós fixamos arbitrariamente $\lambda_\gamma > \lambda_{\gamma'}$, e então $m_\gamma > m_{\gamma'}$. Na Figura 2.5 temos vários exemplos destas soluções que misturam dois autovalores, que passam a existir para $\ell > 4$.

A matriz de estabilidade torna-se

$$M_{a,b}^{(\gamma\gamma')} = \frac{A_{ab}}{\ell^2} - \frac{\beta}{\ell^3} \sum_c A_{ac} \left[1 - (m_c^{\gamma\gamma'})^2 \right] A_{cb} \quad , \quad (2.57)$$

lembrando que neste caso

$$(m_c^{\gamma\gamma'})^2 = \begin{cases} (m_\gamma)^2, & 1 \leq c < \frac{\ell}{2} \\ (m_{\gamma'})^2, & \frac{\ell}{2} \leq c \leq \ell \end{cases} \quad , \quad (2.58)$$

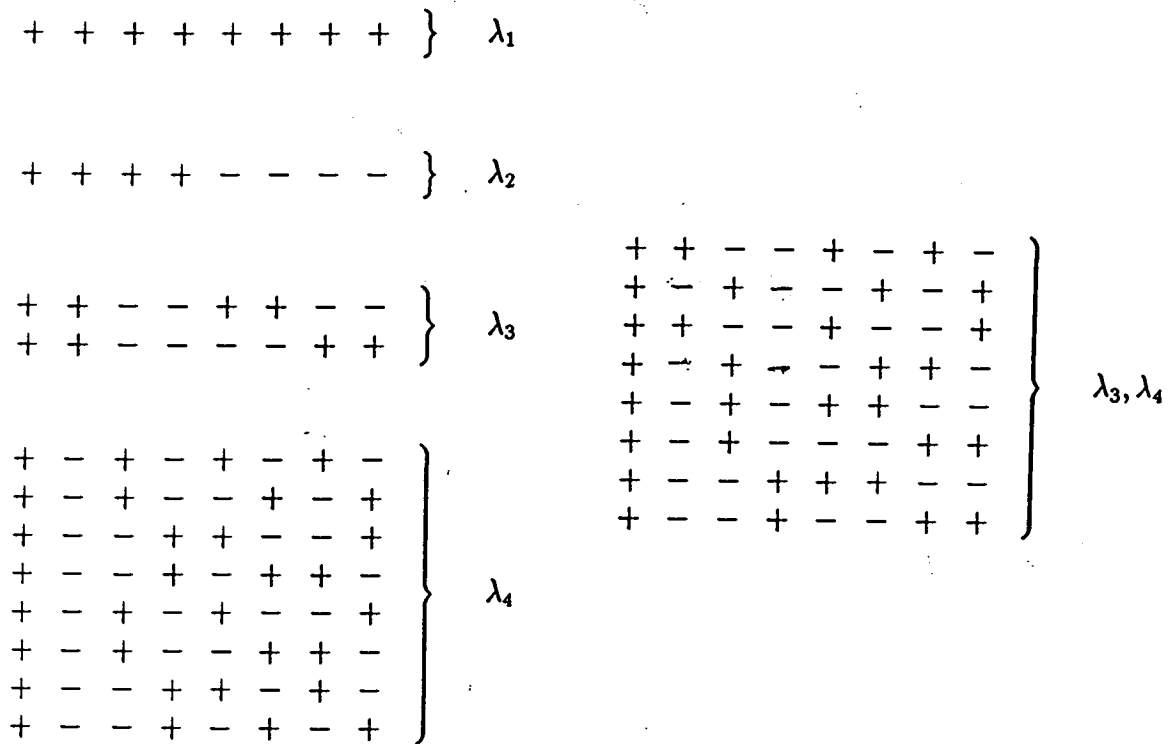


Figura 2.5: Possíveis estados estacionários do sistema com interação hierárquica entre blocos para o caso de $\ell = 8$ a $T = 0$, quando somente um padrão é armazenado. O “+” significa que o bloco está alinhado com a memória e “-” com a anti-memória $\{-\xi^\mu\}$.

o que indica que $\underline{M}^{(\gamma\gamma')}$ não é diagonalizável pela mesma transformação que diagonaliza \underline{A} . Podemos, no entanto, pô-la em uma forma diagonal de blocos através da transformação

$$\sum_{a,b} v_a^{\eta'} M_{a,b}^{(\gamma\gamma')} v_b^\eta = T_{\eta'} \Lambda_{\eta'\eta} \quad , \quad (2.59)$$

com a matriz $\underline{\Lambda}$ na forma

$$\underline{\Lambda} = \begin{bmatrix} \omega_1 & \sigma_{12} & & & & \\ \sigma_{21} & \omega_2 & & & & \\ & & \omega_3 & \sigma_3 & & \\ & & \sigma_3 & \omega_3 & & \\ & & & & \omega_4 & \sigma_4 \\ & & & & \sigma_4 & \omega_4 \\ & & & & & & \omega_4 & \sigma_4 \\ & & & & & & \sigma_4 & \omega_4 \end{bmatrix} \quad , \quad (2.60)$$

onde

$$\Lambda_{\eta\eta} = \omega_s = 1 - \frac{T_s}{T} \left[1 - \frac{m_\gamma^2 + m_{\gamma'}^2}{2} \right] \quad , \quad (2.61)$$

e

$$\Lambda_{\eta'\eta} = \frac{T_s}{T} \left(\frac{m_\gamma^2 - m_{\gamma'}^2}{2} \right) = \sigma_{\eta'\eta} \quad , \quad (2.62)$$

com os índices dados por

$$\begin{aligned} \eta' &= 2\delta - 1, \eta = 2\delta \quad \text{or} \quad \eta' = 2\delta, \eta = 2\delta - 1 \quad , \\ 2^{s-2} + 1 &\leq \eta \leq 2^{s-1} \quad \text{for} \quad s = 2, 3 \dots r \quad , \\ \eta &= 1 \quad \text{for} \quad s = 1 \quad . \end{aligned}$$

De acordo com (2.59), já que $T_\eta \geq 0$, a condição suficiente para a estabilidade das soluções em (2.56) é que os autovalores de $\underline{\Lambda}$ sejam positivos:

$$\begin{aligned} \Lambda_1^\mp &= \frac{1}{2}(\omega_1 + \omega_2) \mp \frac{1}{2} \left\{ (\omega_1 - \omega_2)^2 + 4\sigma_{12}\sigma_{21} \right\}^{\frac{1}{2}} \quad , \\ \Lambda_s^- &= 1 - \frac{\beta}{\beta_s} (1 - m_{\gamma'}^2) \quad , \\ \Lambda_s^+ &= 1 - \frac{\beta}{\beta_s} (1 - m_\gamma^2) \quad , \end{aligned} \quad (2.63)$$

com $\Lambda_1^- < \Lambda_1^+ < \Lambda_3^- < \Lambda_3^+ < \dots$.

Assim as soluções duplamente misturadas são dadas pelo sistema formado pelas duas equações (2.57) acrescido da condição de estabilidade $\Lambda_1^- \geq 0$, já que este é o menor autovalor. A temperatura crítica é obtida igualando este autovalor a zero. Como no caso das soluções que são autovetores, confira a figura 2.3, aqui também existe uma transição descontínua, ou seja, $m_\gamma^* \neq 0$ e $m_{\gamma'}^* \neq 0$ para $T = T_{\gamma\gamma'}^*$, que é a temperatura crítica correspondente a esta solução.

Como exemplo, mostramos os resultados do problema com $\ell = 8$ para o qual escolhemos a normalização de \underline{A} em (2.9) de forma que $T_1 = \lambda_1/\ell = 1$. Os demais autovalores ficam então:

$$T_2 = \frac{J\lambda_2}{\ell} = \frac{\epsilon^3 + 2\epsilon^2 + 4\epsilon}{\epsilon^3 + 2\epsilon^2 + 4\epsilon + 8}, \quad (2.64)$$

$$T_3 = \frac{J\lambda_3}{\ell} = \frac{\epsilon^3 + 2\epsilon^2}{\epsilon^3 + 2\epsilon^2 + 4\epsilon + 8}, \quad (2.65)$$

$$T_4 = \frac{J\lambda_4}{\ell} = \frac{\epsilon^3}{\epsilon^3 + 2\epsilon^2 + 4\epsilon + 8}. \quad (2.66)$$

Podemos comparar na figura 2.6 as temperaturas críticas em função de ϵ para três tipos de soluções:

a) $\bar{m} = m_3 \bar{v}^3$ onde \bar{v}^3 é qualquer um dos autovetores com T_3 .

b) $\bar{m} = m_4 \bar{v}^4$ onde \bar{v}^4 é qualquer um dos autovetores com T_4 .

$$c) m_c^{\gamma\gamma'} = \begin{cases} m_\gamma v_c^\gamma(\frac{\ell}{2}), & 1 \leq c < \frac{\ell}{2} \\ m_{\gamma'} v_c^{\gamma'}(\frac{\ell}{2}), & \frac{\ell}{2} \leq c \leq \ell \end{cases}$$

As equações para T^* nos três casos são:

$$\begin{cases} m_3 = \tanh \left[\frac{T_3}{T_3^*} m_3 \right], \\ 1 - \frac{1}{T_3^*} (1 - m_3^2) = 0 \end{cases}, \quad (2.67)$$

$$\begin{cases} m_4 = \tanh \left[\frac{T_4}{T_4^*} m_4 \right], \\ 1 - \frac{1}{T_4^*} (1 - m_4^2) = 0 \end{cases}, \quad (2.68)$$

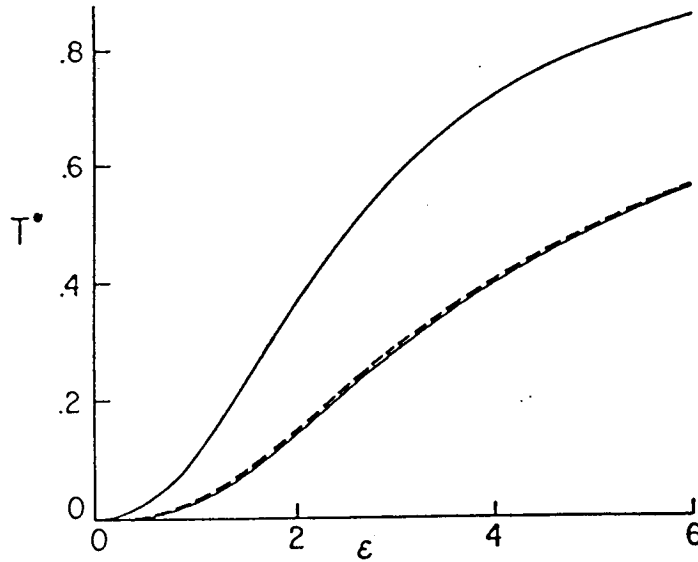


Figura 2.6: Resultados numéricos para a temperatura crítica, T_{34}^* (linha tracejada), de uma solução misturada de autovalores em função do parâmetro ε . As curvas cheias representam as temperaturas críticas T_3^* (superior) e T_4^* (inferior)

$$\begin{cases} m_3 = \tanh \left[\frac{T_3}{T_{34}^*} m_4 \right], \\ m_4 = \tanh \left[\frac{T_4}{T_{34}^*} m_4 \right], \\ \Lambda_1^+ = 0 \end{cases} \quad (2.69)$$

com

$$\Lambda_1^+ = 1 - \frac{T_1 + T_2}{2T_{34}^*} \left(1 - \frac{m_3^2 + m_4^2}{2} \right) - \frac{1}{T_{34}^*} \sqrt{(T_1 + T_2)^2 \left(1 - \frac{m_3^2 + m_4^2}{2} \right) + T_1 T_2 (m_3^2 + m_4^2)}. \quad (2.70)$$

Quanto a energia livre das soluções do tipo 2.39 podemos por substituição direta desta solução em (2.33) obter:

$$f_{\gamma\gamma'}(T) = \frac{1}{2} (f_\gamma(T) + f_{\gamma'}(T)) . \quad (2.71)$$

2.5 Solução com mistura de padrões

Primeiramente vamos voltar a expressão (2.30)

$$m_a^\mu = \left\langle \xi^\mu \tanh \left(\frac{\beta}{\ell} \sum_b A_{ab} \xi \cdot \vec{m}_b \right) \right\rangle_\xi \quad (2.72)$$

que é o conjunto de equações que deve ser resolvido para $\mu = 1, \dots, p$ e $a = 1, \dots, \ell$.

Continuamos usando a notação “ \vec{v} ” para vetor no espaço de memória e “ \bar{v} ” para vetor no espaço de blocos.

Novamente como na seção 2.4 o problema é separável em dois casos.

2.5.1 Soluções onde \vec{m} é autovetor de \underline{A}

Este tipo de solução pode ser representada por um produto externo entre um vetor \vec{m}_γ no espaço de padrões e outro \bar{v}^γ no espaço de blocos, ou seja,

$$\vec{m} = \vec{m}_\gamma \otimes \bar{v}^\gamma \quad (2.73)$$

Como antes

$$\underline{A} \cdot \bar{v}^\gamma = \lambda_\gamma \bar{v}^\gamma$$

e $|\bar{v}_a^\gamma| = 1$.

Assim ao substituírmos em (2.72) obtemos

$$\vec{m} = \left\langle \vec{\xi} \tanh \left(\frac{\beta \lambda_\gamma}{\ell} \vec{\xi} \cdot \vec{m} \right) \right\rangle_\xi \quad (2.74)$$

Daí saem as soluções que misturam memórias, simétricas e assimétricas, completamente equivalente ao que ocorre no modelo de Hopfield [15].

A energia livre é

$$f = \frac{1}{2} \frac{\lambda_\gamma}{\ell} (\vec{m})^2 - \frac{1}{\beta} \left\langle \ln \left(2 \cosh \left(\frac{\beta \lambda_\gamma}{\ell} \vec{m} \cdot \vec{\xi} \right) \right) \right\rangle_\xi, \quad (2.75)$$

e a matriz de estabilidade

$$M_{ab}^{\mu\nu} = \left(\frac{A_{ab}}{\ell^2} - \frac{\beta}{\ell^3} (A_{ab})^2 \right) \delta^{\mu\nu} + \frac{\beta}{\ell^3} (A_{ab})^2 \left\langle \xi^\mu \xi^\nu \tanh^2 \left[\frac{\beta \lambda_\gamma}{\ell} \vec{\xi} \cdot \vec{m}_d \right] \right\rangle_\xi \quad (2.76)$$

Esta matriz é facilmente diagonalizável no espaço de blocos, já no espaço de memórias depende do tipo particular de solução misturada de memórias que tivermos.

2.5.2 Soluções com mistura de autovetores e memórias

Semelhante ao que ocorre na subseção 2.4.2 existem soluções na forma geral

$$\vec{m}(\ell) = \begin{pmatrix} \vec{m}'(\frac{\ell}{2}) \\ \vec{m}''(\frac{\ell}{2}) \end{pmatrix} \quad (2.77)$$

E a estabilidade dependerá da diagonalização de

$$M_{ab}^{\mu\nu} = \left(\frac{A_{ab}}{\ell^2} - \frac{\beta^2}{\ell^3} \sum_c A_{ac} A_{cb} \right) \delta^{\mu\nu} + \frac{\beta}{\ell^3} \sum_c A_{ac} A_{cb} \left\langle \xi^\mu \xi^\nu \tanh^2 \left[\frac{\beta}{\ell} \sum_d A_{cd} \vec{\xi} \cdot \vec{m}_d \right] \right\rangle_\xi, \quad (2.78)$$

que, em geral, não será diagonal no espaço de memória e nem será diagonalizável no espaço de blocos pela mesma matriz que diagonaliza \underline{A} .

Capítulo 3

Propriedades de saturação.

No capítulo anterior, discutimos os tipos de soluções estacionárias que emergem de um sistema onde se combinou a regra de aprendizado de Hebb com uma estrutura especial de agregados, ou blocos, de neurônios. Vimos que esta estrutura, por si própria, provoca o aparecimento de uma família de estados descendentes para cada protótipo armazenado pela regra de Hebb.

Sabemos do estudo do modelo de Hopfield [14] que a regra de Hebb, por sua vez, também proporciona o surgimento de outros estados estacionários além dos p protótipos explicitamente armazenados. A princípio, estes estados são indesejáveis, já que não conhecemos sua estrutura microscópica, e crescem em número à medida que aumentamos p . Existe um limite para este valor a partir do qual os estados protótipos deixam de ser estacionários, ou seja, o sistema não é mais capaz de reconhecê-los como memórias. Isso se deve à concorrência de dois termos na função energia do sistema. Considere que estejamos interessados na recuperação da memória $\mu = \mu_0$, a regra de Hebb, então, pode ser separada em dois termos,

$$J_{ij} = \frac{1}{N} \xi_i^{\mu_0} \xi_j^{\mu_0} + \frac{1}{N} \sum_{\mu \neq \mu_0} \xi_i^{\mu} \xi_j^{\mu} \quad . \quad (3.1)$$

O primeiro deles, o termo de sinal, tenta alinhar o sistema nesta memória e o outro, o

termo de ruído, devido as outras $p - 1$ memórias, tenta tirá-lo dela. Os estados estacionários do sistema são resultado da competição destes dois termos. O termo de ruído cresce com p até um ponto onde supera o termo de sinal, neste momento o sistema perde sua capacidade de recuperar memória. Quando isso ocorre se diz que a rede atingiu sua saturação. Esta transição é descontínua e ocorre para um valor finito da razão $\alpha = \frac{p}{N}$.

O nosso objetivo neste capítulo é investigar as propriedades de saturação de uma rede com hierarquia de interações entre seus neurônios, como foi descrita na seção 2.1. Mais especificamente, vamos estar interessados na determinação de sua fração crítica de armazenagem α_c .

Existem diferentes formas de compararmos os termos de sinal e ruído e, portanto, diversos critérios para determinar esta quantidade. Na primeira parte deste capítulo, usaremos o mesmo método de campo médio desenvolvido na capítulo anterior. Neste caso, o valor de α_c corresponde ao ponto a partir do qual desaparece a solução $m \neq 0$ das equações de estado do sistema. Para um valor de α imediatamente menor que este valor crítico, o sistema recupera os padrões protótipos com erro, ou seja, com uma fração finita de neurônios desalinhados. Alternativamente, podemos estar interessados em calcular o número máximo de padrões que o sistema pode suportar de forma que ainda possa recuperar perfeitamente, ou seja, com erro zero, os padrões armazenados. Agora, o método apropriado é o chamado “método de sinal×ruído” para o qual dedicaremos a última parte deste capítulo.

Demonstraremos ao final que, muito embora os valores de α_c para os dois métodos não sejam comparáveis, o mesmo não ocorre com as razões

$$\left(\frac{\alpha_c}{\alpha_c^{\text{Hopfield}}} \right)$$

3.1 Campo Médio - Cálculo da Função de Energia Livre

A saturação de uma rede de neurônios ocorre quando tentamos armazenar um número de padrões proporcional ao número de sítios da rede, ou seja, $p = \alpha N$. Observe que, neste caso, a condição para que os padrões $\{\xi_i^\mu\}$ sejam automediáveis, $2^p/N \rightarrow 0$, não é mais satisfeita e, portanto, a natureza e a dificuldade do problema mudam bastante.

Com a falta desta propriedade das médias, deveremos calcular a energia livre tendo que realizar explicitamente a média sobre as variáveis aleatórias, ou seja,

$$f = -\frac{1}{\beta N} \langle \ln Z \rangle_\xi \quad . \quad (3.2)$$

O procedimento usual nestes casos é usar o método das réplicas que substitui a média sobre o logaritmo, difícil de ser operada, por

$$\langle \ln x \rangle = \lim_{n \rightarrow 0} \frac{\langle x^n \rangle - 1}{n} \quad .$$

Assim, a nossa expressão para a energia livre se torna

$$f = -\frac{1}{\beta N} \lim_{n \rightarrow 0} \frac{\langle Z^n \rangle_\xi - 1}{n} \quad . \quad (3.3)$$

O termo “réplica” vem do fato de Z^n ser a função de partição de n sistemas idênticos, ou réplicas. Esta função de partição replicada é dada por

$$Z^n = Tr_{\{s_i^\rho\}} \exp \left(\frac{\beta}{2N} \sum_{a,b} A_{ab} \sum_{i,j}^{(a),(b)} \sum_{\mu}^{\alpha N} \sum_{\rho=1}^n \xi_i^\mu \xi_j^\mu S_i^\rho S_j^\rho \right) \quad , \quad (3.4)$$

onde foi usada a mesma expressão usada anteriormente para o hamiltoniano.

O estado da rede $\{S_i\}$ pode estar completamente alinhado com um dos padrões $\{\xi_i^\mu\}$ ou parcialmente alinhado com muitos deles. Quando este alinhamento é tal que a superposição definida em (??) é $o(1)$ essa memória é dita condensar macroscopicamente, se ao contrário ela for $o(1/\sqrt{N})$ ou menor, dizemos que ela não condensa ou está microscopicamente condensada. O número máximo de memórias p' que podem condensar macroscopicamente em um único estado misturado deve ser $\ll \sqrt{N}$ [42].

Assim, vamos considerar que no nosso conjunto de p padrões apenas um número finito p' condense macroscopicamente, de modo que podemos realizar a média $\langle Z^n \rangle_\xi$ na forma de um produto de duas médias:

$$\begin{aligned} \langle Z^n \rangle &= Tr_S \left\langle \exp \left(\frac{\beta}{2N} \sum_{\rho=1}^n \sum_{a,b} A_{ab} \sum_{i,j}^{(a),(b)} \sum_{\mu=1}^{p'} \xi_i^\mu \xi_j^\mu S_i^\rho S_j^\rho \right) \right\rangle_\xi \\ &\times \left\langle \exp \left(\frac{\beta}{2N} \sum_{\rho=1}^n \sum_{a,b} A_{ab} \sum_{i,j}^{(a),(b)} \sum_{\nu=p'+1}^{\alpha N} \xi_i^\nu \xi_j^\nu S_i^\rho S_j^\rho \right) \right\rangle_\xi \end{aligned} \quad (3.5)$$

Como a primeira delas é sobre um número finito de $\{\xi^\mu\}$, estas variáveis são automediáveis e o sinal da média é dispensável. Devemos somente nos preocupar com a parte correspondente às memórias não recuperáveis:

$$\begin{aligned} &\left\langle \exp \left(\frac{\beta}{2N} \sum_{\rho=1}^n \sum_{a,b} A_{ab} \sum_{i,j}^{(a),(b)} \sum_{\nu=p'+1}^{\alpha N} \xi_i^\nu \xi_j^\nu S_i^\rho S_j^\rho \right) \right\rangle_\xi \\ &= \exp \left(\alpha N \ln \left\langle \exp \left(\frac{\beta}{2} \sum_{i,j} \Lambda_{ij} \xi_i \xi_j \right) \right\rangle_\xi \right) \end{aligned} \quad (3.6)$$

Aqui foi utilizado o fato das memórias serem estatisticamente independentes e, portanto, a média do produto ser equivalente ao produto das médias; fizemos $\alpha N - p' \sim \alpha N$ e definimos

$$\Lambda_{ij} = \frac{A_{ab}}{N} \sum_{\rho=1}^n S_i^\rho S_j^\rho, \quad (3.7)$$

com $i \in a$ e $j \in b$.

3.1.1 Cálculo de $\langle \exp(\frac{\beta}{2} \sum_{i,j} \Lambda_{ij} \xi_i \xi_j) \rangle_\xi$

O fator (3.7) é responsável pelo ruído que os padrões que não condensam causam na recuperação daqueles que podem ser recuperáveis. A avaliação correta deste termo, portanto, é muito importante.

No trabalho de Amit e all. [14] onde o método para o estudo de uma rede de neurônios foi originalmente desenvolvido, isto é feito através da introdução direta dos parâmetros de ordem m^ν , como em (2.26). E, como no caso de padrões microscópicos $m^\nu \sim o(1/\sqrt{N})$, podemos expandir (2.27) em primeira ordem deste parâmetro e a seguir integrar sobre ele. No nosso caso, no entanto, este último passo nos levaria a uma integração gaussiana generalizada muito complexa devido a presença da matriz $\underline{\Lambda}$. Assim optamos por um procedimento alternativo, mas muito poderoso, que é a expansão diagramática de (3.6).

Isto é feito do seguinte modo. Considere, primeiramente, a expansão

$$\begin{aligned}
\ln \langle \exp(\frac{\beta}{2} \sum_{i,j} \Lambda_{ij} \xi_i \xi_j) \rangle &= \frac{\beta}{2} \sum_{i,j} \Lambda_{ij} \langle \xi_i \xi_j \rangle \\
&+ \frac{1}{2!} \left(\frac{\beta}{2}\right)^2 \sum_{ijkl} \Lambda_{ij} \Lambda_{kl} (\langle \xi_i \xi_j \xi_k \xi_l \rangle - \langle \xi_i \xi_j \rangle \langle \xi_k \xi_l \rangle) \\
&+ \frac{1}{3!} \left(\frac{\beta}{2}\right)^3 \sum_{ijklmn} \Lambda_{ij} \Lambda_{kl} \Lambda_{mn} (\langle \xi_i \xi_j \xi_k \xi_l \xi_m \xi_n \rangle - 3 \langle \xi_i \xi_j \xi_k \xi_l \rangle \langle \xi_m \xi_n \rangle \\
&+ 2 \langle \xi_i \xi_j \rangle \langle \xi_k \xi_l \rangle \langle \xi_m \xi_n \rangle) + \dots
\end{aligned} \tag{3.8}$$

ou simplesmente

$$\ln \langle \exp(\frac{\beta}{2} \sum_{i,j} \Lambda_{ij} \xi_i \xi_j) \rangle = \frac{\beta}{2} \sum_{i,j} \Lambda_{ij} \langle \xi_i \xi_j \rangle_c$$

$$+ \frac{1}{2!} \left(\frac{\beta}{2}\right)^2 \sum_{ijkl} \Lambda_{ij} \Lambda_{kl} \langle \xi_i \xi_j \xi_k \xi_l \rangle_c + \dots \quad (3.9)$$

onde $\langle \dots \rangle_c$ significa a média cumulante. Observe que para simplificar a notação omitimos o subíndice ξ das média configuracionais.

Cada um dos termos desta expansão pode ser posto na forma de diagramas. A representação diagramática, embora não seja essencial, nos permite visualizar melhor os termos que contribuem dominantemente na soma.

Como exemplo vamos calcular explicitamente as contribuições dos dois primeiros cumulantes em (3.9). Começemos pelas médias

$$\langle \xi_i \xi_j \rangle = \delta_{ij} \quad (3.10)$$

e

$$\langle \xi_i \xi_j \xi_k \xi_l \rangle = \delta_{ij} \delta_{kl} + \delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk} - 2\delta_{ijkl} \quad (3.11)$$

que, introduzidas em (3.8) resultariam nos seguintes termos

$$\frac{\beta}{2} \sum_{ij} \Lambda_{ij} \langle \xi_i \xi_j \rangle = \frac{\beta}{2} \sum_i \Lambda_{ii} \quad (3.12)$$

e

$$\begin{aligned} \frac{1}{2!} \left(\frac{\beta}{2}\right)^2 \sum_{ijkl} \Lambda_{ij} \Lambda_{kl} \langle \xi_i \xi_j \xi_k \xi_l \rangle &= \frac{1}{2} \left(\frac{\beta}{2}\right)^2 \left(\sum_i \Lambda_{ii}\right)^2 \\ + \left(\frac{\beta}{2}\right)^2 \sum_{ij} \Lambda_{ij} \Lambda_{ji} - \left(\frac{\beta}{2}\right)^2 \sum_i (\Lambda_{ii})^2 & \quad (3.13) \end{aligned}$$

As regras para construção dos diagramas que usaremos são as seguintes:

- (i) cada fator $\frac{\theta}{2} \Lambda_{ij}$ é indicado por uma linha não orientada que liga os pontos "i" e "j".
- (ii) cada soma é indicada por um vértice.

Assim temos para (3.12) e (3.13)

$$\text{Diagrama} \tag{3.14}$$

e

$$1/2 (\text{Diagrama 1} \text{ } \text{Diagrama 2}) + \text{Diagrama 3} - \text{Diagrama 4} \tag{3.15}$$

A separação dos termos dominantes na soma em (3.9) deve obedecer a dois fatores importantes:

- Diagramas com vértices de quatro pernas ou mais, são de ordem $o(1/N)$ ou menor. De fato, se lembrarmos de (3.7) que cada Λ_{ij} é $o(1/N)$ veremos que o único modo de termos $o(1)$ é se existir o mesmo número de Λ e de somas, ou seja, número de linhas igual ao número de vértices. Assim diagramas do tipo do último à direita de (3.15) serão desconsiderados. Isto equivale a tomarmos:

$$\langle \xi_1 \xi_2 \dots \xi_n \rangle \cong \sum_P \langle \xi_{k_1} \xi_{k_2} \rangle \langle \xi_{k_3} \xi_{k_4} \rangle \dots \langle \xi_{k_{n-1}} \xi_{k_n} \rangle \tag{3.16}$$

onde P significa todas as possíveis combinações dos índices $1, 2, \dots, n$, dois a dois.

- Termos desconexos desaparecem nos cumulantes [43]. Esta regra geral é facilmente verificável na álgebra de diagramas, veja por exemplo:

$$\langle \xi_i \xi_j \xi_k \xi_l \rangle_c = \langle \xi_i \xi_j \xi_k \xi_l \rangle - \langle \xi_i \xi_j \rangle \langle \xi_k \xi_l \rangle \quad (3.17)$$

$$\left(\begin{array}{c} \text{Diagram 1} \\ \text{Diagram 2} \end{array} + 2 \begin{array}{c} \text{Diagram 3} \\ \text{Diagram 4} \end{array} - 2 \begin{array}{c} \text{Diagram 5} \\ \text{Diagram 6} \end{array} \right) - \left(\begin{array}{c} \text{Diagram 7} \\ \text{Diagram 8} \end{array} \right) = 2 \begin{array}{c} \text{Diagram 9} \end{array}$$

e

$$\begin{aligned} \langle \xi_i \xi_j \xi_k \xi_l \xi_m \xi_n \rangle_c &= \langle \xi_i \xi_j \xi_k \xi_l \xi_m \xi_n \rangle \\ &- 3 \langle \xi_i \xi_j \xi_k \xi_l \rangle \langle \xi_m \xi_n \rangle + 2 \langle \xi_i \xi_j \rangle \langle \xi_k \xi_l \rangle \langle \xi_m \xi_n \rangle \end{aligned} \quad (3.18)$$



Cada cumulante contribuirá apenas com um diagrama. Este diagrama é conexo e tem somente vértices de duas pernas.

O coeficiente C_r , à frente de cada diagrama, representa o número de escolhas dos índices em (3.16) tais que não tenhamos diagramas desconexos. Este cálculo é fácil de ser feito para um diagrama de ordem r .

Considere um segmento orientado \rightarrow como sendo um Λ_{ij} . Cada um destes segmentos é numerado e o diagrama que estamos interessados é a polinomial fechada formada por r destes segmentos.



O que devemos contar é quantas combinações possíveis se pode obter. Primeiramente temos um fator $(r - 1)!$ para as permutações cíclicas, além disso, cada segmento pode ter duas orientações e portanto deve-se ajuntar um termo 2^r . Finalmente, devido a simetria de reflexão dos diagramas devemos dividir o número que obtivemos por dois. Logo

$$C_r = (r - 1)! 2^{r-1} \quad (3.19)$$

Voltando a (3.9), obtemos

$$\ln \langle \exp \left(\frac{\beta}{2} \sum_{i,j} \Lambda_{ij} \xi_i \xi_j \right) \rangle = \frac{1}{2} \left(\text{Diagrama 1} + \text{Diagrama 2} + \text{Diagrama 3} + \dots \right) , \quad (3.20)$$

e, reutilizando as regras de construção dos diagramas, podemos escrever as somas:

$$\begin{aligned} \ln \langle \exp \left(\frac{\beta}{2} \sum_{i,j} \Lambda_{ij} \xi_i \xi_j \right) \rangle &= \frac{1}{2} \left\{ \beta \sum_i \Lambda_{ii} + \frac{\beta^2}{2} \sum_{ij} \Lambda_{ij} \Lambda_{ji} + \frac{\beta^3}{3} \sum_{ijk} \Lambda_{ij} \Lambda_{kl} \Lambda_{mn} + \dots \right\} \\ &= -\frac{1}{2} \text{Tr}_{\{i\}} \ln (1 - \beta \underline{\Lambda}) \quad , \end{aligned} \quad (3.21)$$

onde $\underline{\Lambda}$ é a matriz de elementos Λ_{ij} e o traço é tomado no espaço de sítios.

Para comparar com resultados conhecidos, podemos testar este método para o Modelo de Hopfield. Neste caso, teríamos

$$\Lambda_{ij}^H = \frac{1}{N} \sum_{\rho} S_i^{\rho} S_j^{\rho} \quad (3.22)$$

e, portanto,

$$\text{Tr}_{\{i\}} (\underline{\Lambda}^H)^r = \text{Tr}_{\{\rho\}} (\underline{Q})^r \quad , \quad (3.23)$$

onde \underline{Q} é uma matriz no espaço de réplicas com elementos

$$Q^{\rho\rho'} = \frac{1}{N} \sum_i S_i^{\rho} S_i^{\rho'}$$

O que significa que, para o Modelo de Hopfield, a equação (3.21) pode ser escrita

$$-\frac{1}{2}Tr_{\{i\}} \ln(1 - \beta \underline{\Lambda}) = -\frac{1}{2}Tr_{\{\rho\}} \ln(1 - \beta \underline{Q}) \quad , \quad (3.24)$$

com o segundo traço sendo realizado no espaço de réplicas, o que concorda com os resultados da referência [14].

Para Λ_{ij} definida como em (3.7) não podemos fazer algo tão simples como foi feito acima no modelo de Hopfield. O único modo de representarmos (3.21) como produto de matrizes no espaço de réplicas é através da soma

$$-\frac{1}{2}Tr_{\{i\}} \ln(1 - \beta \underline{\Lambda}) = \sum_{k=1}^{\infty} \left\{ \frac{\beta^k}{k} \sum_{a_1, \dots, a_k} A_{a_1 a_2} \dots A_{a_k a_1} \sum_{\rho_1, \dots, \rho_k} Q_{a_1}^{\rho_1 \rho_2} \dots Q_{a_k}^{\rho_k \rho_1} \right\} \quad , \quad (3.25)$$

com

$$Q_a^{\rho\rho'} = \frac{1}{N} \sum_i^{(a)} S_i^\rho S_i^{\rho'} \quad . \quad (3.26)$$

Assim, temos que :

$$\begin{aligned} & \left\langle \exp\left(\frac{\beta}{2} \sum_{i,j} \Lambda_{ij} \xi_i \xi_j\right) \right\rangle^{\alpha N} = \\ & \times \exp \left\{ \frac{\alpha N}{2} \sum_{k=1}^{\infty} \left(\frac{\beta^k}{k} \sum_{a_1, \dots, a_k} A_{a_1 a_2} \dots A_{a_k a_1} \sum_{\rho_1, \dots, \rho_k} Q_{a_1}^{\rho_1 \rho_2} \dots Q_{a_k}^{\rho_k \rho_1} \right) \right\} \quad (3.27) \end{aligned}$$

3.1.2 Cálculo do traço sobre as variáveis de spin em $\langle Z^n \rangle$.

De (3.5) e (3.27), podemos escrever a média configuracional sobre a função de partição das réplicas como

$$\begin{aligned} \langle Z^n \rangle &= \text{Tr}_S \left\langle \exp \left(\frac{\beta}{2N} \sum_{\rho=1}^n \sum_{a,b} A_{ab} \sum_{i,j}^{(a),(b)} \sum_{\mu=1}^{p'} \xi_i^\mu \xi_j^\mu S_i^\rho S_j^\rho \right) \right\rangle_\xi \\ &\times \exp \left\{ \frac{\alpha N}{2} \sum_{k=1}^{\infty} \left(\frac{\beta^k}{k} \sum_{a_1, \dots, a_k} A_{a_1 a_2} \dots A_{a_k a_1} \sum_{\rho_1, \dots, \rho_k} Q_{a_1}^{\rho_1 \rho_2} \dots Q_{a_k}^{\rho_k \rho_1} \right) \right\} \quad (3.28) \end{aligned}$$

E, usando novamente transformação de Hubbard-Stratonovich [39], podemos escrever o primeiro fator como

$$\begin{aligned} &\exp \left\{ \frac{\beta N}{2\ell} \sum_{a,b} \frac{A_{ab}}{\ell} \left(\frac{\ell}{N} \sum_i^{(a)} \xi_i^\mu S_i^\rho \right) \left(\frac{\ell}{N} \sum_j^{(b)} \xi_j^\mu S_j^\rho \right) \right\} = \\ &\sqrt{|\underline{A}|} (\sqrt{\pi} \beta N / \ell)^\ell \int \prod_a dm_a^{\mu, \rho} \times \\ &\exp \left\{ -\frac{\beta N}{\ell} \left(\frac{1}{2} \sum_{a,b} \frac{A_{ab}}{\ell} m_a^{\mu, \rho} m_b^{\mu, \rho} + \sum_{a,b} \frac{A_{ab}}{\ell} \left(\frac{\ell}{N} \sum_i^{(a)} \xi_i^\mu S_i^\rho \right) m_b^{\mu, \rho} \right) \right\} \quad (3.29) \end{aligned}$$

Para extrair as variáveis de spin de dentro de (3.27), introduz-se as variáveis contínuas $q_a^{\rho\rho'}$ e $r_a^{\rho\rho'}$ através das identidades

$$\frac{\alpha \beta^2 N_0}{2} \int_{-i\infty}^{i\infty} \frac{dr_a^{\rho\rho'}}{2\pi i} \int_{-\infty}^{\infty} dq_a^{\rho\rho'} \exp \left\{ \frac{\alpha \beta^2 N}{2} r_a^{\rho\rho'} \left(\frac{q_a^{\rho\rho'}}{\ell} - Q_a^{\rho\rho'} \right) \right\} = 1 \quad , \quad (3.30)$$

para cada $a = 1, \dots, \ell$ e $\rho, \rho' = 1, \dots, n$.

Assim, obtemos o seguinte resultado

$$\langle Z^n \rangle = \Gamma_n \int_{-\infty}^{\infty} \prod_a \left\{ \prod_{\mu\rho} dm_a^{\mu\rho} \prod_{\rho<\rho'} dr_a^{\rho\rho'} dq_a^{\rho\rho'} \right\} e^{-\beta N n \mathcal{F}_n(m, q, r)} , \quad (3.31)$$

onde nós agregamos todas constantes multiplicativas em Γ_n e

$$\begin{aligned} \diamond[B\mathcal{F}_n(m, q, r)] &= \frac{1}{2n\ell} \sum_{\mu\rho} \sum_{a,b} \frac{A_{ab}}{\ell} m_a^{\mu\rho} m_b^{\mu\rho} + \frac{1}{2n\ell} \alpha\beta \sum_{\rho\neq\rho'} \sum_a r_a^{\rho\rho'} q_a^{\rho\rho'} \\ &\quad - \frac{\alpha}{2n\beta} \sum_{k=2}^{\infty} \frac{1}{k} \left(\frac{\beta}{\ell}\right)^k \sum_{a_1 \dots a_k} A_{a_1 a_2} \dots A_{a_k a_1} \text{Tr}_{\{\rho\}} \left[\underline{q}_{a_1} \underline{q}_{a_2} \dots \underline{q}_{a_k} \right] \\ &\quad - \frac{1}{\beta N n} \ln \left\{ \text{Tr}_{\{S\}} e^{-\beta H_{eff}} \right\} , \end{aligned} \quad (3.32)$$

com

$$H_{eff} = \frac{N}{\ell} \sum_{\mu\rho} \sum_{a,b} \frac{A_{ab}}{\ell} m_a^{\mu\rho} S_b^{\mu\rho} + \frac{\alpha\beta N}{2\ell} \sum_a \sum_{\rho\neq\rho'} r_a^{\rho\rho'} Q_a^{\rho\rho'} . \quad (3.33)$$

Tomando a solução no ponto de sela para (3.31), a energia livre fica, de (3.3),

$$f = -\frac{1}{\beta N} \lim_{n \rightarrow 0} \frac{\Gamma^n e^{-\beta N n \mathcal{F}_n(\hat{m}, \hat{q}, \hat{r})} - 1}{n} , \quad (3.34)$$

ou

$$f = \lim_{n \rightarrow 0} (\mathcal{F}_n(\hat{m}, \hat{q}, \hat{r}) + o(n)) + o(1/N) . \quad (3.35)$$

O chapéu em m, q e r indica que estes são valores no ponto de sela. Por simplicidade, omitiremos este sinal no decorrer do texto. É importante lembrar que o limite termodinâmico está sendo subentendido e, este limite, comuta com o limite do número de réplicas indo a zero.

3.2 Teoria simétrica de réplicas

A teoria simétrica de réplicas consiste em supor que as soluções de ponto de sela da equação (3.31) sejam tais que não dependam das réplicas. Os parâmetros m , q e r devem ser tais que satisfaçam

$$\begin{aligned} m_a^{\mu\rho} &= m_a^\mu \\ q_a^{\rho\rho'} &= q_a & \rho \neq \rho' \\ r_a^{\rho\rho'} &= r_a & \rho \neq \rho', \end{aligned} \quad (3.36)$$

e, substituídos em (3.32), resultem

$$\begin{aligned} \mathcal{F}_n(m, q, r) &= \frac{1}{2\ell} \sum_{a,b} \frac{A_{ab}}{\ell} \vec{m}_a \cdot \vec{m}_b + \frac{n-1}{2\ell} \alpha\beta \sum_a r_a q_a \\ &\quad - \frac{\alpha}{2\beta n} \sum_{k=2}^{\infty} \frac{1}{k} \left(\frac{\beta}{\ell}\right)^k \sum_{a_1 \dots a_k} A_{a_1 a_2} \dots A_{a_k a_1} \\ &\quad \times \sum_{\rho_1 \dots \rho_k} (\delta^{\rho_1 \rho_2} + (1 - \delta^{\rho_1 \rho_2}) q_{a_1}) \dots (\delta^{\rho_k \rho_1} + (1 - \delta^{\rho_k \rho_1}) q_{a_k}) \\ &\quad - \frac{1}{\beta N n} \ln \left\{ \text{Tr}_{\{\sigma\}} e^{-\beta H_{eff}} \right\} . \end{aligned} \quad (3.37)$$

Rearranjando alguns termos e usando

$$\exp \left\{ \frac{\alpha\beta^2}{2} r_a \left(\sum_{\rho} S^{\rho} \right)^2 \right\} = \int Dz \exp \left(\beta \sqrt{\alpha r_a} \sum_{\rho} S^{\rho} z \right) \quad (3.38)$$

com $\int Dz \dots \equiv \int \frac{dz}{\sqrt{2\pi}} \exp(-\frac{1}{2}z^2) \dots$, obtemos, após tomarmos o traço sobre spins e o limite parcial do número de réplicas indo a zero, a seguinte expressão para a energia livre:

$$\begin{aligned}
f(m, q, r) &= \frac{1}{2\ell} \sum_{a,b} \frac{A_{ab}}{\ell} \vec{m}_a \vec{m}_b + \frac{\alpha\beta}{2\ell} \sum_a r_a (1 - q_a) \\
&\quad - \frac{\alpha}{2\beta} \lim_{n \rightarrow 0} \frac{1}{n} \sum_{k=2}^{\infty} \frac{1}{k} \left(\frac{\beta}{\ell} \right)^k \sum_{a_1 \dots a_k} A_{a_1 a_2} \dots A_{a_k a_1} \\
&\quad \times \sum_{\rho_1 \dots \rho_k} (\delta^{\rho_1 \rho_2} + (1 - \delta^{\rho_1 \rho_2}) q_{a_1}) \dots (\delta^{\rho_k \rho_1} + (1 - \delta^{\rho_k \rho_1}) q_{a_k}) \\
&\quad - \frac{1}{\beta\ell} \int Dz \left\langle \ln \left(2 \cosh \left(\beta \sum_b \frac{A_{ab}}{\ell} \vec{m}_b \cdot \vec{\xi} + \beta \sqrt{\alpha r_a} z \right) \right) \right\rangle. \quad (3.39)
\end{aligned}$$

O último limite, que está apenas indicado, pode ser feito da seguinte forma: Considere a soma em ρ_1, \dots, ρ_k escrita como

$$\sum_{\rho_1 \dots \rho_k} ((1 - q_{a_1}) \delta^{\rho_1 \rho_2} + q_{a_1}) \dots ((1 - q_{a_k}) \delta^{\rho_k \rho_1} + q_{a_k}) \quad .$$

Se abrirmos o produto e fizermos as somas, veremos que existirão termos de ordens n , n^2 , \dots , n^k , mas somente os de ordem n permanecerão após o limite. Estes termos são aqueles resultantes do produto acima que contenham k ou $k+1$ deltas de Kronecker (pois nestes casos restará apenas uma soma), ou seja,

$$n \{ (1 - q_{a_1}) \dots (1 - q_{a_r}) + (1 - q_{a_1}) \dots (1 - q_{a_{r-1}}) q_r + \text{demais combinações} \} \quad .$$

Substituindo em (3.39), obtemos

$$\begin{aligned}
f(m, q, r) &= \frac{1}{2\ell^2} \sum_{a,b} A_{ab} \vec{m}_a \vec{m}_b + \frac{\alpha\beta}{2\ell} \sum_a r_a (1 - q_a) \\
&\quad - \frac{\alpha}{2\beta} \sum_{k=2}^{\infty} \frac{1}{k} \left(\frac{\beta}{\ell} \right)^k \sum_{a_1 \dots a_k} A_{a_1 a_2} \dots A_{a_k a_1} \\
&\quad \times \left\{ \prod_{j=1}^k (1 - q_{a_j}) - \sum_{k'} q_{k'} \prod_{j \neq k'} (1 - q_{a_j}) \right\}
\end{aligned}$$

$$-\frac{1}{\beta\ell} \int Dz \left\langle \ln \left(2 \cosh \left(\beta \sum_b \frac{A_{ab}}{\ell} \bar{m}_b \cdot \vec{\xi} + \beta \sqrt{\alpha r_a} z \right) \right) \right\rangle \quad (3.40)$$

onde $\prod_{j=1}^0 \equiv 1$.

As equações do ponto de sela são :

$$m_a^\mu = \int Dz \left\langle \xi^\mu \tanh \left(\beta \sum_b \frac{A_{ab}}{\ell} \bar{m}_b \cdot \vec{\xi} + \beta \sqrt{\alpha r_a} z \right) \right\rangle, \quad (3.41)$$

$$q_a = \int Dz \left\langle \tanh^2 \left(\beta \sum_b \frac{A_{ab}}{\ell} \bar{m}_b \cdot \vec{\xi} + \beta \sqrt{\alpha r_a} z \right) \right\rangle, \quad (3.42)$$

e

$$r_c = -\frac{\alpha}{\beta^2} \sum_{k=2}^{\infty} \left(\frac{\beta}{\ell} \right)^k \sum_{a_2 \dots a_k} A_{ca_2} \dots A_{a_k c} \sum_{k'=2}^k \frac{q_{a_{k'}} \prod_{j=2}^k (1 - q_{a_j})}{1 - q_{a_{k'}}}. \quad (3.43)$$

3.2.1 Soluções homogêneas

Vimos que os estados estacionários do sistema com α finito, no caso de simetria de réplicas, são dados pela solução conjunta das equações (3.41), (3.42) e (3.43). Serão chamadas de soluções homogêneas aquelas que tiverem $q_a = q$ para todo "a".

Primeiramente, vamos verificar se realmente tais soluções existem. Começaremos por (3.43), se $q_a = q$, então:

$$\begin{aligned} r_c &= \frac{\ell}{\beta^2} \sum_{k=2}^{\infty} \left(\frac{\beta}{\ell} \right)^k [\underline{A}^k]_{cc} (k-1)q(1-q)^{k-2} \\ &= \frac{1}{\ell} \left[\frac{q \underline{A}^2}{(1 - \beta/\ell \underline{A}(1-q))^2} \right]_{cc}. \end{aligned} \quad (3.44)$$

Como a matriz \underline{A} é tal que, seus elementos diagonais são todos iguais, podemos escrever

$$r = \sum_{\delta=1}^{\ell} \frac{q(\lambda_{\delta}/\ell)^2}{[1 - \beta\lambda_{\delta}(1-q)]^2} \quad r_c = r \quad \forall c . \quad (3.45)$$

Vemos que a homogeneidade em q implica na homogeneidade em r . As demais equações de ponto de sela ficam

$$m_{\gamma}^{\mu} = \left\langle \xi^{\mu} \int Dz \tanh \left(\beta \sum_b \frac{A_{ab}}{\ell} \bar{m}_b \bar{\xi} + \beta \sqrt{\alpha r} z \right) \right\rangle_{\xi} , \quad (3.46)$$

$$q = \left\langle \int Dz \tanh^2 \left(\beta \sum_b \frac{A_{ab}}{\ell} \bar{m}_b \bar{\xi} + \beta \sqrt{\alpha r} z \right) \right\rangle_{\xi} . \quad (3.47)$$

Uma possível solução para este sistema de equações, é aquela para a qual $\{\bar{m}^{\mu}\}$ é um autovetor de \underline{A} , ou seja,

$$\bar{m} = \bar{m}_{\gamma} \otimes \bar{v}^{\gamma} .$$

onde \otimes representa o produto externo entre os espaços de memórias e de blocos.

Entre estas soluções, vamos estudar com mais detalhe apenas aquelas que envolve uma só memória, ou seja, que satisfaçam

$$m_{\gamma} = \int Dz \tanh \left(\beta \frac{\lambda_{\gamma}}{\ell} m_{\gamma} + \beta \sqrt{\alpha r} z \right) , \quad (3.48)$$

$$q = \int Dz \tanh^2 \left(\beta \frac{\lambda_{\gamma}}{\ell} m_{\gamma} + \beta \sqrt{\alpha r} z \right) , \quad (3.49)$$

$$r = \sum_{\delta=1}^{\ell} \frac{q(\lambda_{\delta}/\ell)^2}{[1 - \beta\lambda_{\delta}/\ell(1-q)]^2} , \quad (3.50)$$

com a função energia livre dada por

$$\begin{aligned}
 f(\gamma) = & \frac{1}{2} \frac{\lambda_\gamma}{\ell} m_\gamma^2 + \frac{\alpha\beta}{2} r (1 - q) \\
 & + \frac{\alpha}{2\beta} \sum_{\delta=1}^{\ell} \left\{ \ln \left[1 - \beta \frac{\lambda_\delta}{\ell} (1 - q) \right] - \beta \frac{\lambda_\delta}{\ell} q \left[1 - \beta \frac{\lambda_\delta}{\ell} (1 - q) \right]^{-1} \right\} \\
 & - \frac{1}{\beta} \int Dz \ln \left[2 \cosh \left(\beta \frac{\lambda_\gamma}{\ell} m_\gamma + z \beta \sqrt{\alpha r} \right) \right] . \quad (3.51)
 \end{aligned}$$

Em geral, nem todas as soluções do sistema de equações (3.41),(3.42) e (3.43) serão homogêneas. É imediato verificar que, um exemplo típico de solução não homogênea seria uma que envolvesse mistura de dois autovetores de \underline{A} como foi discutido na seção 2.4.2 do capítulo anterior.

3.3 Diagrama de fases $\alpha \times T$ para as soluções homogêneas

Como a equação para r é apenas uma equação algébrica podemos substituir sua expressão em (3.48) e obter o seguinte sistema para determinar m e q :

$$\begin{cases}
 m_\gamma = \int Dz \tanh \left(\beta \frac{\lambda_\gamma}{\ell} m_\gamma + \beta \sqrt{\alpha \sum_{\delta=1}^{\ell} \frac{q(\lambda_\delta/\ell)^2}{[1-\beta\lambda_\delta(1-q)]^2}} z \right) \\
 q = \int Dz \tanh^2 \left(\beta \frac{\lambda_\gamma}{\ell} m_\gamma + \beta \sqrt{\alpha \sum_{\delta=1}^{\ell} \frac{q(\lambda_\delta/\ell)^2}{[1-\beta\lambda_\delta(1-q)]^2}} z \right)
 \end{cases}$$

Uma vez que definimos λ_γ esse sistema apresenta três tipos de soluções estáveis, dependendo dos valores dos parâmetros α e T . Cada uma delas determina uma fase do sistema. Assim, se $m_\gamma(\alpha, T) \neq 0$ e $q_\gamma(\alpha, T) \neq 0$, dizemos que o sistema está na fase de recuperação da memória γ , se $m_1(\alpha, T) = 0$ e $q_1(\alpha, T) \neq 0$, onde $\gamma = 1$ é o ancestral, o sistema está na fase de vidro de spin e, finalmente, se $m_1(\alpha, T) = 0$ e $q_1(\alpha, T) = 0$, a fase é paramagnética.

Os diagramas de fase das figuras 3.1, 3.2 e 3.3 mostram os resultados numéricos para um caso especial em que a rede só possua uma hierarquia, ou dois blocos, e \underline{A} é dada pela matriz 2×2 :

$$\underline{A} = \frac{2}{2 + \epsilon} \begin{bmatrix} 1 + \epsilon & 1 \\ 1 & 1 + \epsilon \end{bmatrix} \quad (3.52)$$

Como sempre, o fator $2/(2 + 3)$ tem a função de normalizar esta matriz de forma que seu autovalor máximo seja igual a um. (A consequência disto é que as propriedades a $\alpha = 0$ da solução ancestral permanecem iguais as que ocorre no modelo de Hopfield).

Para dois blocos existe apenas uma solução descendente. Esta solução vamos denominar $(+-)$, pois é contruída a partir do ancestral pela inversão dos estados dos últimos $N/2$ sítios.

Nos diagramas as curvas T_M^1 e T_M^2 delimitam as regiões de recuperação para o ancestral e descendente $(+-)$; T_G mostra a transição vidro de spin - paramagnetismo e T_C^1 e T_C^2 são tais que abaixo delas $f_1 < f_{vs}$ e $f_2 < f_{vs}$ respectivamente, onde f_γ é a energia livre da solução γ e f_{vs} daquela de vidro de spin. As curvas pontilhadas seriam as verdadeiras curvas de transição de fase, mas como já foi alertado, em redes de neurônios, a região de recuperação é determinada a partir do aparecimento da solução estável de recuperação.

O procedimento numérico para a obtenção destes diagramas consistiu em resolver cada uma das equações de (3.52) pelo método de Newton, e o sistema, em si, por iteração linear. Ou seja, fixa-se o valor de q na primeira equação, calcula-se m por Newton, substitui-se o valor encontrado para m na segunda equação e calcula-se agora q por Newton, e assim sucessivamente. Algum cuidado deve ser dispensado no caso da segunda equação pois, para m fixo, esta possui mais de uma solução $q \neq 0$. A solução de interesse é separada das demais através do uso das condições de estabilidade deduzidas no apêndice 1.

Apesar dos diagramas apresentados, específicos para um sistema com dois blocos, darem uma boa noção de como ocorres o aparecimento dos estados de recuperação seria interessante que tivéssemos alguns resultados gerais, ou seja para o caso de um número

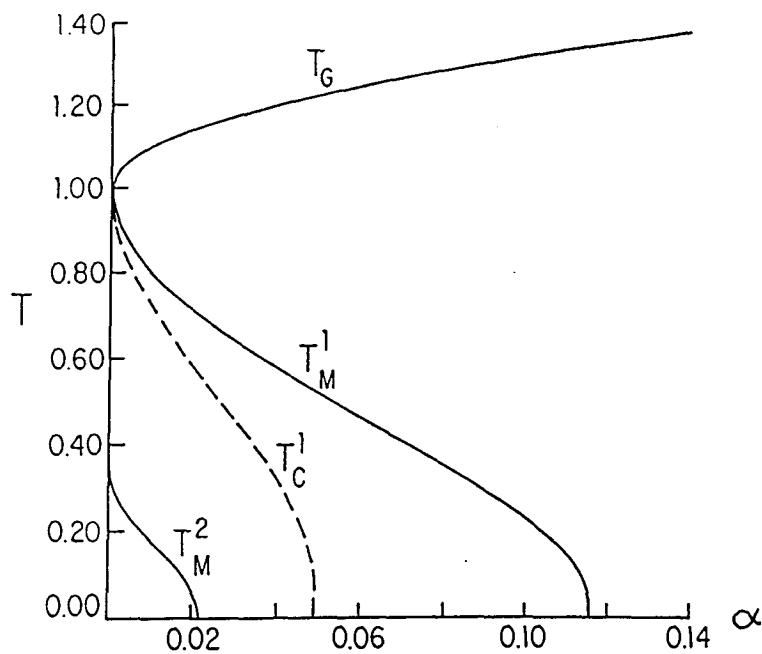


Figura 3.1: Diagrama de fases para $\varepsilon = 2$ ($\lambda_2/\ell = 0.5$).

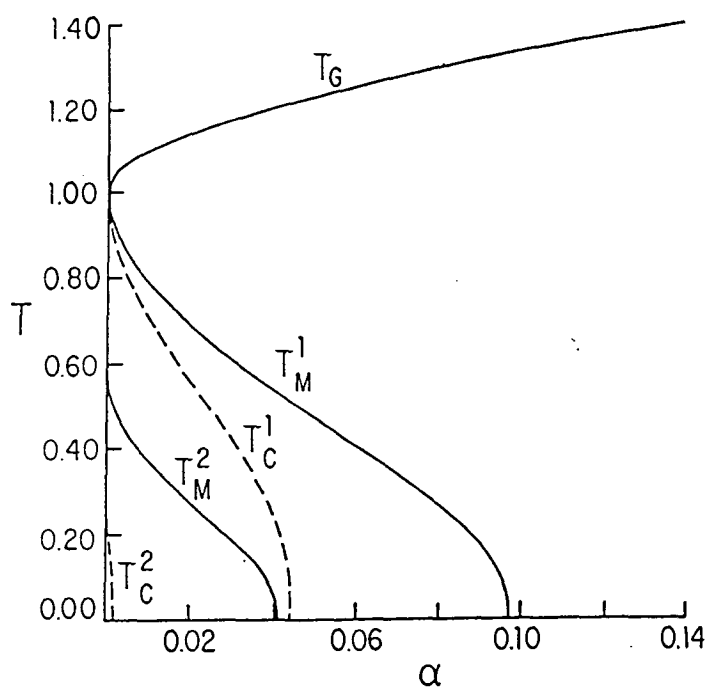


Figura 3.2: Diagrama de fases para $\varepsilon = 4.6$ ($\lambda_2/\ell = 0.7$).

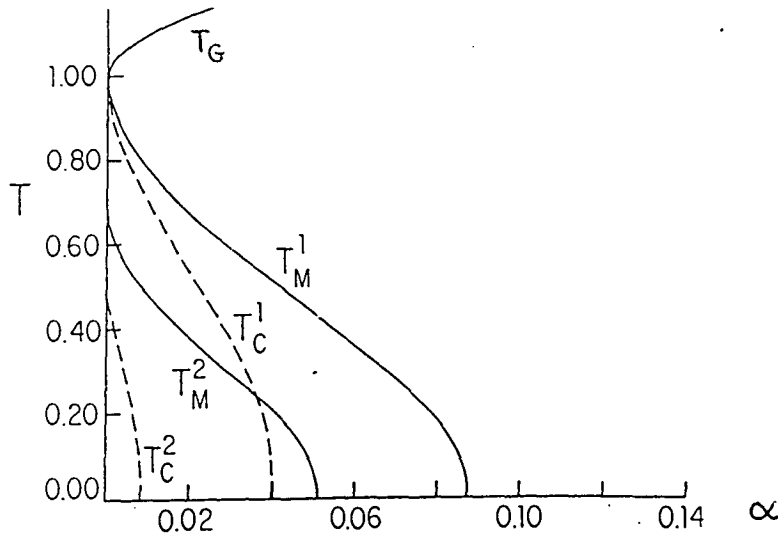


Figura 3.3: Diagrama de fases para $\varepsilon = 8$ ($\lambda_2/\ell = 0.8$).

qualquer de hierarquias. Nas seções seguintes discutimos alguns limites nos quais é possível obter detalhadamente estes resultados gerais.

3.3.1 T_G - Transição de fases Vidro de Spin \Leftrightarrow Paramagnetismo

Colocando $m = 0$ nas equações (3.52) acima, obtemos a equação para q na fase de vidro de spin,

$$q = \int Dz \tanh^2 \left(\beta \sqrt{\alpha \sum_{\delta=1}^{\ell} \frac{q(\lambda_{\delta}/\ell)^2}{[1 - \beta \lambda_{\delta}(1-q)]^2}} z \right) \quad (3.53)$$

Como a transição da fase de vidro de spin para a paramagnética é contínua, a curva $T_G = T_G(\alpha)$ que marca esta transição pode ser obtida pela expansão de (3.53) no limite em que $q \rightarrow 0$, o que resulta em

$$q = \int Dz \beta^2 z^2 \alpha q \sum_{\delta} \left(\frac{\lambda_{\delta}/\ell}{1 - \lambda_{\delta}/\ell/T(1-q)} \right)^2 + o(q^3) \quad (3.54)$$

Fazendo $q = 0$, obtemos a seguinte relação entre T_G e α :

$$1 = \alpha \sum_{\delta} \left(\frac{\lambda_{\delta}/\ell}{T_G - \lambda_{\delta}/\ell} \right)^2 \quad (3.55)$$

Esta relação, como está, não determina uma função $T_G = T(\alpha)$ unívoca, isto porque para um dado α existem várias possibilidades de T_G que a satisfazem. Se juntarmos, no entanto, a primeira das condições de estabilidade em (A.78), que pode ser escrita como

$$1 - \beta \frac{\lambda_1}{\ell} (1 - q) > 0 \quad , \quad (3.56)$$

para $q = 0$ obteremos a função, agora unívoca, definida como:

$$\begin{cases} 1 = \alpha \sum_{\delta} \left(\frac{\lambda_{\delta}/\ell}{T_G - \lambda_{\delta}/\ell} \right)^2 \\ T_G > \frac{\lambda_1}{\ell} \end{cases}$$

Esta função no limite que α cresce é

$$T_G(\alpha) \simeq \sqrt{\sum_{\delta} \left(\frac{\lambda_{\delta}}{\ell} \right)^2} \sqrt{\alpha} \quad (3.57)$$

Observe que este é um resultado válido para qualquer regra de aprendizagem do tipo (2.7).

3.3.2 Temperaturas críticas em $\alpha = 0$

Vamos considerar aqui os valores nos quais as curvas T_M cortam o eixo ordenado nos diagramas apresentados. Dois pontos importantes devem ser observados:

- Em $\alpha = 0$ uma dada solução determinada pelo autovalor λ_{γ} não dependerá dos outros autovalores de \underline{A} e nem do número de blocos existentes (vide seção 2.4).
- De todas as diversas temperaturas críticas, uma para cada escolha de λ_{γ} , somente a maior representará uma transição contínua.

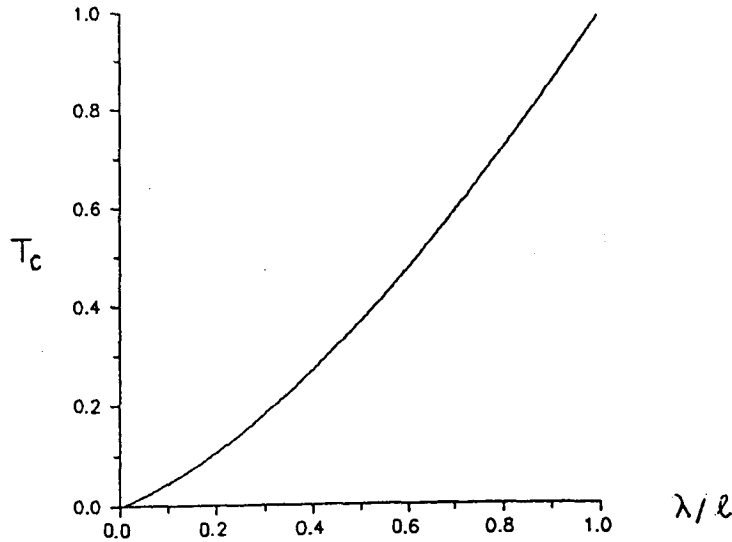


Figura 3.4: Temperatura crítica das soluções descendentes em função de λ_γ a $\alpha = 0$

Efetivamente, de (2.42), vemos que o sistema que determina T_M é

$$\begin{cases} m = \tanh\left(\frac{T_\gamma}{T}m\right) \\ 1 - \frac{T_\gamma}{T}(1 - m^2) > 0 \end{cases}, \quad (3.58)$$

onde voltamos a usar a notação $T_\gamma = \lambda_\gamma/\ell$.

Podemos observar que, se $\lambda_\gamma = \lambda_1$, as duas equações (note que agora a segunda é derivada em relação a “m” da primeira) terão soluções arbitrariamente próximas de $m = 0$. De uma forma mais clara, podemos ver que a condição para que haja solução $m \neq 0$ na primeira delas é $T < T_\gamma$, que é uma condição menos restritiva que

$$1 - \frac{T_\gamma}{T}(1 - m^2) > 0, \quad ,$$

se $\lambda_\gamma \neq \lambda_1$ e λ_1 é o autovalor máximo.

Tendo isto em vista, podemos calcular a temperatura crítica em função de T_γ resolvendo numericamente o sistema (3.58). Os resultados são apresentados na figura 3.4 para a qual escolhemos $T_1 = 1$.

3.3.3 Diagrama a $T = 0$ - Determinação do armazenamento crítico

O limite das equações (3.48) para $T \rightarrow 0$ resulta :

$$m_\gamma = \operatorname{erf} \left(\frac{1}{\sqrt{2\alpha r_\gamma}} \frac{\lambda_\gamma}{\ell} m_\gamma \right) , \quad (3.59)$$

$$r_\gamma = \sum_{\delta=1}^{\ell} \left[\frac{\lambda_\delta}{\ell} - C_\gamma \right]^{-2} , \quad (3.60)$$

$$C_\gamma = \lim_{\beta \rightarrow \infty} \beta(1 - q) = \left(\frac{2}{\pi \alpha r_\gamma} \right)^{1/2} \exp \left\{ -\frac{1}{2\alpha r_\gamma} \left(\frac{\lambda_\gamma}{\ell} m_\gamma \right)^2 \right\} , \quad (3.61)$$

Nos interessa soluções de α para $m \neq 0$, assim, podemos combinar as equações acima em uma única:

$$\alpha(u)^\gamma = \frac{1}{2} \left\{ \sum_{\delta=1}^{\ell} \left[\frac{\lambda_\gamma}{\lambda_\delta} \frac{\operatorname{erf}(u)}{u} - \frac{2}{\sqrt{\pi}} e^{-u^2} \right]^{-2} \right\}^{-1} , \quad (3.62)$$

onde

$$u = \frac{\lambda_\gamma}{\ell} \frac{m_\gamma}{\sqrt{2\alpha r}}$$

O alfa crítico α_c^γ , que é função de λ_γ e dos demais autovalores, é o máximo valor de α para o qual a equação (3.62) tem solução, ou seja, é simplesmente o valor máximo da função $\alpha^\gamma(u)$.

Como este valor não depende isoladamente de λ_γ devemos escolher uma matriz \underline{A} parametrizada e estudar seu comportamento com a variação do parâmetro. Por simplicidade, consideramos novamente o problema de dois blocos com \underline{A} definida em (3.52).

As figuras 3.5 e 3.6 mostram o comportamento de $\alpha(u)$ que é neste caso

$$\alpha = \frac{1}{2} \left\{ \left[\frac{\operatorname{erf}(u)}{u} - \frac{2}{\sqrt{\pi}} e^{-u^2} \right]^{-2} + \left[\frac{\varepsilon}{2 + \varepsilon} \frac{\operatorname{erf}(u)}{u} - \frac{2}{\sqrt{\pi}} e^{-u^2} \right]^{-2} \right\}^{-1} \quad (3.63)$$

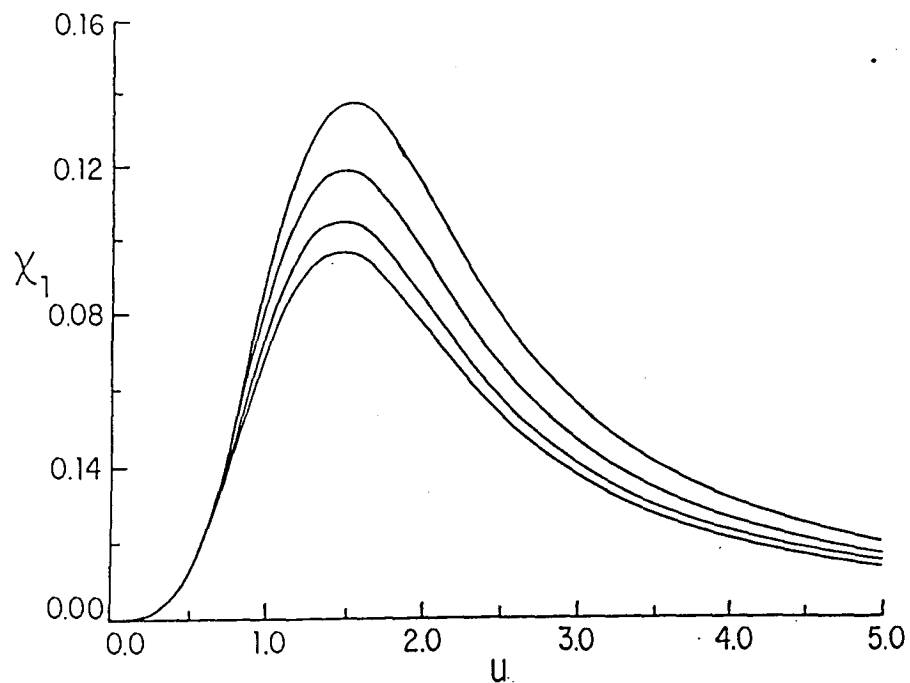


Figura 3.5: Curvas de $\alpha^1(u)$ para (de cima para baixo) $\varepsilon = 0.1, 1.6, 3.1, 4.6$

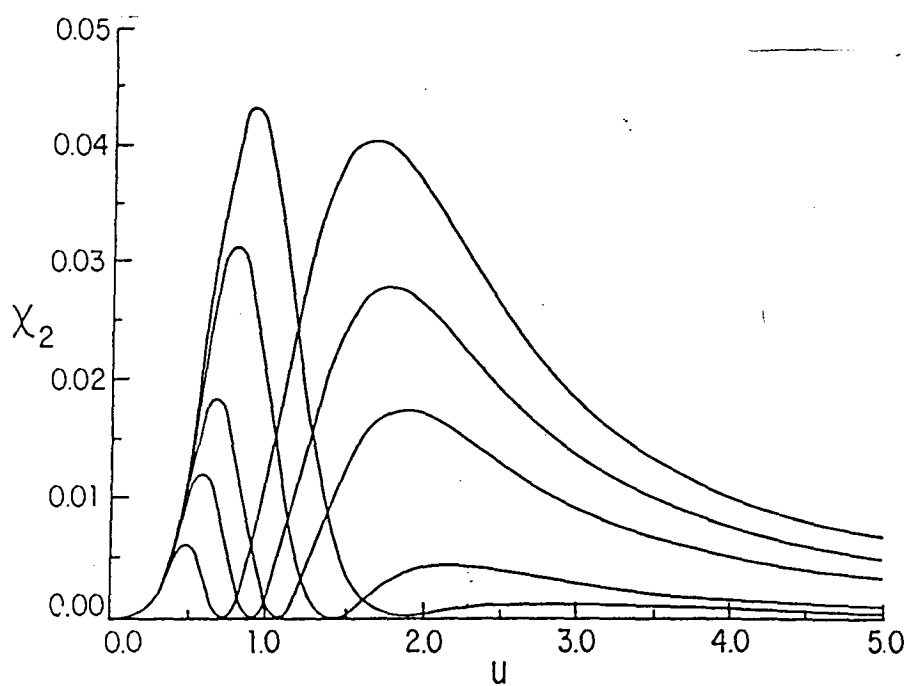


Figura 3.6: Curvas de $\alpha^2(u)$ para (de cima para baixo, na direita) $\varepsilon = 4.6, 3.1, 1.6, 0.6, 0.1$

para alguns valores de λ_1 e λ_2 .

Para saber se as soluções que aparecem nas figuras 3.5 e 3.6 são realmente estáveis devemos novamente recorrer às condições (A.78) que para $T = 0$ é

$$1 - T_1 C_\gamma > 0 \quad \gamma = 1, 2 \quad , \quad (3.64)$$

onde C_γ é dado por (3.59).

Para o caso específico das soluções em questão as condições podem ser escritas como

$$\frac{\operatorname{erf}(u)}{u} - \frac{2}{\sqrt{\pi}} e^{-u^2} > 0 \quad \text{para } T_1 = 1, \quad (3.65)$$

$$\frac{\varepsilon}{2 + \varepsilon} \frac{\operatorname{erf}(u)}{u} - \frac{2}{\sqrt{\pi}} e^{-u^2} > 0 \quad \text{para } T_2 = \frac{\varepsilon}{2 + \varepsilon}, \quad (3.66)$$

A condição (3.65) para $T_1 = 1$ (solução de recuperação ++) é sempre satisfeita salvo para $u = 0$ ou ∞ . Assim, o alfa crítico é simplesmente o máximo de cada uma das curvas que aparecem na figura 3.5.

A coisa é um pouco diferente no caso da solução descendente +-. Tipicamente, as curvas que aparecem na figura 3.6 possuem dois máximos relativos separados por um vale, onde a função assume o valor zero. Esse zero nada mais é que o valor de u^* para o qual a desigualdade (3.66) torna-se uma igualdade e, desde que a função

$$\chi(u) = \frac{\varepsilon}{2 + \varepsilon} \frac{\operatorname{erf}(u)}{u} - \frac{2}{\sqrt{\pi}} e^{-u^2}$$

é tal que é sempre positiva para $u > u^*$ e sempre negativa para $u < u^*$, a condição (3.66) significa que, na fig. 3.6, somente as soluções à direita dos vales são estáveis, e os máximos da direita são aqueles que interessam para o cálculo de α_c^2 .

Com estes resultados em mente, já é possível traçar as curvas de α_c^2 em função de ε . As figuras que seguem são respectivamente as curvas que correspondem ao problema de uma

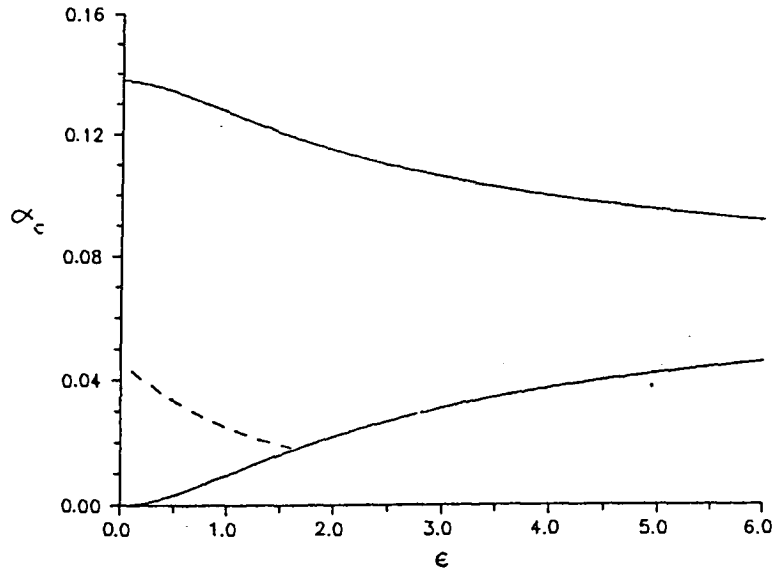


Figura 3.7: Curvas de $\alpha_c^1(\epsilon)$ e $\alpha_c^2(\epsilon)$ para uma única hierarquia . A linha pontilhada significa soluções instáveis

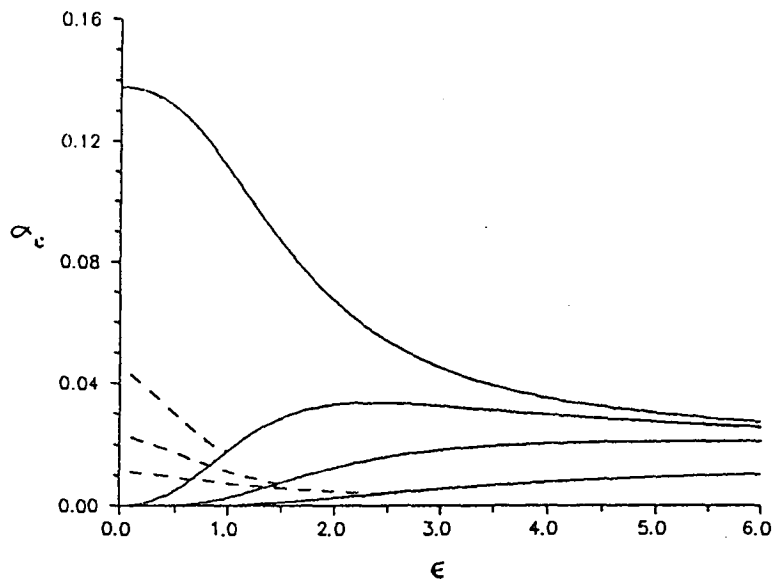


Figura 3.8: Curvas de (de cima para baixo) α_c^1 , α_c^2 , α_c^3 e α_c^4 para 3 hierarquias . A linha pontilhada significa soluções instáveis

hierarquia (dois blocos) estudado acima e aquelas para um sistema de três hierarquias (oito blocos) com a matriz \underline{A} definida tal como (2.9) com os autovalores (2.64).

O limite assintótico da figura 3.7 quando $\varepsilon \rightarrow \infty$ é

$$\alpha_c = \frac{1}{2}(0.138)$$

pois, neste caso, vamos recair no problema de duas redes de Hopfield cada uma com $\frac{N}{2}$ neurônios. Para a figura 3.8 este limite será

$$\alpha_c = \frac{1}{8}(0.138)$$

3.3.4 Condição para a existência da curva T_C^2

Observamos que, ao passarmos da figura 3.1 para a figura 3.2, há o aparecimento da curva T_C^2 . Como foi dito anteriormente esta curva marca os pontos a partir dos quais a energia livre do estado (+-) torna-se menor que aquela da solução de vidro de spin.

Nesta seção, discutiremos qual a condição, sobre o valor de λ_γ que determina a existência desta curva para uma solução geral γ . Isto será feito através da comparação entre f_γ e f_{vs} no limite em que $\alpha \rightarrow 0$.

Do capítulo 2 temos que

$$\begin{cases} f_\gamma = \frac{1}{2} \frac{\lambda_\gamma}{\ell} m^2 - \frac{1}{\beta} \ln 2 \cosh(\beta \frac{\lambda_\gamma}{\ell} m) \\ \text{com } m = \tanh(\beta \frac{\lambda_\gamma}{\ell} m) \end{cases} \quad (3.67)$$

No caso da solução de vidro de spin, temos de (3.53)

$$q = \int Dz \tanh^2 \left(\beta \sqrt{\alpha \sum_{\delta=1}^{\ell} \frac{q(\lambda_\delta/\ell)^2}{[1 - \beta \lambda_\delta(1-q)]^2}} z \right) \quad (3.68)$$

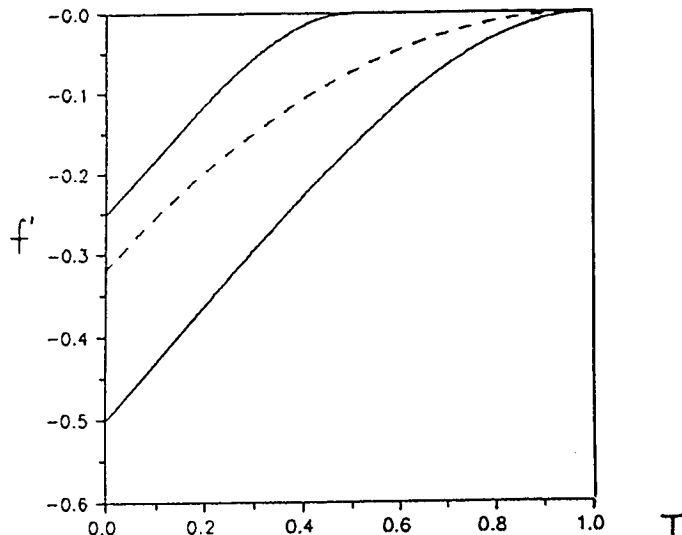


Figura 3.9: Gráficos de $f'_\gamma(T)$ para $\lambda_\gamma/\ell = 0.5, 1$ (de cima para baixo) e $f'_{vs}(T)$ (linha tracejada), onde $f' = f + T \ln 2$.

no limite $\alpha \rightarrow 0$, o Ansatz apropriado para q é

$$q_\sigma = \left(1 - T \frac{\ell}{\lambda_\sigma}\right) + C_\sigma \sqrt{\alpha} \quad (3.69)$$

onde C_σ deverá ser calculado de

$$\left(1 - T \frac{\ell}{\lambda_\sigma}\right) = \int Dz \tanh^2 \left(\frac{\sqrt{1 - T \frac{\ell}{\lambda_\sigma}}}{C_\sigma} z \right) \quad (3.70)$$

Como a condição $1 - \beta(1 - q) > 0$ deve ser sempre satisfeita, a única solução que interessa é $\sigma = 1$. A energia livre resultante daí pode ser escrita nos mesmos moldes de (3.67) como

$$\begin{cases} f_{vs} = \frac{1}{2} \left(\frac{X}{\beta}\right)^2 - \frac{1}{\beta} \int Dz \ln 2 \cosh(Xz) \\ \text{com} \quad (1 - T) = \int Dz \tanh^2(Xz) \end{cases} \quad (3.71)$$

A solução numérica de (3.67), para $\frac{\lambda_\gamma}{\ell} = 0.5, 1$ e (3.71) são apresentadas na figura 3.9. Observe que f_{vs} para $\alpha \rightarrow 0$ é independente da arquitetura da rede, representada por $\underline{\Lambda}$

e seus autovalores. Sendo assim, a condição para que f_γ a intercepte em algum ponto é

$$f_\gamma(0) < f_{vs}(0)$$

e como, de [42], sabemos que

$$f_\gamma(0) = -\frac{1}{2} \frac{\lambda}{\ell} \quad f_{sg} = -\frac{1}{\pi} \simeq -0.32 \quad ,$$

a condição final para existência de T_C^γ é apenas

$$\frac{\lambda_\gamma}{\ell} > \frac{2}{\pi} \quad . \quad (3.72)$$

Assim, para todas as soluções que satisfazem (3.72) existe uma região em $\alpha = 0$ onde suas energias livres são menores que a de vidro de spin. E portanto, existe a curva T_C^γ .

3.4 Análise de Sinal×ruído

Nesta seção, será apresentado um método alternativo para o cálculo da capacidade máxima de armazenamento de padrões em uma rede de neurônios. Ao final, será possível comparar os resultados obtidos aqui com aqueles calculados na seção anterior por campo médio.

O método consiste em calcularmos a probabilidade de que todos os padrões armazenados na rede sejam estáveis no limite $N \rightarrow \infty$. Observa-se que, esta é uma função descontínua do parâmetro $\alpha = \frac{p}{N}$, de forma que, acima de um certo valor $\alpha = \alpha_c$, ela passa descontinuamente a assumir o valor zero.

A condição para que uma dada configuração $\{S_i\}$ seja estável no sítio i é que o valor S_i esteja alinhado com seu campo local, ou seja,

$$S_i = \text{sgn}(h_i) \quad , \quad (3.73)$$

ou ainda,

$$h_i S_i > 0 \quad , \quad (3.74)$$

No nosso modelo, o campo local é dado por

$$h_i = \sum_b A_{ab} \sum_j^{(b)} J_{ij} S_j \quad , \quad i \in a \quad (3.75)$$

De acordo com nossos resultados anteriores, para cada padrão original armazenado ξ_μ o sistema é capaz de recuperar, além destes, uma família de padrões descendentes $\{\eta^{\mu\gamma}\}$, $\gamma = 1, 2, \dots, \ell$, onde

$$\eta_i^{\mu\gamma} \equiv v_a^\gamma \xi_i^\mu \quad , \quad i \in a \quad (3.76)$$

e

$$\underline{A} \cdot \bar{v}^\gamma = \lambda_\eta \bar{v}^\gamma \quad , \quad (3.77)$$

Substituindo estas soluções em (3.74), usando (3.75), obtemos para a condição de estabilidade a seguinte expressão

$$v_a^\gamma \xi_i^\mu \sum_b A_{ab} \sum_i^{(b)} \sum_\nu \xi_i^\nu \xi_j^\nu v_b^\gamma \xi_j^\mu > 0 \quad . \quad (3.78)$$

Se separarmos o termo de sinal ($\mu = \nu$) do termo de ruído, ou seja,

$$N_0 v_a^\gamma \sum_b A_{ab} v_b^\gamma + v_a^\gamma \sum_b A_{ab} v_b^\gamma \sum_j^{(b)} \sum_{\nu(\nu \neq \mu)} \xi_i^\nu \xi_j^\nu \xi_j^\mu > 0$$

$$N_0 \lambda_\gamma + v_a^\gamma \sum_b A_{ab} v_b^\gamma \chi_b > 0 \quad , \quad (3.79)$$

observamos que, como os ξ_i^μ são aleatórios com probabilidade 1/2 de ser +1 ou -1, a variável

$$\chi_b = \sum_j^{(b)} \sum_\nu \xi_i^{\nu(\nu \neq \mu)} \xi_j^\nu \xi_j^\mu$$

possuirá uma distribuição binomial que no limite em que o número de termos da soma cresce (neste caso χ_b é uma soma de $(p-1)N_0$ termos) pode ser escrita, pelo Teorema do Limite Central, como:

$$P(\chi_a) = \sqrt{\frac{1}{2\pi p N_0}} \exp\left(\frac{-\chi_a^2}{2p N_0}\right) \quad . \quad (3.80)$$

Portanto, a probabilidade de ter o neurônio do sítio $i \in a$ alinhado com seu campo local , quando toda a rede está na configuração $\{S\} = \{\eta^{\mu\gamma}\}$, é,

$$P_i^{\mu\gamma} = \int_{-\infty}^{\infty} \prod_b d\chi_b P(\chi_b) \theta \left(N_0 \lambda_\gamma + v_a^\gamma \sum_b A_{ab} \chi_b v_b^\gamma \right) , \quad (3.81)$$

que fica mais convenientemente escrita, se definirmos as novas variáveis $\chi'_b = (v_a^\gamma \chi_b v_b^\gamma) / (\sqrt{2pN_0})$, da forma

$$P_i^{\mu\gamma} = \int_{-\infty}^{\infty} \prod_b \frac{d\chi'_b}{\sqrt{\pi}} e^{-\Sigma_b (\chi'_b)^2} \theta \left(\lambda_\gamma + (2\alpha\ell)^{1/2} \sum_b A_{ab} \chi'_b \right) . \quad (3.82)$$

A integral múltipla acima pode ser transformada em uma integral simples por meio da identidade

$$1 = \int_{-\infty}^{\infty} d\xi \int_{-\infty}^{\infty} \frac{d\rho}{2\pi} \exp \{ i\rho(\xi - \Sigma_b A_{ab} \chi_b) \} . \quad (3.83)$$

Assim,

$$P_i^{\mu\gamma} = \int_{-\infty}^{\infty} d\xi \int_{-\infty}^{\infty} \frac{d\rho}{2\pi} e^{i\rho\xi} \exp \left(-\frac{\rho^2}{2} \alpha\ell \sum_b A_{ab}^2 \right) \theta(\lambda_\gamma + \xi) \quad (3.84)$$

e, finalmente, usando $z = \xi / (\sqrt{\sum_b A_{ab}^2})$ resulta

$$P_i^{\mu\gamma} = \int_{-\infty}^{\infty} dz e^{-z^2} \theta \left(z + \frac{\lambda_\gamma}{\sqrt{2\alpha\ell \sum_b A_{ab}^2}} \right) , \quad (3.85)$$

ou seja,

$$P_i^{\mu\gamma} = \frac{1}{2} \left\{ 1 + \operatorname{erf} \left(\lambda_\gamma / \left(2\alpha\ell \sum_b A_{ab}^2 \right)^{1/2} \right) \right\} \quad (3.86)$$

onde

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x dz e^{-z^2} . \quad (3.87)$$

Como sabemos, $P_i^{\mu\gamma}$ é a probabilidade que o estado $\eta^{\mu\gamma}$ seja estável no sítio i . A probabilidade que todos os $p = \alpha N$ sejam estáveis em todos os N sítios é então:

$$P(\lambda_\gamma) = \prod_{\mu} \prod_i P_i^{\mu\gamma} = (P_i^{\mu\gamma})^{\alpha N^2} \quad (3.88)$$

O único modo de $P(\lambda_\gamma)$ não ser nulo no limite termodinâmico é se $P_i^{\mu\gamma}$ tender a 1 neste limite. Pela forma funcional de (3.86) isto significa que podemos expandir $P(\lambda_\gamma)$ para valores grandes do argumento da função erro, ou seja,

$$\begin{aligned} P(\lambda_\gamma) &= \left(1 - \frac{\sqrt{\alpha}}{\sqrt{2\pi C}} e^{-\frac{C^2}{2\alpha}} + \dots \right)^{\alpha N^2} \\ &= 1 - \frac{\alpha^{3/2} N}{\sqrt{2\pi C}} e^{-\frac{C^2}{2\alpha}} + \dots \end{aligned} \quad (3.89)$$

onde $C = \lambda_\gamma / \sqrt{\sum_b A_{ab}^2}$. Se exigimos que o segundo termo acima seja menor que um certo κ finito obtemos a condição :

$$\alpha < \frac{C^2}{4 \ln N} \quad (3.90)$$

Uma das propriedades da matriz \underline{A} é que a soma $\sum_b A_{ab}^2$ não depende do índice de bloco "a", assim, podemos escrever

$$\sum_b A_{ab} A_{ba} = \frac{1}{\ell} \sum_{a,b} A_{ab} A_{ba} = \frac{1}{\ell} \text{Tr} \underline{A}^2 = \frac{1}{\ell} \sum_{\delta=1}^{\ell} \lambda_\delta^2 \quad , \quad (3.91)$$

e então,

$$[\alpha_c^\gamma]_{SR} = \frac{1}{2 \ln N} \frac{\lambda_\gamma^2}{\sum_{\delta=1}^{\ell} \lambda_\delta^2} \quad , \quad (3.92)$$

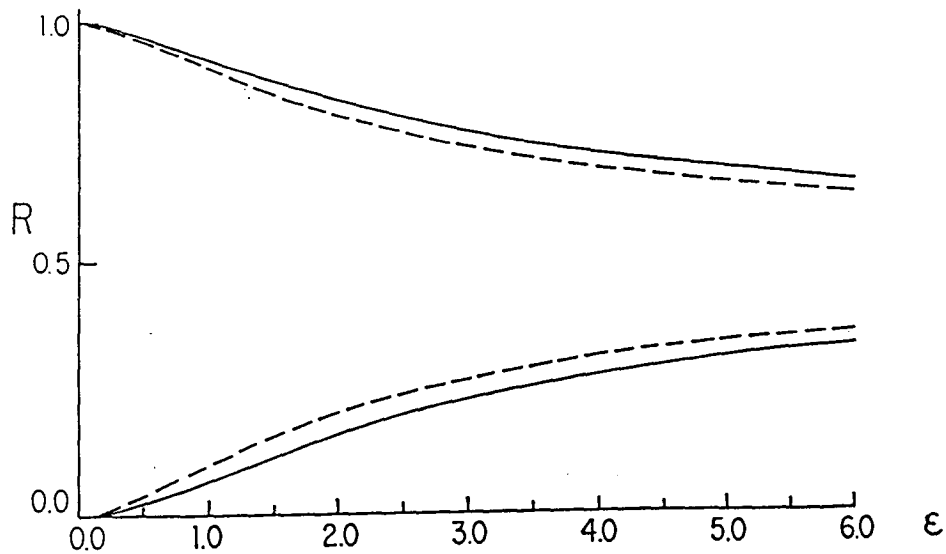


Figura 3.10: Comparação entre os resultados de campo médio (linha cheia) e sinal-ruído (linha pontilhada) para a razão $\left(\frac{\alpha_c^\gamma}{\alpha_c^{Hopfield}}\right)$ no caso de (de cima para baixo) $\gamma = 1$ e 2, para um sistema de 2 blocos.

que pode ser escrito como a razão

$$\left[\frac{\alpha_c^\gamma}{\alpha_c^{Hopfield}}\right]_{SR} = \lambda_\gamma^2 / \left(\sum_{\delta=1}^L \lambda_\delta^2\right) \quad (3.93)$$

onde nós usamos o resultado de sinal-ruído do modelo de Hopfield

$$[\alpha_c^{Hopfield}]_{SR} = (2 \ln N)^{-1} .$$

Uma boa comparação com os resultados de campo médio pode ser feita simplesmente superpondo as curvas de $\left(\frac{\alpha_c^\gamma}{\alpha_c^{Hopfield}}\right)$ em função do parâmetro ϵ , obtidos da relação acima, com aquelas das figuras 3.7 e 3.8. Isso é feito nas figuras 3.10 e 3.11.

Para o caso de uma hierarquia, os autovalores em função de ϵ são

$$\frac{\lambda_1}{2} = 1 \quad \frac{\lambda_2}{2} = \frac{\epsilon}{2 + \epsilon} ,$$

e, para o caso de três hierarquias eles estão listados em (2.64).

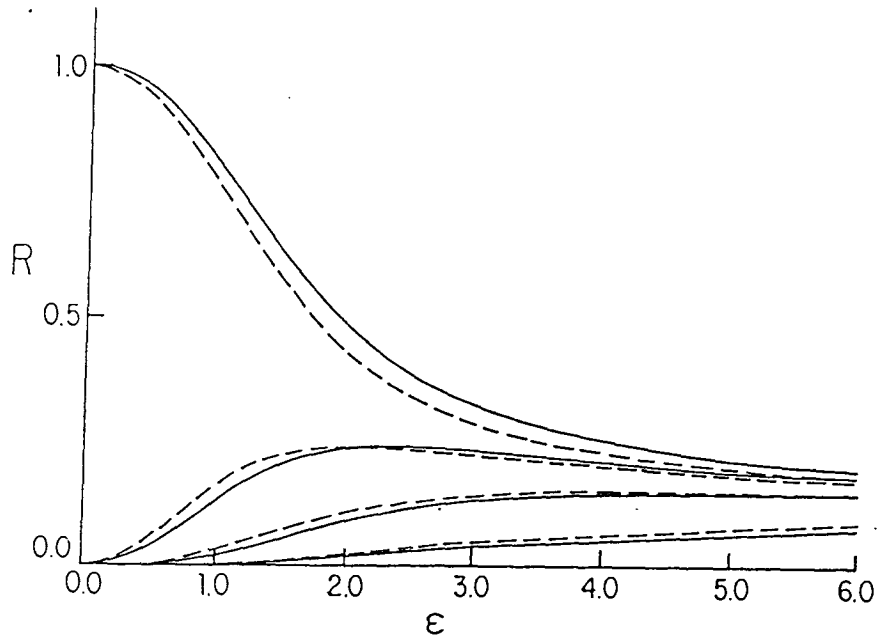


Figura 3.11: Comparação entre os resultados de campo médio (linha cheia) e sinal x ruído (linha pontilhada) para a razão $\left(\frac{\alpha_c^\gamma}{H_{op} \gamma \alpha_c}\right)$ no caso de (de cima para baixo) $\gamma = 1, 2, 3$ e 4 , para um sistema de 8 blocos.

Capítulo 4

Investigação do espaço de interações

4.1 O problema de Gardner

Até este momento, temos considerado um sistema com uma função de energia bem definida, do tipo

$$\mathcal{H} = -\frac{1}{2} \sum_{i,j} \mathcal{J}_{ij} S_i S_j \quad , \quad (4.1)$$

do qual nos interessa investigar propriedades de armazenamento e recuperação de padrões. As interações sinápticas $\{\mathcal{J}_{ij}\}$ são variáveis fixas determinadas pelo conjunto de padrões $\{\xi_i^\mu\}$ para $\mu = 1, \dots, p$ que queremos armazenar e também por algumas características arquitetônicas da rede. As variáveis $\{S_i\}$, que representam os estados de cada um dos neurônios, são as variáveis dinâmicas do sistema.

Usando técnicas de mecânica estatística, podemos, definidas as $\{\mathcal{J}_{ij}\}$, encontrar os possíveis estados estacionários do sistema. Isto é feito através da solução da equação de estado onde a quantidade a ser resolvida é o parâmetro de ordem macroscópico associado a S_i

$$m^\mu = \frac{1}{N} \sum_{\mu} \langle \xi_i^\mu \langle S_i \rangle_T \rangle_\xi \quad , \quad (4.2)$$

com $\langle \dots \rangle_T$ e $\langle \dots \rangle_\xi$ significando as médias térmica e configuracional.

Uma das mais importantes informações que podemos extrair da solução desta equação é o número máximo de padrões, $p_c = \alpha_c N$, que a rede pode armazenar, de forma que ainda tenha $m^\mu \neq 0$ como estado estacionário. Assim, em resumo, o método consiste em definir o sistema através de $\{J_{ij}\}$ e a partir daí extrair as propriedades, como por exemplo α_c .

O método de Gardner [44] busca resolver o problema inverso. Considera que tenhamos como variáveis fixas os estados de neurônios, que são dados pelas memórias que desejamos armazenar $S_i = \xi_i^\mu$. As quantidades dinâmicas, aquelas que permitiremos variarem, passam a ser os acoplamentos $\{J_{ij}\}$, que a princípio não são necessariamente simétricos, isto é, em geral $J_{ij} \neq J_{ji}$. Assim constrói-se equações de estado para quantidades macroscópicas relacionadas com J_{ij} , a partir da condição que $S_i = \xi_i^\mu$ sejam os estados estacionários da dinâmica

$$S_i(t) = \text{sgn}(h_i(t + \Delta t)) \quad , \quad (4.3)$$

com o campo h_i dado por

$$h_i(t) = \frac{1}{\sqrt{N}} \sum_j J_{ij} S_j(t) \quad , \quad (4.4)$$

para todo i e μ .

A solução destas equações também nos leva a determinação de uma fração crítica de armazenamento α_c . Este valor é um limite superior para qualquer modelo. Isso porque o método consiste em varrer o espaço $N(N-1)$ -dimensional dos $\{J_{ij}\}$, que é na verdade

o espaço de todos os possíveis modelos, a procura do modelo ótimo, o qual armazenará o maior número de padrões.

Apesar da determinação deste limite superior ser muito importante, a sua validade na prática é pouca pois, já que \mathcal{J}_{ij} são variáveis dinâmicas, não temos acesso ao valores delas que caracterizariam o modelo ótimo. O modo de contornar este problema é usar este método em conjunto com algum algoritmo [44] que permita a construção das soluções $\{\mathcal{J}_{ij}\}$, coisa que não faremos.

A aplicação do método de Gardner para o caso do armazenamento de padrões decorrelacionados, ou seja, tais que

$$\langle \xi^\mu \xi^\nu \rangle = 0 \quad \mu \neq \nu ,$$

onde o espaço dos possíveis modelos está restrito à condição esférica

$$\sum_j \mathcal{J}_{ij}^2 = N \quad i = 1, \dots, N \quad , \quad (4.5)$$

tem como principal resultado que existe um conjunto de escolha das sinápses $\{\mathcal{J}_{ij}\}$ para o qual é possível armazenar um número máximo de $p_c = 2N$ padrões. Observe que o valor $\alpha_c = 2$ é bem maior que aquele previsto pela análise de sinal×ruído da regra de Hebb, o que mostra que esta regra está longe de ser a ótima.

O que faremos neste capítulo é calcular o armazenamento crítico para o caso em que o espaço das possíveis escolhas das interações sinápticas é mais restrito que aquele originado da condição esférica (4.5). Esta condição mais restrita tem como objetivo simular uma estrutura hierárquica semelhante aquela introduzida na seção 2.1.

4.2 Método de Gardner para um sistema com hierarquia nas interações

A principal característica do modelo que temos estudado é sua estrutura espacial. Esta foi bem representada no capítulo 2 pela dependência, via a matriz \underline{A} , da sua intensidade de memorização por sinapse (2.15) com os índices dos blocos onde estão os neurônios que se ligam pela respectiva sinapse. Agora, por conveniência, vamos trabalhar com outra quantidade igualmente representativa, também definida na referência [36], denominada “intensidade sináptica acumulada”, no nosso modelo dada por

$$\| \mathcal{J} \|_{ij}^2 \equiv \langle \mathcal{J}_{ij}^2 \rangle_{\xi} = \alpha (A_{ab})^2 , \quad (4.6)$$

onde foi usado

$$\mathcal{J}_{ij} = A_{ab} J_{ij}^H , \quad (4.7)$$

com a regra de Hebb devidamente normalizada

$$J_{ij}^H = \frac{1}{\sqrt{N}} \sum_{\mu} \xi_i^{\mu} \xi_j^{\mu} . \quad (4.8)$$

Podemos estabelecer um paralelo entre esta quantidade e o tipo de vínculo que deve ser utilizado se quisermos determinar o conjunto de todos os pontos $\{\mathcal{J}_{ij}\}$ que representam modelos com interações hierárquicas. No modelo de Hopfield, a soma sobre “i” da intensidade sináptica resulta

$$\sum_i \| \mathcal{J} \|_{ij}^2 = \alpha N , \quad (4.9)$$

e a generalização para um espaço maior, que contenha este modelo, é imediata [44], dada por

$$\sum_j \mathcal{J}_{ij} = N \quad , \quad (4.10)$$

onde α foi absorvido em \mathcal{J}_{ij} e N é essencial para que o campo (4.4) sobre cada neurônio seja $o(1)$ no limite termodinâmico. No nosso modelo o modo que consideramos ideal para esta generalização é quebrar o vínculo esférico (4.10), que acopla N variáveis, em ℓ vínculos esféricos sobre $N_0 = N/\ell$ variáveis cada um deles dado por

$$\sum_j^{(b)} \mathcal{J}_{ij} = (A_{ab})^2 N_0 \quad \cdot i \in a \quad . \quad (4.11)$$

É fácil verificar que o modelo definido por (4.7) está contido no espaço delimitado por estes vínculos. O volume deste espaço é:

$$\Omega_0 = \int \prod_{i,j} d\mathcal{J}_{ij} \prod_{a,b} \prod_i^{(b)} \delta \left(\sum_j^{(a)} \mathcal{J}_{ij}^2 - A_{ab}^2 N_0 \right) \quad . \quad (4.12)$$

O próximo passo seria determinar qual a probabilidade que, dados os p padrões $\{\xi_i^\mu\}$, um certo ponto $\{\mathcal{J}_{ij}\}$ escolhido neste volume os tenha como estados metaestáveis da dinâmica (4.3). Esta probabilidade é proporcional a fração de volume que satisfaz a condição

$$\xi_i^\mu h_i^\mu > \kappa \quad \forall i, \mu \quad , \quad (4.13)$$

onde $\kappa > 0$ significa que estamos interessados em que estas configurações além de serem estáveis também possuam uma bacia de atração finita.

Esta fração de volume é

$$V_T = \frac{\Omega(\alpha, \kappa)}{\Omega_0} \quad , \quad (4.14)$$

onde

$$\Omega(\alpha, \kappa) = \int \prod_{i,j} d\mathcal{J}_{ij} \prod_{i,\mu} \Theta(\xi_i^\mu h_i^\mu - \kappa) \prod_{a,b} \prod_i^{(b)} \delta \left(\sum_j^{(a)} \mathcal{J}_{ij}^2 - A_{ab}^2 N_0 \right) \quad . \quad (4.15)$$

A fração V_T , e conseqüentemente $\Omega(\alpha, \kappa)$, devem se colapsar em zero para uma dada curva $\alpha_c = \alpha_c(\kappa)$ que representa a fração crítica de armazenamento. O valor máximo desta função será atingido quanto permitirmos armazenagem com bacia de atração zero, ou seja, no caso de $\kappa = 0$.

Apesar de podermos extrair de V_T toda a informação necessária para obter α_c vamos, optar por apresentar um formalismo ligeiramente diferente, mas que propicia um melhor entendimento do método em termos da mecânica estatística. Considere o problema acima como sendo equivalente aquele de determinar as propriedades macroscópicas de um sistema isolado. Para isso devemos utilizar o ensemble microcanônico, que nesse caso é composto de um grande número de sistemas idênticos, dos quais a única informação disponível é que, para cada um deles, as p configurações $\{\xi_i^\mu\}$ satisfazem a condição (4.13). A questão fundamental neste caso é determinar qual é a probabilidade de, ao escolher um destes sistemas, obtermos a configuração sináptica $\{\mathcal{J}_{ij}\}$. Esta densidade de probabilidade é

$$\mathcal{P}(\{\mathcal{J}_{ij}\}) = \frac{1}{\Omega(\alpha, \kappa)} \prod_{i,\mu} \Theta(\xi_i^\mu h_i^\mu - \kappa) \prod_{a,b} \prod_i^{(b)} \delta \left(\sum_j^{(a)} \mathcal{J}_{ij}^2 - A_{ab}^2 N_0 \right) \quad . \quad (4.16)$$

O valor de \mathcal{P} é zero para todo os pontos $\{\mathcal{J}_{ij}\}$ que não satisfaçam (4.13) ou que não estejam contidos no volume (4.12). Qualquer informação macroscópica deve ser obtida a partir das médias de ensemble

$$\bar{O} = \left\langle \int d\mathcal{J}_{ij} O(\{\mathcal{J}_{ij}\}) \mathcal{P}(\{\mathcal{J}_{ij}\}) \right\rangle_\xi \quad , \quad (4.17)$$

que podem ser calculadas de uma única quantidade básica que é a função geradora destas médias, dada por

$$G = \frac{1}{N} \langle \ln \Omega(\alpha, \kappa, \{y^{ij}\}) \rangle_{\xi} \quad , \quad (4.18)$$

onde y^{ij} é o campo associado a J_{ij} .

A função G , para $y^{ij} = 0$, é uma função análoga à entropia associada ao espaço de estados definido pelo volume $\Omega(\alpha, \kappa)$. Como o sistema é completamente conectado, podemos calcular esta função exatamente com auxílio do método de ponto de sela. As equações de ponto de sela, por sua vez, serão como equações de estado, derivadas a partir do critério de máxima entropia, das quais será possível tirarmos as informações que buscamos.

4.2.1 Determinação de α_c no caso de armazenamento de padrões descorrelacionados

Estabelecido o método, vamos passar a sua aplicação. Observando que o volume $\Omega(\alpha, \kappa)$ é separável na forma de um produtório sobre o sítio "i" podemos escrever

$$G = \frac{1}{N} \langle \ln \Omega(\alpha, \kappa) \rangle_{\xi} = \frac{1}{\ell} \sum_a \langle \ln \Omega_a(\alpha, \kappa) \rangle_{\xi} \quad . \quad (4.19)$$

O cálculo de G , analogamente com o que aconteceu com f no capítulo 3, exige que usemos réplicas, ou seja

$$G = \frac{1}{\ell} \sum_a \lim_{n \rightarrow \infty} \frac{\langle \Omega_a^n \rangle_{\xi} - 1}{n} \quad , \quad (4.20)$$

onde omitimos, por economia, os argumentos de Ω_a .

O volume replicado é, então,

$$\langle \Omega_a^n \rangle_\xi = \int \prod_{j,\rho} dJ_j^\rho \prod_\mu \langle \prod_\rho \Theta(\xi^\mu h^{\mu,\rho} - \kappa) \rangle_\xi \prod_{b,\rho} \delta \left(\sum_j^{(b)} (J_j^\rho)^2 - A_{ab}^2 N_0 \right) \quad (4.21)$$

e, usando

$$\Theta(y - \kappa) = \int_\kappa^\infty d\lambda \int_{-\infty}^\infty \frac{dx}{2\pi} e^{ix(\lambda-y)} \quad , \quad (4.22)$$

podemos realizar a média sobre os padrões resultando

$$\begin{aligned} \textcircled{a} &= \langle \prod_\rho \Theta(\xi h^\rho - \kappa) \rangle_\xi \\ &= \int_\kappa^\infty \prod_\rho d\lambda^\rho \int_{-\infty}^\infty \prod_\rho \frac{dx^\rho}{2\pi} e^{i \sum_\rho x^\rho \lambda^\rho} \prod_j \cos\left(\frac{1}{\sqrt{N}} \sum_\rho x^\rho J_j^\rho\right) \quad . \end{aligned} \quad (4.23)$$

Expandindo o cosseno até ordem $o(1/N)$ obtemos

$$\begin{aligned} \textcircled{a} &= \int_\kappa^\infty \prod_\rho d\lambda^\rho \int_{-\infty}^\infty \prod_\rho \frac{dx^\rho}{2\pi} \times \\ &\exp\left(i \sum_\rho x^\rho \lambda^\rho - \frac{1}{2} \sum_{\rho,\rho'} x^\rho x^{\rho'} \frac{1}{\ell} \sum_b (A_{ab}^2 \delta^{\rho\rho'} + q_{ab}^{\rho\rho'} (1 - \delta^{\rho\rho'}))\right) \quad , \end{aligned} \quad (4.24)$$

onde foi introduzindo o parâmetro

$$q_{ab}^{\rho\rho'} = \frac{1}{N_0} \sum_j^{(b)} \langle J_j^\rho J_j^{\rho'} \rangle_\xi$$

através de

$$1 = \int dq_{ab}^{\rho\rho'} \delta\left(q_{ab}^{\rho\rho'} - \frac{1}{N_0} \sum_j^{(b)} J_j^\rho J_j^{\rho'}\right) \quad . \quad (4.25)$$

Assim, finalmente obtemos

$$\langle \Omega_a^n \rangle_\xi = \int d\underline{E} d\underline{F} d\underline{q} \exp N \left\{ \frac{1}{2\ell} \sum_b (A_{ab}^2 \sum_\rho E_{ab}^\rho - \sum_{\rho \neq \rho'} q_{ab}^{\rho\rho'} F_{ab}^{\rho\rho'} + 2 \mathcal{V}_b) + \alpha \mathcal{Q} \right\} \quad (4.26)$$

onde $d\underline{E}$ significa a integral sobre todas as variáveis E_{ab}^ρ , o mesmo valendo para $d\underline{F}$ e $d\underline{q}$, e

$$\mathcal{V}_b = \ln \int d\underline{\mathcal{J}} \exp \left(-\frac{1}{2} \left(\sum_\rho E_{ab}^\rho (\mathcal{J}^\rho)^2 + \sum_{\rho \neq \rho'} F_{ab}^{\rho\rho'} \mathcal{J}^\rho \mathcal{J}^{\rho'} \right) \right) . \quad (4.27)$$

A região de α para recuperação perfeita é tal que a solução de simetria de réplicas é sempre estável [45]. Assim, fazendo $E_{ab}^\rho = E_{ab}$, $F_{ab}^{\rho\rho'} = F_{ab}$ e $q_{ab}^{\rho\rho'} = q_{ab}$ para as soluções do ponto de sela, podemos escrever

$$\mathcal{Q} = \int Dz \left\{ \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} \frac{d\lambda}{2\pi} \exp \left(ixy - \frac{1}{2} x^2 \frac{1}{\ell} \sum_b (A_{ab}^2 - q_{ab}) + i \sqrt{\frac{1}{\ell} \sum_b q_{ab}} x z \right) \right\}^n, \quad (4.28)$$

onde linearizamos a variável x^ρ através de

$$\exp \left\{ \frac{1}{\ell} \sum_b q_{ab} (\sum_\rho x^\rho)^2 \right\} = \int Dz \exp \left\{ i \sqrt{\frac{1}{\ell} \sum_b q_{ab}} \sum_\rho x^\rho z \right\} , \quad (4.29)$$

com a notação

$$\int Dz \dots = \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} e^{\frac{1}{2} z^2} .$$

Se considerarmos primeira ordem em “n” obtemos

$$\mathcal{Q} = n \int Dz \ln H \left(\frac{\sqrt{\frac{1}{\ell} \sum_b q_{ab} z + \kappa}}{\sqrt{\frac{1}{\ell} \sum_b (A_{ab}^2 - q_{ab})}} \right) , \quad (4.30)$$

onde

$$H(x) = \int_x^\infty Dz e^{-\frac{1}{2}x^2} . \quad (4.31)$$

De forma semelhante, podemos integrar \mathcal{J} em \mathcal{V}_b resultando

$$\mathcal{V}_b = \frac{n}{2} \left\{ \ln \left(\frac{2\pi}{E_{ab} + F_{ab}} \right) + \frac{F_{ab}}{E_{ab} - F_{ab}} \right\} , \quad (4.32)$$

e, portanto,

$$\begin{aligned} \langle \Omega_a^n \rangle_\xi &= \int d\underline{E} d\underline{F} d\underline{q} \times \\ &\exp nN \left\{ \frac{1}{2\ell} \sum_b \left(A_{ab}^2 E_{ab} + q_{ab} F_{ab} + \ln \left(\frac{2\pi}{E_{ab} + F_{ab}} \right) + \frac{F_{ab}}{E_{ab} - F_{ab}} \right) + \right. \\ &\left. \alpha \int Dz \ln H(\zeta) \right\} \end{aligned} \quad (4.33)$$

com

$$\zeta = \frac{\sqrt{\frac{1}{\ell} \sum_b q_{ab} z + \kappa}}{\sqrt{\frac{1}{\ell} \sum_b (A_{ab}^2 - q_{ab})}} . \quad (4.34)$$

Substituindo a expressão (4.33) em (4.20) , usando o limite termodinâmico e após, o limite de réplicas, obtemos para a grandeza $g = \frac{G}{N}$:

$$g = \frac{1}{2\ell} \sum_b \left(A_{ab}^2 E_{ab} + q_{ab} F_{ab} + \ln \left(\frac{2\pi}{E_{ab} + F_{ab}} \right) + \frac{F_{ab}}{E_{ab} - F_{ab}} \right) + \alpha \int Dz \ln H(\zeta). \quad (4.35)$$

As equações de ponto de sela para as variáveis E_{ab} e F_{ab} são algébricas. Logo podemos resolvê-las e substituí-las em g , fazendo que esta função dependa somente de q_{ab} . Essas equações são:

$$\frac{\partial g}{\partial F_{ac}} = \frac{1}{2\ell} \left(q_{ac} - \frac{F_{ac}}{(E_{ac} + F_{ac})^2} \right) = 0 , \quad (4.36)$$

$$\frac{\partial g}{\partial E_{ac}} = \frac{1}{2\ell} \left(A_{ac}^2 - \frac{1}{E_{ac} + F_{ac}} - \frac{F_{ac}}{(E_{ac} + F_{ac})^2} \right) = 0 . \quad (4.37)$$

E o resultado da substituição é

$$F_{ac} = \frac{q_{ac}}{(A_{ac}^2 - q_{ac})^2} , \quad (4.38)$$

$$E_{ac} = \frac{1}{A_{ac}^2 - q_{ac}} - \frac{q_{ac}}{(A_{ac}^2 - q_{ac})^2} , \quad (4.39)$$

Voltando a g temos

$$g = \frac{1}{\ell} \sum_b \left\{ \ln \frac{A_{ab}^2}{A_{ab}^2 - q_{ab}} - \frac{q_{ab}}{A_{ab}^2 - q_{ab}} \right\} + \alpha \int Dz \ln H(\zeta) , \quad (4.40)$$

com q_{ab} dado por

$$\frac{\partial g}{\partial q_{ac}} = \frac{1}{2\ell} \frac{q_{ab}}{(A_{ab}^2 - q_{ab})^2} - \alpha \int Dz \frac{e^{\frac{1}{2}\zeta^2}}{\sqrt{2\pi} H(\zeta)} \left(\frac{\partial \zeta}{\partial q_{ac}} \right) = 0 , \quad (4.41)$$

onde

$$\frac{\partial \zeta}{\partial q_{ac}} = \frac{1}{\left(\sqrt{\frac{1}{\ell} \sum_b (A_{ab}^2 - q_{ab})} \right)^3} \left(\frac{\frac{1}{\ell} \sum_b A_{ab}^2}{\sqrt{\frac{1}{\ell} \sum_b q_{ab}}} z + \kappa \right) . \quad (4.42)$$

É imediato verificar que

$$\frac{\partial \zeta}{\partial q_{ac}} = \frac{1}{\sqrt{\frac{1}{\ell} \sum_b (A_{ab}^2 - q_{ab})} \sqrt{\frac{1}{\ell} \sum_b q_{ab}}} \frac{\partial}{\partial z} \left(\frac{1}{2} z^2 + \frac{1}{2} \zeta^2 \right) , \quad (4.43)$$

logo podemos integrar z por partes em (4.41), e obter uma expressão mais compacta para o conjunto de equações de estado que determinam q_{ab} , ou seja,

$$\frac{1}{2\ell} \frac{q_{ab}}{(A_{ab}^2 - q_{ab})^2} = \frac{1}{\frac{1}{\ell} \sum_b (A_{ab}^2 - q_{ab})} \frac{\alpha}{2\pi} \int Dz \frac{1}{H^2(\zeta)} e^{-\zeta^2} \quad b = 1, \dots, \ell. \quad (4.44)$$

Considerando que o lado direito de cada uma das equações é completamente independente do índice "b", resulta que podemos escrevê-las como

$$\frac{q_{a1}}{(A_{a1}^2 - q_{a1})^2} = \frac{q_{a2}}{(A_{a2}^2 - q_{a2})^2} = \dots = \frac{q_{a\ell}}{(A_{a\ell}^2 - q_{a\ell})^2} = \frac{1}{\epsilon^2} , \quad (4.45)$$

onde definimos

$$\frac{1}{\epsilon^2} = \frac{1}{\frac{1}{\ell} \sum_b (A_{ab}^2 - q_{ab})} \frac{\alpha}{2\pi} \int Dz e^{-\frac{1}{2}\zeta^2} H^2(\zeta) . \quad (4.46)$$

Estamos interessados na determinação do valor máximo de α_c . Esta informação é resultado de tomarmos o limite $\epsilon \rightarrow 0$, ou seja, q_{ab} no seu valor máximo A_{ab}^2 .

De (4.45) nós temos

$$q_{ab} = A_{ab}^2 + \frac{\epsilon^2}{2} \pm \frac{1}{2} \sqrt{4A_{ab}^2 \epsilon^2 + \epsilon^4} , \quad (4.47)$$

onde devemos escolher o sinal negativo, pois sabemos que $q_{ab} \leq A_{ab}^2$. Para ϵ próximo de zero

$$q_{ab} \simeq A_{ab}^2 - A_{ab}\epsilon \quad . \quad (4.48)$$

A quantidade ζ torna-se

$$\zeta \simeq \frac{\sqrt{\frac{1}{\ell} \sum_b A_{ab}^2} z + \kappa}{\sqrt{\frac{1}{\ell} \sum_b A_{ab}} \epsilon} \quad , \quad (4.49)$$

e, se usarmos a expressão assintótica de $H(x)$ para argumentos grandes,

$$H(x) \simeq \Theta(-x) + \frac{e^{\frac{1}{2}x^2}}{\sqrt{2\pi} x} \quad , \quad (4.50)$$

nós finalmente obtemos a expressão de como varia α_c com o parâmetro κ , que dá uma medida do tamanho da bacia de atração

$$\alpha_c^{-1} = \frac{\frac{1}{\ell} \sum_b A_{ab}^2}{\left(\frac{1}{\ell} \sum_b A_{ab}\right)^2} \int_{-\kappa'}^{\infty} Dz (z - \kappa')^2 \quad , \quad (4.51)$$

onde $\kappa' = \kappa / \sqrt{\frac{1}{\ell} \sum_b A_{ab}^2}$.

Usando a expressão dos autovalores da matriz \underline{A} podemos ainda escrever

$$\frac{1}{\ell} \sum_b A_{ab}^2 = \frac{1}{\ell^2} \sum_{\gamma} \lambda_{\gamma}^2$$

e

$$\frac{1}{\ell} \sum_b A_{ab} = \frac{\lambda_1}{\ell}.$$

No limite que o tamanho da bacia de atração vai a zero temos

$$\alpha_c(0) = 2 \frac{\lambda_1^2}{\sum_{\gamma} \lambda_{\gamma}^2} \quad , \quad (4.52)$$

que é perfeitamente comparável com o que obtivemos na análise de sinal \times ruído, para o caso da recuperação do padrão ancestral. Vemos que

$$\left[\frac{\alpha_c}{\alpha_c^{Gardner}} \right]_{\text{Método de Gardner}} = \left[\frac{\alpha_c}{\alpha_c^{Hopfield}} \right]_{\text{Sinal-ruído}} \quad (4.53)$$

Como $\frac{\lambda_1^2}{\sum_{\gamma} \lambda_{\gamma}^2}$ é sempre menor ou igual a um, concluímos que a introdução da estrutura espacial, pelo menos nos moldes que viemos aplicando, sempre se traduz numa piora da capacidade de armazenamento da rede. No caso específico do problema de Gardner, isso mostra que a escolha ótima de padrões, aquela que possuiria uma fração máxima de armazenagem $\alpha_c^{max} = 2$, não se encontra no volume definido por (4.12).

A semelhança entre o resultado obtido aqui e aquele da análise de sinal \times ruído nos estimula a seguir um pouco mais adiante.

4.2.2 O Método de Gardner e a Armazenagem de Padrões Descendentes

Na análise de campo médio do hamiltoniano

$$\mathcal{H} = -\frac{1}{2} \sum_{ab} A_{ab} \sum_i^{(a)} \sum_j^{(b)} J_{ij} S_i S_j \quad , \quad (4.54)$$

observou-se que, além das p memórias $\{\xi_i^{\mu}\}$ armazenadas via regra de Hebb, que denominamos padrões ancestrais, o sistema possuía outros estados metaestáveis de configurações

microscópicas conhecidas. Para um tipo particular deles, que são aqueles escritos como

$$\eta_i^{\mu,\gamma} = \xi_i^\mu v_a^\gamma \quad i \in a \quad , \quad (4.55)$$

onde \bar{v}^γ é autovetor de \underline{A} com autovalor λ^γ , foi utilizada a análise sinal×ruído para determinar a capacidade máxima de armazenamento no caso de uma recuperação perfeita.

O resultado obtido foi

$$\alpha_c^\gamma = \frac{\lambda_\gamma^2}{\sum_\delta \lambda_\delta^2} \alpha_c^{Hopfield} \quad , \quad (4.56)$$

onde α_c^γ é a armazenagem crítica para $\{\eta^{\mu,\gamma}\}$.

Duas interessantes questões que se poderia colocar são: Existe possibilidade de investigarmos o armazenamento de descendentes usando o método de Gardner? E se caso isso for possível, o resultado não seria de alguma forma, a exemplo do que obtivemos para o caso dos padrões ancestrais, semelhante a (4.56)?

Se substituíssemos (4.55) diretamente na expressão para a condição de metaestabilidade (4.13) e seguíssemos calculando, chegaríamos ao mesmo resultado que obtivemos usando os padrões $\{\xi_i^\mu\}$. Isto porque o método de Gardner não distingue entre eles. Tanto $\{\eta^{\mu,\gamma}\}$ como $\{\xi_i^\mu\}$ não passam de dois conjuntos de p padrões descorrelacionados entre si. No caso da análise de sinal \times ruído, tal distinção é feita explicitamente quando usamos a regra de Hebb, a qual especifica quais são as configurações armazenadas como padrões ancestrais. Sendo assim, se quisermos estudar as propriedades de armazenagem de descendentes pelo método de Gardner devemos não só exigir a condição de metaestabilidade para estes como também para os seus ancestrais. Além disso, o problema deve, de alguma forma, carregar a informação de quais padrões são ancestrais e quais são descendentes.

Nesta seção estudaremos o caso mais simples de armazenagem de ancestrais e descendentes, que consiste em armazenarmos p padrões $\{\xi_i^\mu\}$ junto com outros p padrões $\{\eta^{\mu,2}\}$,

que são construídos a partir dos primeiros pela inversão do estado dos últimos $N/2$ sítios, ou seja:

$$\eta_i^{\mu,2} = \begin{cases} +\xi_i^\mu & i \leq \frac{N}{2} \\ -\xi_i^\mu & i > \frac{N}{2} \end{cases} . \quad (4.57)$$

A distinção entre quais são os padrões ancestrais e quais os descendentes faremos através do tamanho da bacia de atração. A razão disto é que observamos no capítulo 2 que os primeiros possuem sempre menor energia livre e, conseqüentemente [19], maior bacia de atração que os outros.

Assim, o volume de interesse no espaço de interações deverá ser

$$\begin{aligned} \Omega(\alpha, \kappa_1, \kappa_2) &= \int \prod_{i,j} d\mathcal{J}_{ij} \prod_{i,\mu} \Theta(\xi_i^\mu h_i^{1,\mu} - \kappa_1) \\ &\times \Theta(\eta_i^{2,\mu} h_i^{2,\mu} - \kappa_2) \prod_{a,b} \prod_i^{(b)} \delta \left(\sum_j^{(a)} \mathcal{J}_{ij}^2 - A_{ab}^2 N_0 \right) , \end{aligned} \quad (4.58)$$

com $\kappa_1 > \kappa_2$, $a, b = 1, 2$ e

$$h_i^{1,\mu} = \frac{1}{\sqrt{N}} \sum_j \mathcal{J}_{ij} \xi_j^\mu , \quad h_i^{2,\mu} = \frac{1}{\sqrt{N}} \sum_j \mathcal{J}_{ij} \eta_j^{2,\mu} .$$

Como estamos considerando um sistema com uma única hierarquia (dois blocos), a matriz \underline{A} é de dimensão dois, dada por

$$\underline{A} = \begin{bmatrix} A_1 & A_2 \\ A_2 & A_1 \end{bmatrix} , \quad (4.59)$$

com $A_1 > A_2$.

Seguindo os passos da seção anterior devemos calcular $\langle \Omega^n(\alpha, \kappa_1, \kappa_2) \rangle_\xi$ que pode ser posta na forma mais conveniente para aplicação do ponto de sela,

$$\langle \Omega_a^n \rangle_\xi = \int d\underline{E} d\underline{F} d\underline{q} \exp N \left\{ \frac{1}{2\ell} \sum_b (A_{ab}^2 \sum_\rho E_{ab}^\rho - \sum_{\rho \neq \rho'} q_{ab}^{\rho\rho'} F_{ab}^{\rho\rho'} + 2 \eta_b) + \alpha @^* \right\} \quad (4.60)$$

Esta expressão é idêntica a (4.26), salvo que temos $@^*$ ao invés de $@$ definido em (4.23), e as somas são sobre apenas dois blocos. Neste caso $@^*$ é dado por

$$@^* = \langle \prod_\rho \Theta(\xi_i^\mu h_i^{1,\mu} - \kappa_1) \Theta(\eta_i^{2,\mu} h_i^{2,\mu} - \kappa_2) \rangle_\xi \quad . \quad (4.61)$$

Expressando as funções $\Theta(x)$ em termos de integrais (4.22), operando as médias sobre as variáveis ξ e tomando primeira ordem em $1/N$ obtemos

$$@^* = \int_{\kappa_1}^\infty \prod_\rho d\lambda_1^\rho \int_{\kappa_2}^\infty \prod_\rho d\lambda_2^\rho \int_{-\infty}^\infty \prod_\rho \frac{dx_1^\rho}{2\pi} \frac{dx_2^\rho}{2\pi} \exp \left(i \sum_\rho (x_1^\rho \lambda_1^\rho + x_2^\rho \lambda_2^\rho) - \frac{1}{2} \sum_{\rho, \rho'} \sum_b \frac{x_1^\rho + v_a^2 v_b^2 x_2^\rho}{\sqrt{2}} \frac{x_1^{\rho'} + v_a^2 v_b^2 x_2^{\rho'}}{\sqrt{2}} (A_{ab}^2 \delta^{\rho\rho'} + q_{ab}^{\rho\rho'} (1 - \delta^{\rho\rho'})) \right) \quad (4.62)$$

Se for suposto simetria de réplicas $E_{ab}^\rho = E_{ab}$, $F_{ab}^{\rho\rho'} = F_{ab}$ e $q_{ab}^{\rho\rho'} = q_{ab}$, considerarmos $a = 1$, suprimindo este índice, podemos fazer a seguinte troca de variáveis

$$\bar{x}_1^\rho = \frac{x_1^\rho + x_2^\rho}{\sqrt{2}} \quad \bar{x}_2^\rho = \frac{x_1^\rho - x_2^\rho}{\sqrt{2}} \quad (4.63)$$

$$\bar{\lambda}_1^\rho = \frac{\lambda_1^\rho + \lambda_2^\rho}{\sqrt{2}} \quad \bar{\lambda}_2^\rho = \frac{\lambda_1^\rho - \lambda_2^\rho}{\sqrt{2}} \quad (4.64)$$

que resulta em

$$\begin{aligned}
 @^* &= \int_{\frac{\kappa_1 + \kappa_2}{\sqrt{2}}}^{\infty} \prod_{\rho} d\bar{\lambda}_1^{\rho} \int_{\sqrt{2}\kappa_1 - \bar{\lambda}_1}^{\bar{\lambda}_1 - \sqrt{2}\kappa_2} \prod_{\rho} d\bar{\lambda}_2^{\rho} \int_{-\infty}^{\infty} \prod_{\rho} \frac{dx_1^{\rho}}{2\pi} \frac{dx_2^{\rho}}{2\pi} \\
 &\times \exp \left(i \sum_{\rho, c} (x_c^{\rho} \lambda_c^{\rho} - \frac{1}{2} (A_c^2 - q_c) (\bar{x}_c^{\rho})^2) \right) \\
 &\times \int Dz_1 Dz_2 \exp \left(i \sum_c \sqrt{q_c} z_c \sum_{\rho} \bar{x}_b^{\rho} \right) . \tag{4.65}
 \end{aligned}$$

onde a introdução de z_1 e z_2 permite a integração nas variáveis \bar{x} ,

$$@^* = \int Dz_1 Dz_2 \left\{ \int_1 \frac{d\bar{\lambda}_1}{\sqrt{2\pi\epsilon_1}} \int_2 \frac{d\bar{\lambda}_2}{\sqrt{2\pi\epsilon_2}} \exp \left(-\frac{1}{2} \sum_c \frac{1}{\epsilon_c} (\bar{\lambda}_c + \sqrt{q_c} z_c)^2 \right) \right\}^n \tag{4.66}$$

e $\epsilon_c = A_c^2 - q_c$.

A expressão final para a entropia g^* , considerando a aproximação de ponto de sela e levando em conta que as equações para E e F são as mesmas do caso anterior, é:

$$g^* = \frac{1}{4} \sum_c \left\{ \ln \frac{A_c^2}{A_c^2 - q_c} - \frac{q_c}{A_c^2 - q_c} \right\} + \alpha \int Dz_1 Dz_2 \ln \mathcal{W} , \tag{4.67}$$

com a função \mathcal{W} dada por

$$\mathcal{W} = \int_{C_1}^{\infty} D\lambda \left\{ H \left(\sqrt{\frac{\epsilon_1}{\epsilon_2}} (-\lambda + C_1 + C_2) \right) - H \left(\sqrt{\frac{\epsilon_1}{\epsilon_2}} (\lambda - C_1 + C_2) \right) \right\} , \tag{4.68}$$

onde definimos

$$C_1 = \frac{1}{\sqrt{\epsilon_1}} \left(\frac{\kappa_1 + \kappa_2}{\sqrt{2}} + \sqrt{q_1} z_1 \right) \quad (4.69)$$

$$C_2 = \frac{1}{\sqrt{\epsilon_2}} \left(\frac{\kappa_1 - \kappa_2}{\sqrt{2}} + \sqrt{q_2} z_2 \right) \quad (4.70)$$

A obtenção das equações de ponto de sela para os parâmetro q_1 e q_2 , a partir das derivadas de g , envolvem uma grande quantidade de cálculos sem absolutamente nenhuma novidade, de forma que apresentamos somente as expressões finais, deixando para o apêndice 2 os passos intermediários. Estas são:

$$\frac{q_1}{\epsilon_1^2} = \frac{2}{\epsilon_1 + \epsilon_2} \frac{\alpha}{2\pi} \int Dz_1 Dz_2 \left(\frac{e^{-\frac{1}{2} \frac{\epsilon_1}{\epsilon_1 + \epsilon_2} (C_1 + C_2)^2} H(\zeta_-) + e^{-\frac{1}{2} \frac{\epsilon_1}{\epsilon_1 + \epsilon_2} (C_1 - C_2)^2} H(\zeta_+)}{\mathcal{W}} \right)^2 \quad (4.71)$$

$$\frac{q_2}{\epsilon_2^2} = \frac{2}{\epsilon_1 + \epsilon_2} \frac{\alpha}{2\pi} \int Dz_1 Dz_2 \left(\frac{e^{-\frac{1}{2} \frac{\epsilon_1}{\epsilon_1 + \epsilon_2} (C_1 + C_2)^2} H(\zeta_-) - e^{-\frac{1}{2} \frac{\epsilon_1}{\epsilon_1 + \epsilon_2} (C_1 - C_2)^2} H(\zeta_+)}{\mathcal{W}} \right)^2 \quad (4.72)$$

onde

$$\zeta_{\pm} = \frac{\epsilon_2 C_1 \pm \epsilon_1 C_2}{\sqrt{\epsilon_2(\epsilon_1 + \epsilon_2)}} \quad .$$

O primeiro teste é verificar se estas equações se reduzem a (4.44) na situação que armazenamos somente ancestrais. Comparando (4.23) com (4.61) vemos que armazenar somente ancestrais corresponde a tomarmos $k_2 \rightarrow -\infty$. Neste limite observamos que

$$\left\{ \begin{array}{l} (C_1 + C_2) = \frac{1}{\sqrt{\epsilon_1}} (\sqrt{2}\kappa_1 + \sqrt{q_1}z_1 + \sqrt{q_2}z_2) \\ (C_1 - C_2) \rightarrow \infty \\ \zeta_+ \rightarrow \infty \\ \zeta_- \rightarrow -\infty \end{array} \right. , \quad (4.73)$$

e portanto

$$\begin{aligned} \mathcal{W} &\rightarrow \int_{-\infty}^{\infty} D\lambda H\left(-\lambda + \frac{1}{\sqrt{\epsilon_1}}(\sqrt{2}\kappa_1 + \sqrt{q_1}z_1 + \sqrt{q_2}z_2)\right) \\ &= H\left(\frac{1}{\sqrt{\epsilon_1}}(\sqrt{2}\kappa_1 + \sqrt{q_1}z_1 + \sqrt{q_2}z_2)\right) \quad , \end{aligned} \quad (4.74)$$

de forma que as equações para o ponto de sela ficam

$$\frac{q_1}{\epsilon_1^2} = \frac{2}{\epsilon_1 + \epsilon_2} \frac{\alpha}{2\pi} \int Dz_1 Dz_2 \left(\frac{e^{-\frac{1}{2} \frac{1}{\epsilon_1 + \epsilon_2} (\sqrt{2}\kappa_1 + \sqrt{q_1}z_1 + \sqrt{q_2}z_2)^2}}{H\left(\frac{1}{\sqrt{\epsilon_1}}(\sqrt{2}\kappa_1 + \sqrt{q_1}z_1 + \sqrt{q_2}z_2)\right)} \right)^2 \quad (4.75)$$

$$\frac{q_2}{\epsilon_2^2} = \frac{2}{\epsilon_1 + \epsilon_2} \frac{\alpha}{2\pi} \int Dz_1 Dz_2 \left(\frac{e^{-\frac{1}{2} \frac{1}{\epsilon_1 + \epsilon_2} (\sqrt{2}\kappa_1 + \sqrt{q_1}z_1 + \sqrt{q_2}z_2)^2}}{H\left(\frac{1}{\sqrt{\epsilon_1}}(\sqrt{2}\kappa_1 + \sqrt{q_1}z_1 + \sqrt{q_2}z_2)\right)} \right)^2 \quad (4.76)$$

e, se usarmos, de (3.85), que

$$\int Dz_1 Dz_2 F(K + Az_1 + Bz_2) = \int Dz F(K + \sqrt{A^2 + B^2}z) \quad ,$$

finalmente recuperamos as equações do problema anterior.

Como na seção anterior, nos interessa resolver as equações (4.71) no limite em que

$$q_1 \rightarrow A_1^2 \quad , \quad q_2 \rightarrow A_2^2 \quad ,$$

ou seja,

$$\epsilon_1 \rightarrow 0 \quad , \quad \epsilon_2 \rightarrow 0 \quad ,$$

pois dessa forma se obtém o armazenamento crítico [44].

Os detalhes da aplicação deste limite se encontram no apêndice B , os resultados finais são:

$$q_1 = \frac{\epsilon_1^2}{\epsilon_1 + \epsilon_2} \frac{\alpha}{\pi} \int Dz_1 Dz_2 K^+ \quad (4.77)$$

$$q_2 = \frac{\epsilon_2^2}{\epsilon_1 + \epsilon_2} \frac{\alpha}{\pi} \int Dz_1 Dz_2 K^- \quad (4.78)$$

com

$$\begin{aligned}
K^+ &= \frac{2\pi\epsilon_1}{\epsilon_1 + \epsilon_2} \left\{ \Theta(-C_1)\Theta(-|C_1| + C_2)(|C_1| - C_2)^2 \right. \\
&\quad + \Theta(-C_1)\Theta(-|C_1| - C_2)(|C_1| + C_2)^2 \\
&\quad + \Theta(C_1)\Theta\left(-\frac{\epsilon_2}{\epsilon_1} |C_1| + C_2\right)(|C_1| + C_2)^2 \\
&\quad \left. + \Theta(C_1)\Theta\left(-\frac{\epsilon_2}{\epsilon_1} |C_1| - C_2\right)(|C_1| - C_2)^2 \right\} \\
&\quad + \frac{2\pi(\epsilon_1 + \epsilon_2)}{\epsilon_1} \Theta(C_1)\Theta\left(\frac{\epsilon_2}{\epsilon_1} |C_1| + C_2\right)\Theta\left(\frac{\epsilon_2}{\epsilon_1} |C_1| - C_2\right)(C_1)^2 \quad (4.79)
\end{aligned}$$

$$\begin{aligned}
K^- &= \frac{2\pi\epsilon_1}{\epsilon_1 + \epsilon_2} \left\{ \Theta(-C_1)\Theta(-|C_1| + C_2)(|C_1| - C_2)^2 \right. \\
&\quad + \Theta(-C_1)\Theta(-|C_1| - C_2)(|C_1| + C_2)^2 \\
&\quad + \Theta(C_1)\Theta\left(-\frac{\epsilon_2}{\epsilon_1} |C_1| + C_2\right)(|C_1| + C_2)^2 \\
&\quad \left. + \Theta(C_1)\Theta\left(-\frac{\epsilon_2}{\epsilon_1} |C_1| - C_2\right)(|C_1| - C_2)^2 \right\} \\
&\quad + \frac{2\pi\epsilon_1(\epsilon_1 + \epsilon_2)}{\epsilon_2} \Theta(C_1)\Theta\left(\frac{\epsilon_2}{\epsilon_1} |C_1| + C_2\right)\Theta\left(\frac{\epsilon_2}{\epsilon_1} |C_1| - C_2\right)(C_1)^2 \quad (4.80)
\end{aligned}$$

E para $\kappa_1 = \kappa_2 = 0$, que corresponde ao α_c máximo, temos:

$$q_1 = \frac{2\epsilon_1^2}{\epsilon_1 + \epsilon_2} \alpha \left\{ \frac{2}{\epsilon_1 + \epsilon_2} \frac{A_2^2}{\pi} \int_{-A}^{A/E} dt \left(\frac{t - A}{t^2 + 1} \right)^2 \right\}$$

$$+ \frac{\epsilon_1 + \epsilon_2}{\epsilon_1^2} \frac{2a_1^2}{\pi} \int_0^{A/E} dt \left(\frac{1}{t^2 + 1} \right)^2 \} \quad (4.81)$$

$$q_2 = \frac{2\epsilon_2^2}{\epsilon_1 + \epsilon_2} \alpha \left\{ \frac{2}{\epsilon_1 + \epsilon_2} \left(\frac{A_2^2}{\pi} \int_{-A}^{E/A} dt \left(\frac{t-A}{t^2+1} \right)^2 \right) \right. \\ \left. + \frac{\epsilon_1 + \epsilon_2}{\epsilon_1^2} \frac{2A_1^2}{\pi} \int_0^{E/A} dt \left(\frac{t}{t^2+1} \right)^2 \right\} \quad (4.82)$$

Aqui definimos $E = \epsilon_2/\epsilon_1$ (observe que este parâmetro nada tem a ver com o E variável de integração definida anteriormente) e $A = A_2/A_1$.

Combinando estas duas equações sabendo que $q_1 = A_1^2$ e $q_2 = A_2^2$, resulta em um sistema de equações para as variáveis E e α , cujas soluções dependerão do parâmetro A , que dá a informação sobre a modulação do sistema. Este sistema é

$$\frac{\pi}{2\alpha} = \frac{2}{1+E^2} \int_{-A}^{A/E} dt \left(\frac{t-A}{t^2+1} \right)^2 + 2 \int_0^{A/E} dt \left(\frac{1}{t^2+1} \right)^2 \quad (4.83)$$

$$\frac{\pi}{2\alpha} = \frac{2E^2}{A^2(1+E^2)} \int_{-A}^{A/E} dt \left(\frac{t-A}{t^2+1} \right)^2 + 2 \int_0^{A/E} dt \left(\frac{t^2}{t^2+1} \right)^2 \quad (4.84)$$

A solução numérica da função $\alpha_c = \alpha_c(A)$ que sai deste sistema é apresentada na figura 4.1. Observamos que o armazenamento crítico é praticamente insensível a modulação do sistema, ficando sempre em torno de 1 que é o valor correspondente ao caso em que armazenamos αN padrões em uma rede de $N/2$ neurônios.

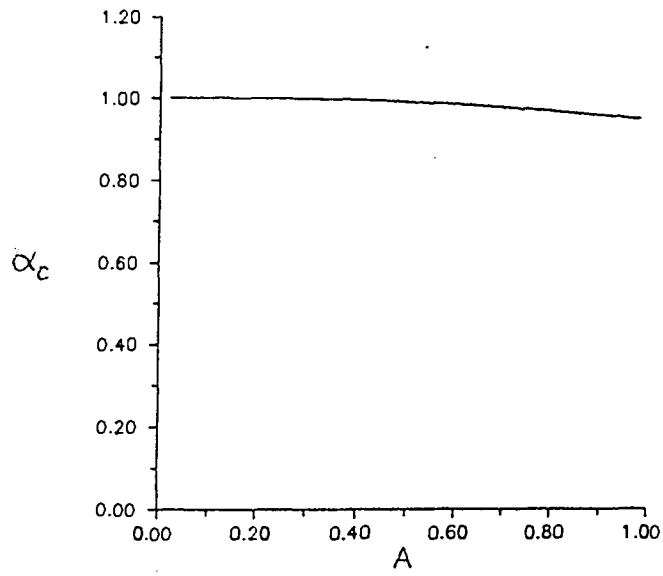


Figura 4.1: Variação de α_c com A_2/A_1 para o caso de armazenamento de ancestrais e descendentes.

Capítulo 5

Conclusões

No decorrer dos três capítulos anteriores, estudamos quais seriam as propriedades de uma rede de Neurônios que possuísse uma modulação espacial na intensidade de suas ligações sinápticas. Esta característica estrutural da rede pode ser descrita da seguinte forma: O sistema é dividido igualmente em ℓ partes ou blocos, e a quantidade

$$\langle \mathcal{J}_{ij}^2 \rangle_{\epsilon}$$

deve variar de acordo com os blocos a que pertencem os neurônios “i” e “j”. A matriz que dá esta variação, ou modulação, foi escolhida como sendo gerada pela equação de recorrência

$$\underline{\underline{A}}(\ell) = \begin{bmatrix} \epsilon \underline{\underline{A}}(\frac{\ell}{2}) + \underline{\underline{U}}(\frac{\ell}{2}) & \underline{\underline{U}}(\frac{\ell}{2}) \\ \underline{\underline{U}}(\frac{\ell}{2}) & \epsilon \underline{\underline{A}}(\frac{\ell}{2}) + \underline{\underline{U}}(\frac{\ell}{2}) \end{bmatrix},$$

o que caracteriza uma relação entre blocos do tipo hierárquica (figura 2.2). O parâmetro ϵ nos auxilia a escrever alguns resultados quantitativos para $\ell > 2$ mas, cabe ressaltar, que tal parametrização não é essencial. Todos os resultados obtidos nos capítulos 2 e 3 valem para qualquer matriz que tenha a forma ultramétrica descrita anteriormente. Por

exemplo, para o caso de $\ell = 8$ esta seria

$$\underline{\underline{A}} = \begin{bmatrix} a_1 & a_2 & & & & & & \\ & a_2 & a_1 & & & & & \\ & & & a_3 & & & & \\ & & & & a_1 & a_2 & & \\ & & a_3 & & & & & a_4 \\ & & & a_2 & a_1 & & & \\ & & & & & a_1 & a_2 & \\ & & & & & & & a_3 \\ & & & & & & & & a_4 \\ & & & & & & & & & a_1 & a_2 \\ & & & & & & & & & & a_3 \\ & & & & & & & & & & & a_1 & a_2 \\ & & & & & & & & & & & & a_2 & a_1 \end{bmatrix} \quad (5.1)$$

com quatro parâmetros independentes $a_1 > a_2 > a_3 > a_4$. No caso genérico de uma matriz $\ell \times \ell$ com $\ell = 2^r$ teríamos $r + 1$ parâmetros independentes.

5.1 Resultados

Este trabalho como um todo pode ser dividido em duas partes principais.

A primeira delas consistiu na investigação das alterações decorrentes da introdução da modulação em uma rede atratora do tipo Hopfield. Com rede do tipo Hopfield quero me referir ao fato de haver simetria nas ligações sinápticas $\mathcal{J}_{ij} = \mathcal{J}_{ji}$ e das memórias $\{\xi^\mu\}$ serem incorporadas ao sistema através da regra de Hebb. Para isso foram utilizadas duas técnicas, campo médio e análise de sinal×ruído, cujos resultados discutiremos a seguir.

5.1.1 Redes Atratoras

1. Armazenamento de um número finito de padrões. $\alpha = 0$

- O cálculo de campo médio mostra que, além dos estados de recuperação e estados espúrios, também presentes no modelo de Hopfield, o sistema é capaz de recuperar uma família de novos estados os quais chamamos de descendentes. A característica principal deles é justamente a sua descendência, ou seja, o fato de serem construídos a partir do conjunto original de p padrões $\{\xi^\mu\}$ [49] [50].

- Os estados descendentes são tais que os p vetores \bar{m}^μ , que tem como componentes as magnetizações do sistema em cada bloco, são autovetores ou combinações lineares de autovetores de \underline{A} . Estas soluções podem descender de um ou mais padrões $\{\xi^\mu\}$. Quando existe apenas um ancestral o número delas foi calculado para $\ell \gg 1$ resultando [50]

$$n \sim e^{0.45\ell} .$$

- Os estados descendentes são sempre metaestáveis e suas propriedades básicas, como energia livre e temperatura crítica, dependem exclusivamente do autovalor ou autovalores de \underline{A} a que correspondem. Para o caso de soluções que são autovetores de \underline{A} observa-se que $f_a < f_b$ e $T_a^* > T_b^*$ se $\lambda_a > \lambda_b$, onde f é energia livre, T^* temperatura crítica e λ autovalor de \underline{A} [50].

- Os cálculos de estabilidade para estas soluções mostram que todas elas desaparecem descontinuamente ao aumentarmos a temperatura [50].

2. Armazenamento de um número extensivo de padrões. $\alpha \neq 0$

Nesta parte, investigou-se as propriedades de recuperação de somente dois tipos de estados: ancestrais e descendentes que fossem autovetores de \underline{A} . Os principais objetivos eram o cálculo do diagrama de fases $\alpha \times T$ e a determinação do armazenamento crítico para cada uma das soluções.

No que se refere ao cálculo de α_c também usamos, alternativamente à teoria de campo médio, uma análise de sinal×ruído. Os resultados absolutos desta quantidade nos dois métodos não são comparáveis, pois diferem na sua definição. Afortunadamente, o mesmo não ocorre para as razões

$$\left[\frac{\alpha_c}{\alpha_c^{Hopfield}} \right]$$

- Os resultados mostram que, nos dois casos, o armazenamento crítico depende explicitamente da matriz \underline{A} através de seus autovalores. Além disso, vemos que a introdução da modulação sempre representa uma diminuição da capacidade de armazenamento dos

ancestrais $\{\xi^\mu\}$, que é tanto mais intensa quanto maior for o valor correspondente a esta mesma capacidade dos descendentes [51].

5.1.2 Investigação do Espaço de Modelos

Na segunda parte do trabalho estudamos, através do método de Gardner, o espaço de todos os modelos que possuíssem modulação através da condição que suas ligações sinápticas satisfizessem:

$$\sum_j^{(b)} \mathcal{J}_{ij}^2 = A_{ij}^2 \frac{N}{\ell} \quad . \quad (5.2)$$

O objeto de investigação agora é encontrar a capacidade de armazenagem que teria o melhor modelo, ou seja, aquele que memorizasse o maior número de padrões. Os resultados foram:

- Para o caso de armazenamento de ancestrais, existe uma diminuição deste valor nos mesmos moldes do que ocorreu para uma rede atratora. Além disso podemos comparar diretamente este resultado com aquele obtido na análise de sinal×ruído no problema anterior. Mais do que isso, obtemos a seguinte igualdade:

$$\left[\frac{\alpha_c}{\alpha_c^{Hopfield}} \right]_{\text{signal-ruído}} = \left[\frac{\alpha_c}{\alpha_c^{Gardner}} \right]_{\text{Gardner}} \quad .$$

Esse resultado é muito importante, pois mostra que a condição escolhida para o Método de Gardner (5.2) é adequada para representar a modulação das intensidades sinápticas que usamos na primeira parte do trabalho [52].

- No caso do armazenamento conjunto de ancestrais e descendentes observou-se que a modulação não parece facilitar esta tarefa. Ao contrário do que era de se esperar, dados os resultados anteriores, o armazenamento crítico máximo α_c é, ao menos no caso de $\ell = 2$, praticamente insensível a modulação [52].

5.2 Comentários Finais

A investigação de sistemas de redes de neurônios com estrutura espacial, como já foi dito, é importante tanto do ponto de vista puramente científico, quando o interesse é o entendimento dos cérebros biológicos, como do tecnológico para o qual o objetivo é o projeto de dispositivos. Contudo, existem poucos modelos que exploram esta característica e isto se deve principalmente ao fato que, nestes casos, o campo médio deixa de dar a solução exata do problema.

No modelo que propusemos contornamos essa dificuldade considerando uma estrutura espacial macroscópica. Observe que, mesmo que alguns dos elementos de matriz A_{ab} fossem iguais a zero, significando a não existência de ligações entre certos grupos de neurônios, ainda assim cada neurônio se ligaria com um número extensivo de outros neurônios, o que permitiria a utilização de campo médio como solução exata.

As principais consequências da introdução da modulação foram: O aparecimento de novos estados atratores diferentes daqueles originalmente armazenados e a diminuição da capacidade de armazenamento destes últimos.

Quanto ao primeiro aspecto, devo dizer que os estados descendentes, apesar de não terem sido explicitamente armazenados na regra de aprendizado, não podem ser considerados como tendo o mesmo status dos estados espúrios. Isso porque estes estados são conhecidos microscopicamente, enquanto que os outros não. Observe que, se dizemos que uma rede está no estado expúrio que mistura as memórias 1 e 2, representado como $\vec{m} = m(1, 1, 0, \dots, 0)$, a única informação microscópica que dispomos é que existem $3/4N$ sítios que concordam com $\{\xi^1\}$ e outros $3/4N$ que concordam com $\{\xi^2\}$, mas não sabemos quais são eles.

Já a diminuição da capacidade de armazenamento, infelizmente, parece estar fortemente relacionada com o fato de haver modulação na rede, não importa de que tipo. Observe que todos os resultados que obtivemos nas três técnicas apontam para isso e, pelo menos dois deles, valem para uma matriz \underline{A} mais geral que (5.1).

Encontramos muitas semelhanças formais entre o nosso modelo e modelos para ar-

mazenamento de memórias com diferentes intensidades [41] e modelos de memórias de trabalho [47]. Isso porque os descendentes no nosso modelo são, na verdade, padrões mais fracamente memorizados. Para aqueles modelos, este tipo de memória mais fraca também surge, a $\alpha = 0$, através de uma transição de primeira ordem a medida que baixamos a temperatura [41]. Além disso sua capacidade de armazenamento também prova ser sempre menor que aquela do Modelo de Hopfield [47].

O último comentário a fazer é sobre a relação existente entre nossos ancestrais e descendentes. Como escolhemos trabalhar com os padrões ancestrais ortogonais, resulta que descendentes de diferentes ancestrais também são ortogonais, isto porque suas superposições são zero em cada bloco. Ocorre, porém, que muitos descendentes de um mesmo ancestral também são ortogonais. A razão disso é que alguns dos autovetores de \underline{A} são ortogonais entre si. Isto mostra que o significado de ancestral e descendente aqui é fundamentalmente diferente daquele usado nos modelos de armazenagem de padrões hierárquicos [18] [24] discutidos na introdução deste trabalho.

5.2.1 Possíveis Extensões

- Seguindo a linha do que foi realizado nos capítulos 1 e 2, podemos ainda investigar o comportamento de uma rede definida pela função de energia (2.10) quando os padrões ancestrais não são mais ortogonais, ou seja, quando possuem atividade.

- Uma outra possibilidade é explorar a forma da matriz \underline{A} , buscando outros arranjos que não o hierárquico. Na maioria dos cálculos que desenvolvemos, a única exigência sobre a forma desta matriz era que seus autovalores fossem positivos, de modo que ainda temos muita liberdade na sua escolha.

- A hipótese mais tentadora, no entanto, é fazer uma análise dinâmica de um modelo com modulação. A vantagem deste tipo de abordagem é que não está atrelada a condições como simetria da matriz sináptica, ou então, a rede ser completamente conectada, exigidas pelo tratamento da mecânica estatística de equilíbrio; além disso permite a obtenção das bacias de atração para os estados estacionários [46], que nesse caso seriam padrões

ancestrais e descendentes.

- Um modo possível de introduzir assimetria é, mantendo a regra de Hebb, fazer cortes em algumas sinapses. Além de causar assimetria, estes cortes provocam a diminuição da conectividade da rede (diluição) [27]. O modo de introduzir o conceito de modulação em um modelo diluído do tipo discutido no final da introdução deste trabalho seria trocar a distribuição (1.26) por

$$P(a_{ij}) = A_{ab}\delta(a_{ij} - 1) + (1 - A_{ab})\delta(a_{ij}) \quad , \quad (5.3)$$

com $i \in a$, $j \in b$ e a, b índices de blocos. Desta forma, podemos diluir diferentemente as ligações dependendo dos blocos onde estão os neurônios. Isto nos permitirá estudar sistemas onde grupos de neurônios altamente conectados são ligados entre si de forma esparsa. Isto é importante pois é o tipo de geometria que predomina nos sistemas nervosos biológicos [30] [32].

Apêndice A

Estabilidade de Soluções

O sistema de equações integrais, não lineares, formado pelas equações (3.41), (3.42) e (3.43) possui um certo número de soluções que representam os possíveis estados do sistema. Isto não garante, no entanto que estes estados sejam estáveis [48]. Como estamos interessados apenas em estados que possuam uma bacia de atração finita (que são os estados estáveis e metaestáveis) deveremos ajuntar a este sistema condições que determinem a estabilidade de soluções.

Estas condições são aquelas que garantem que todos os autovalores da matriz formada pelas derivadas segundas da energia livre em relação a todos os parâmetros de ordem sejam positivos .

A hipótese de simetrias de réplicas já é uma escolha de um tipo de solução entre as várias que o problema pode ter. Para o cálculo da matriz de estabilidade devemos pegar a função de energia livre na sua forma mais geral, no caso a equação (3.39) (antes de tomarmos o limite das réplicas), e derivá-la em relação a $m_a^{\mu,\rho}$, $q_a^{\rho\rho'}$ e $r_a^{\rho\rho'}$.

Relembrando essa equação : =====

$$\mathcal{F}_n(m, q, r) = \frac{1}{2} \sum_{\rho=1}^n \sum_{a,b} \frac{A_{ab}}{\ell^2} \vec{m}_a^\rho \vec{m}_a^\rho + \frac{1}{2} \frac{\alpha\beta}{\ell} \sum_a \sum_{\rho,\rho'} r_a^{\rho\rho'} q_a^{\rho\rho'}$$

$$- \frac{\alpha}{2\beta} \sum_{r'=1}^{\infty} \frac{1}{r'} \left(\frac{\beta}{\ell} \right)^{r'} \sum_{a_1 \dots a_{r'}} A_{a_1 a_2} \dots A_{a_{r'} a_1}$$

$$\begin{aligned} & \times \sum_{\rho_1 \dots \rho_{r'}} (\delta^{\rho_1 \rho_2} + q_{a_1}^{\rho_1 \rho_2}) \dots (\delta^{\rho_{r'} \rho_1} + q_{a_{r'}}^{\rho_{r'} \rho_1}) \\ & \times \frac{1}{\beta \ell} \sum_a \langle \log(\text{Tr}_{\{S_\rho\}} \exp(\beta \sum_b \frac{A_{ab}}{\ell} \bar{\xi} \bar{m}_b^\rho S^\rho + \frac{\alpha \beta^2}{2} \sum_{\rho \rho'} r_a^{\rho \rho'} S^\rho S^{\rho'})) \rangle_\xi \quad (\text{A.1}) \end{aligned}$$

Assim ,

$$\begin{aligned} \frac{\partial \mathcal{F}_n}{\partial m_c^{\mu, \gamma}} &= \sum_{b=1}^{\ell} \frac{A_{cb}}{\ell^2} m_b^{\mu, \gamma} - \\ & \sum_a \frac{A_{ac}}{\ell} \langle \xi^\mu \frac{(\text{Tr}_{\{S_\rho\}} S^\gamma \exp(\beta \sum_b \frac{A_{ab}}{\ell} \bar{\xi} \bar{m}_b^\rho S^\rho + \frac{\alpha \beta^2}{2} \sum_{\rho \rho'} r_a^{\rho \rho'} S^\rho S^{\rho'}))}{\text{Tr}_{\{S_\rho\}} e^{\dots}} \rangle_\xi \\ &= \sum_{b=1}^{\ell} \frac{A_{cb}}{\ell^2} m_b^{\mu, \gamma} - \sum_a^{\text{ell}} \frac{A_{ac}}{\ell} \langle \xi^\mu \langle S^\gamma \rangle_a \rangle_\xi \quad (\text{A.2}) \end{aligned}$$

onde $\langle \dots \rangle_a$ representa o quociente na primeira expressão de A.2.

$$\begin{aligned} \frac{\partial^2 \mathcal{F}_n}{\partial m_c^{\mu, \gamma} \partial m_d^{\nu, \delta}} &= \frac{A_{cd}}{\ell^2} \delta^{\mu\nu} \delta^{\gamma\delta} \\ & \frac{\beta}{\ell^3} \sum_a \frac{\ell}{a} A_{ac} A_{ad} \langle \xi^\mu \xi^\nu (\langle S^\gamma S^\delta \rangle_a - \langle S^\gamma \rangle_a \langle S^\delta \rangle_a) \rangle_\xi \quad (\text{A.3}) \end{aligned}$$

$$\frac{\partial^2 \mathcal{F}_n}{\partial m_c^{\mu, \gamma} \partial r_d^{\rho \rho'}} = -\alpha \frac{\beta^2}{2} A_{cd} \langle \xi^\mu (\langle S^\gamma S^\rho S^{\rho'} \rangle_a - \langle S^\gamma \rangle_a \langle S^\rho S^{\rho'} \rangle_a) \rangle_\xi \quad (\text{A.4})$$

para $\rho \neq \rho'$. E

$$\frac{\partial^2 \mathcal{F}_n}{\partial m_c^{\mu, \gamma} \partial q_c^{\rho \rho'}} = 0 \quad (\text{A.5})$$

também se $\rho \neq \rho'$. A primeira derivada em relação a $r_c^{\rho \rho'}$ é:

$$\frac{\partial \mathcal{F}_n}{\partial r_c^{\rho\rho'}} = \frac{\alpha\beta}{\ell} (q_c^{\rho\rho'} - \langle\langle S^\rho S^{\rho'} \rangle\rangle_c) \quad (\text{A.6})$$

e as subsequentes derivadas segundas

$$\frac{\partial^2 \mathcal{F}_n}{\partial r_c^{\rho\rho'} \partial r_d^{\gamma\delta}} = -\frac{\alpha^2 \beta^3}{\ell} \delta^{cd} \langle\langle S^\rho S^{\rho'} S^\gamma S^\delta \rangle\rangle_c - \langle S^\rho S^{\rho'} \rangle_c \langle S^\gamma S^\delta \rangle_c \quad (\text{A.7})$$

com $\rho \neq \rho'$ e $\gamma \neq \delta$.

$$\frac{\partial^2 \mathcal{F}_n}{\partial r_c^{\rho\rho'} \partial q_d^{\gamma\delta}} = \frac{\alpha\beta}{\ell} \delta^{cd} (\delta^{\rho\gamma} \delta^{\rho'\delta} + \delta^{\rho\delta} \delta^{\rho'\gamma}) \quad (\text{A.8})$$

com $\rho \neq \rho'$ e $\gamma \neq \delta$.

Finalmente temos

$$\begin{aligned} \frac{\partial \mathcal{F}_n}{\partial q_c^{\rho\rho'}} &= \frac{\alpha\beta}{\ell} r_c^{\rho\rho'} - \frac{\alpha}{2\beta} \frac{\partial}{\partial q_c^{\rho\rho'}} \times \\ &\sum_{r'=1}^{\infty} \frac{1}{r'} \left(\frac{\beta}{\ell}\right)^{r'} \sum_{a_1 \dots a_{r'}} A_{a_1 a_2} \dots A_{a_{r'} a_1} \times \\ &\left\{ \sum_{\rho_1 \dots \rho_{r'}} (\delta^{\rho_1 \rho_2} + q_{a_1}^{\rho_1 \rho_2}) \dots (\delta^{\rho_{r'} \rho_1} + q_{a_{r'}}^{\rho_{r'} \rho_1}) \right\} \end{aligned} \quad (\text{A.9})$$

$$\begin{aligned} \frac{\partial^2 \mathcal{F}_n}{\partial q_c^{\rho\rho'} \partial q_d^{\gamma\delta}} &= -\frac{\alpha}{2\beta} \frac{\partial^2}{\partial q_c^{\rho\rho'} \partial q_d^{\gamma\delta}} \times \\ &\sum_{r'=1}^{\infty} \frac{1}{r'} \left(\frac{\beta}{\ell}\right)^{r'} \sum_{a_1 \dots a_{r'}} A_{a_1 a_2} \dots A_{a_{r'} a_1} \times \\ &\left\{ \sum_{\rho_1 \dots \rho_{r'}} (\delta^{\rho_1 \rho_2} + q_{a_1}^{\rho_1 \rho_2}) \dots (\delta^{\rho_{r'} \rho_1} + q_{a_{r'}}^{\rho_{r'} \rho_1}) \right\} \end{aligned} \quad (\text{A.10})$$

A.1 Cálculo específico de $\frac{\partial^2 \mathcal{F}_n}{\partial q_c^{\rho\rho'} \partial q_d^{\gamma\delta}}$

$$\frac{\partial \mathcal{F}_n}{\partial q_c^{\rho\rho'}} = \frac{\partial}{\partial q_c^{\rho\rho'}} \sum_{r'=1}^{\infty} \frac{1}{r'} \left(\frac{\beta}{\ell}\right)^{r'} \sum_{a_1 \dots a_{r'}} A_{a_1 a_2} \dots A_{a_{r'} a_1} \times$$

$$\left\{ \sum_{\rho_1 \dots \rho_{r'}} (\delta^{\rho_1 \rho_2} + q_{a_1}^{\rho_1 \rho_2}) \dots (\delta^{\rho_{r'} \rho_1} + q_{a_{r'}}^{\rho_{r'} \rho_1}) \right\} \quad (\text{A.11})$$

$$= \frac{1}{2} \left(\frac{\beta}{\ell}\right)^2 \left\{ 2q_{a_1}^{\rho\rho'} \sum_{a_2} A_{a_1 c} A_{c a_1} + 2q_{a_2}^{\rho\rho'} \sum_{a_2} A_{c a_2} A_{a_2 c} \right\} +$$

$$\frac{1}{3} \left(\frac{\beta}{\ell}\right)^3 \left\{ \sum^{a_2, a_3} A_{c a_2} A_{a_2 a_3} A_{a_3 c} \times \right.$$

$$\left. \sum_{\rho_3} \left((\delta^{\rho_3 \rho_3} + q_{a_2}^{\rho_3 \rho_3}) (\delta^{\rho_3 \rho'} + q_{a_2}^{\rho_3 \rho'}) + (\delta^{\rho' \rho_3} + q_{a_2}^{\rho' \rho_3}) (\delta^{\rho_3 \rho} + q_{a_2}^{\rho_3 \rho}) \right) + \dots \right\}$$

$$+ \dots \quad (\text{A.12})$$

Observa-se que na derivada do r' -ésimo termo de \mathcal{F}_n aparecerá um fator “ r' ” devido ao fato de neste termo existirem r' somas em blocos e além disso sempre teremos um fator 2 resultante da simetria $q_a^{\rho\rho'} = q_a^{\rho'\rho}$. Assim obtemos :

$$\frac{\partial \mathcal{F}_n}{\partial q_c^{\rho\rho'}} = \sum_{r'=2}^{\infty} \left(\frac{\beta}{\ell}\right)^{r'} \sum_{a_2 \dots a_{r'}} A_{c a_2} \dots A_{a_{r'} c} \times$$

$$\left\{ \sum_{\rho_1 \dots \rho_{r'}} (\delta^{\rho_1 \rho_2} + q_{a_2}^{\rho_1 \rho_2}) \dots (\delta^{\rho_{r'} \rho_1} + q_{a_{r'}}^{\rho_{r'} \rho_1}) \right\} \quad (\text{A.13})$$

Prosseguindo temos:

$$\frac{\partial^2 \mathcal{F}_n}{\partial q_c^{\rho\rho'} \partial q_d^{\gamma\delta}} = 2 \left\{ \left(\frac{\beta}{\ell}\right) A_{c d} \delta^{\rho\delta} + \left(\frac{\beta}{\ell}\right)^2 \sum_{a_1} A_{c a_1} A_{a_1 d} (\delta^{\rho\gamma} + q_{a_1}^{\rho\gamma}) + \right.$$

$$\begin{aligned}
 & + \left(\frac{\beta}{\ell} \right)^3 \sum_{a_1 a_2} A_{ca_1} A_{a_1 a_2} A_{a_2 d} \sum_{\rho_1} (\delta^{\rho \rho_1} + q_{a_1}^{\rho \rho_1}) (\delta^{\rho_1 \gamma} + q_{a_1}^{\rho_1 \gamma}) + \dots \Big) \times \\
 & \left(\frac{\beta}{\ell} A_{cd} \delta^{\delta \rho'} + \left(\frac{\beta}{\ell} \right)^2 \sum_{a_1} A_{ca_1} A_{a_1 d} (\delta^{\delta \rho'} + q_{a_1}^{\delta \rho'}) + \right. \\
 & \left. + \left(\frac{\beta}{\ell} \right)^3 \sum_{a_1 a_2} A_{ca_1} A_{a_1 a_2} A_{a_2 d} \sum_{\rho_1} (\delta^{\delta \rho_1} + q_{a_1}^{\delta \rho_1}) (\delta^{\rho_1 \rho'} + q_{a_1}^{\rho_1 \rho'}) + \dots \right) \\
 & + (\rho \delta)(\gamma \delta) \} \tag{A.14}
 \end{aligned}$$

que pode ser escrito

$$\frac{\partial^2 \mathcal{F}_0}{\partial q_c^{\rho \rho'} \partial q_d^{\gamma \delta}} = 2 \left(U_{cd}^{\rho \gamma} U_{cd}^{\delta \rho'} + U_{cd}^{\rho \delta} U_{cd}^{\delta \gamma} \right) \tag{A.15}$$

onde

$$\begin{aligned}
 U_{cd}^{\rho \gamma} = & \left(\frac{\beta}{\ell} A_{cd} \delta^{\rho \delta} + \left(\frac{\beta}{\ell} \right)^2 \sum_{a_1} A_{ca_1} A_{a_1 d} (\delta^{\rho \gamma} + q_{a_1}^{\rho \gamma}) + \right. \\
 & \left. + \left(\frac{\beta}{\ell} \right)^3 \sum_{a_1 a_2} A_{ca_1} A_{a_1 a_2} A_{a_2 d} \sum_{\rho_1} (\delta^{\rho \rho_1} + q_{a_1}^{\rho \rho_1}) (\delta^{\rho_1 \gamma} + q_{a_1}^{\rho_1 \gamma}) + \dots \right) \tag{A.16}
 \end{aligned}$$

De modo que podemos escrever

$$\frac{\partial^2 \mathcal{F}_n}{\partial q_c^{\rho \rho'} \partial q_d^{\gamma \delta}} = -\frac{\alpha}{\beta} \left(U_{cd}^{\rho \gamma} U_{cd}^{\delta \rho'} + U_{cd}^{\rho \delta} U_{cd}^{\delta \gamma} \right) \tag{A.17}$$

com $\rho \neq \rho'$ e $\delta \neq \gamma$.

O cálculo dos autovalores da matriz de elementos (A.3), (A.4), (A.5), (A.7), (A.8) e (A.17) é muito trabalhoso. Somente será possível fazê-lo para alguns tipos de soluções.

A.2 Solução com simetria de réplicas

Vamos começar pelo cálculo das médias do tipo $\langle S^{\rho_1} \dots S^{\rho_k} \rangle_a$ no caso em que vale (3.36).

Considere

$$\langle S^\gamma \rangle_a = \frac{\text{Tr}_{\{S^\rho\}} S^\gamma \exp \left(\beta \sum_b \frac{A_{ab}}{\ell} \vec{\xi} \cdot \sum_\rho \vec{m}_b^\rho S^\rho + \frac{\alpha\beta^2}{2} \sum_{\rho, \rho'} r_a^{\rho\rho'} S^\rho S^{\rho'} \right)}{\text{Tr}_{\{S^\rho\}} \exp \left(\beta \sum_b \frac{A_{ab}}{\ell} \vec{\xi} \cdot \sum_\rho \vec{m}_b^\rho S^\rho + \frac{\alpha\beta^2}{2} \sum_{\rho, \rho'} r_a^{\rho\rho'} S^\rho S^{\rho'} \right)} \quad (\text{A.18})$$

se introduzirmos um campo podemos escrever

$$\langle S^\gamma \rangle_a = \left[\frac{\partial}{\partial h^\gamma} \left(\ln \text{Tr}_{\{S^\rho\}} \exp \left(\sum_\rho \left(\beta \sum_b \frac{A_{ab}}{\ell} \vec{\xi} \cdot \vec{m}_b^\rho + h^\rho \right) S^\rho + \frac{\alpha\beta^2}{2} \sum_{\rho, \rho'} r_a^{\rho\rho'} S^\rho S^{\rho'} \right) \right) \right]_{h^\rho=0} \quad (\text{A.19})$$

usando a simetria de réplicas é possível linearizar o expoente e aplicar o traço.

$$\langle S^\gamma \rangle_a = \left[\frac{\partial}{\partial h^\gamma} \left(\ln \int Dz \prod_\rho \cosh \left(\beta \left(\sum_b \frac{A_{ab}}{\ell} \vec{\xi} \cdot \vec{m}_b^\rho + \sqrt{\rho r_a} z \right) + h^\rho \right) \right) \right]_{h^\rho=0} \quad (\text{A.20})$$

fazendo a derivada e tomando $h^\rho = 0$ resulta:

$$\langle S^\gamma \rangle_a = \frac{\int Dz \sinh \left(\beta \left(\sum_b \frac{A_{ab}}{\ell} \vec{\xi} \cdot \vec{m}_b^\rho + \sqrt{\alpha r_a} z \right) \right) \cosh^{n-1} \left(\beta \left(\sum_b \frac{A_{ab}}{\ell} \vec{\xi} \cdot \vec{m}_b^\rho + \sqrt{\alpha r_a} z \right) \right)}{\int Dz \cosh^n \left(\beta \left(\sum_b \frac{A_{ab}}{\ell} \vec{\xi} \cdot \vec{m}_b^\rho + \sqrt{\alpha r_a} z \right) \right)} \quad (\text{A.21})$$

Agora convém lembrar que ao final de tudo devemos fazer o limite do número de réplicas indo a zero para obtermos a energia livre e suas derivadas. Prevendo isto vou considerar (A.21) em ordem zero nesta expansão, ou seja

$$\langle S^\gamma \rangle_a = \int Dz \tanh \left(\beta \left(\sum_b \frac{A_{ab}}{\ell} \vec{\xi} \cdot \vec{m}_b^p + \sqrt{\alpha r_a} z \right) \right) + o(n) \quad (\text{A.22})$$

Para a correlação entre vários spins é imediato verificar que

$$\langle S^{\gamma_1} \dots S^{\gamma_k} \rangle_a \simeq \int Dz \tanh^k \left(\beta \left(\sum_b \frac{A_{ab}}{\ell} \vec{\xi} \cdot \vec{m}_b^p + \sqrt{\alpha r_a} z \right) \right) \quad (\text{A.23})$$

para $\gamma_1 \neq \gamma_2$, $\gamma_1 \neq \gamma_3$, ...

Vamos estabelecer daqui para frente a seguinte notação

$$\langle S^{\gamma_1} \dots S^{\gamma_k} \rangle_a = \langle (M_a)^k \rangle_z \quad (\text{A.24})$$

onde

$$M_a = \tanh \left(\beta \left(\sum_b \frac{A_{ab}}{\ell} \vec{\xi} \cdot \vec{m}_b^p + \sqrt{\alpha r_a} z \right) \right) \quad (\text{A.25})$$

e

$$\langle \dots \rangle_z = \int Dz \dots$$

Feito isto, podemos representar (A.3),(A.4) e (A.7) em termos destas novas médias.

$$\langle S^\gamma \rangle_a = \langle M_a \rangle_z \quad (\text{A.26})$$

$$\langle S^\gamma S^\delta \rangle_a = \delta^{\gamma\delta} + (1 - \delta^{\gamma\delta}) \langle M_a^2 \rangle_z \quad (\text{A.27})$$

$$\begin{aligned} \langle S^\gamma S^\rho S^{\rho'} \rangle_a &= \delta^{\gamma\rho\rho'} + (\delta^{\gamma\rho}(1 - \delta^{\gamma\rho'}) + 2\text{perm.}) \langle M_a^2 \rangle_z \\ &+ ((1 - \delta^{\gamma\rho})(1 - \delta^{\rho\rho'})(1 - \delta^{\gamma\rho'})) \langle M_a^3 \rangle_z \end{aligned} \quad (\text{A.28})$$

$$\begin{aligned} \langle S^\gamma S^\delta S^\rho S^{\rho'} \rangle_a &= \delta^{\gamma\delta\rho\rho'} + (\delta^{\gamma\delta}\delta^{\rho\rho'}(1 - \delta^{\gamma\rho'}) + 2\text{perm.}) \\ &+ (\delta^{\gamma\delta}(1 - \delta^{\rho\rho'})(1 - \delta^{\gamma\rho})(1 - \delta^{\gamma\rho'}) + 5\text{perm.}) \langle M_a^2 \rangle_z \\ &+ (\delta^{\gamma\delta\rho'}(1 - \delta^{\rho\rho'}) + 3\text{perm.}) \langle M_a \rangle_z \\ &+ ((1 - \delta^{\gamma\delta})(1 - \delta^{\gamma\rho})(1 - \delta^{\gamma\rho'})(1 - \delta^{\delta\rho})(1 - \delta^{\delta\rho'})(1 - \delta^{\rho\rho'})) \langle M_a^4 \rangle_z \end{aligned} \quad (\text{A.29})$$

substituindo em (A.3)

$$\begin{aligned} \frac{\partial^2 \mathcal{F}_n}{\partial m_c^{\mu,\gamma} \partial m_d^{\nu,\delta}} &= \left(\frac{A_{cd}}{\ell^2} - \beta \frac{(A^2)_{cd}}{\ell^3} \right) \delta^{\mu\nu} \delta^{\gamma\delta} \\ &- \frac{\beta}{\ell^3} \sum_a^\ell A_{ac} A_{ad} \langle \xi^\mu \xi^\nu (\langle M_a^2 \rangle_z \delta^{\gamma\delta} - \langle M_a^2 \rangle_z + \langle M_a \rangle_z^2) \rangle_\xi \end{aligned} \quad (\text{A.30})$$

de (A.4)

$$\begin{aligned} \frac{\partial^2 \mathcal{F}_n}{\partial m_c^{\mu,\gamma} \partial r_d^{\rho\rho'}} &= -\alpha \frac{\beta^2}{2} A_{cd} \{ (\delta^{\gamma\rho} + \delta^{\gamma\rho'}) \langle \xi^\mu \langle M_d(1 - M_d^2) \rangle_z \rangle_\xi \\ &+ \langle \xi^\mu (\langle M_d^3 \rangle_z - \langle M_d \rangle_z \langle M_d^2 \rangle_z) \rangle_\xi \} \end{aligned} \quad (\text{A.31})$$

e de (A.7)

$$\begin{aligned}
\frac{\partial^2 \mathcal{F}_n}{\partial r_c^{\rho\rho'} \partial r_d^{\gamma\delta}} &= -\frac{\alpha^2 \beta^3}{\ell} \delta^{cd} \left(\langle M_d^4 \rangle_z - \langle M_d^2 \rangle_z^2 \right) \\
&\quad + (\delta^{\rho\gamma} \delta^{\rho'\delta} + \delta^{\rho\delta} \delta^{\rho'\gamma}) (\langle M_c^4 \rangle_z - 2 \langle M_c^2 \rangle_z + 1) \\
&\quad + (\delta^{\rho\gamma} + \delta^{\rho\delta} + \delta^{\rho'\gamma} \delta^{\rho'\delta}) (\langle M_c^2 \rangle_z - \langle M_c^4 \rangle_z) >_\xi
\end{aligned} \tag{A.32}$$

Quanto as modificações em (A.17) devemos lembrar que com simetria de réplicas

$$q_a^{\rho\rho'} = (1 - \delta^{\rho\rho'}) q_a \tag{A.33}$$

assim podemos escrever

$$U_{cd}^{\rho\gamma} = \mathcal{R}_{cd} + \delta^{\rho\gamma} \mathcal{R}'_{cd} \tag{A.34}$$

onde

$$\mathcal{R}_{cd} = \sum_{k=2}^{\infty} \left(\frac{\beta}{\ell} \right)^k \sum_{a_2, \dots, a_k} A_{ca_2} \cdots A_{a_k d} \sum_{k'=2}^k \frac{q_{a_{k'}} \prod_{j=2}^k (1 - q_{a_j})}{1 - q_{a_{k'}}} \tag{A.35}$$

e

$$\mathcal{R}'_{cd} = \frac{\beta}{\ell} + \sum_{k=2}^{\infty} \left(\frac{\beta}{\ell} \right)^k \sum_{a_2, \dots, a_k} A_{ca_2} \cdots A_{a_k d} \prod_{j=2}^k (1 - q_{a_j}) \tag{A.36}$$

e portanto

$$\begin{aligned} \frac{\partial^2 \mathcal{F}_n}{\partial q_c^{\rho\rho'} \partial q_d^{\gamma\delta}} &= -\frac{\alpha}{\beta} \left(2(\mathcal{R}_c d)^2 + (\delta^{\rho\gamma} + \delta^{\rho\delta} + \delta^{\rho'\gamma} \delta^{\rho\delta}) \mathcal{R}_{cd} \mathcal{R}'_{cd} \right) \\ &\quad - \frac{\alpha}{\beta} (\delta^{\rho\gamma} \delta^{\rho'\delta} + \delta^{\rho\delta} \delta^{\rho'\gamma}) (\mathcal{R}'_c d)^2 \end{aligned} \quad (\text{A.37})$$

Quando buscamos calcular os autovalores da matriz formada pelos elementos (A.4), (A.5), (A.7), (A.8) e (A.17) o que fazemos é diagonalizar o termo da segunda ordem da expansão de $\mathcal{F}_n(m, q, r)$ em torno do ponto de sela. Fazer a diagonalização explicitamente completando quadrados é bem mais simples que montar a matriz e calcular os autovalores a partir da equação secular e, é isto que faremos daqui para frente.

A expressão geral para \mathcal{F}_n em segunda ordem é :

$$\begin{aligned} \mathcal{F}_n(m, q, r) &= \mathcal{F}_n(m, q, r) + \left[\sum_{a,b} \sum_{\nu,\mu} \sum_{\rho,\rho'} \frac{\partial^2 \mathcal{F}_n}{\partial m_a^{\mu,\rho} \partial m_b^{\nu,\rho'}} \right]_{m,q,r} \delta m_a^{\mu,\rho} \delta m_b^{\nu,\rho'} \\ &+ \left[2 \sum_{a,b} \sum_{\mu} \sum_{\rho,\rho' < \rho''} \frac{\partial^2 \mathcal{F}_n}{\partial m_a^{\mu,\rho} \partial r_b^{\rho',\rho''}} \right]_{m,q,r} \delta m_a^{\mu,\rho} \delta r_b^{\rho',\rho''} \\ &+ \left[2 \sum_{a,b} \sum_{\mu} \sum_{\rho,\rho' < \rho''} \frac{\partial^2 \mathcal{F}_n}{\partial m_a^{\mu,\rho} \partial q_b^{\rho',\rho''}} \right]_{m,q,r} \delta m_a^{\mu,\rho} \delta q_b^{\rho',\rho''} \\ &+ \left[\sum_{a,b} \sum_{\rho < \rho'} \sum_{\sigma < \sigma'} \frac{\partial^2 \mathcal{F}_n}{\partial r_a^{\rho,\rho'} \partial r_b^{\sigma,\sigma'}} \right]_{m,q,r} \delta r_a^{\rho,\rho'} \delta r_b^{\sigma,\sigma'} \\ &+ \left[\sum_{a,b} \sum_{\rho < \rho'} \sum_{\sigma < \sigma'} \frac{\partial^2 \mathcal{F}_n}{\partial q_a^{\rho,\rho'} \partial q_b^{\sigma,\sigma'}} \right]_{m,q,r} \delta q_a^{\rho,\rho'} \delta q_b^{\sigma,\sigma'} \\ &+ \left[2 \sum_{a,b} \sum_{\rho < \rho'} \sum_{\sigma < \sigma'} \frac{\partial^2 \mathcal{F}_n}{\partial r_a^{\rho,\rho'} \partial q_b^{\sigma,\sigma'}} \right]_{m,q,r} \delta r_a^{\rho,\rho'} \delta q_b^{\sigma,\sigma'} \end{aligned} \quad (\text{A.38})$$

Substituindo as devidas expressões para as derivadas obtemos

$$\delta^2 \mathcal{F}_n = \sum_{a,b} (A_{ab} - \frac{\beta}{\ell} (A^2)_{ab}) \sum_{\mu} \sum_{\rho} \frac{\delta m_a^{\mu,\rho}}{\ell} \frac{\delta m_b^{\mu,\rho}}{\ell}$$

$$\begin{aligned}
& + \sum_{a,b} \frac{\beta}{\ell} \sum_c A_{ac} A_{cb} \sum_{\mu,\nu} \langle \xi^\mu \xi^\nu \langle M_a^2 \rangle_z \rangle_\xi \sum_\rho \frac{\delta m_a^{\mu,\rho}}{\ell} \frac{\delta m_b^{\nu,\rho}}{\ell} \\
& + \sum_{a,b} \frac{\beta}{\ell} \sum_c A_{ac} A_{cb} \sum_{\mu,\nu} \langle \xi^\mu \xi^\nu (\langle M_a \rangle_z^2 - \langle M_a^2 \rangle_z) \rangle_\xi \left\{ \sum_\rho \frac{\delta m_a^{\mu,\rho}}{\ell} \right\} \left\{ \sum_\rho' \frac{\delta m_b^{\nu,\rho'}}{\ell} \right\} \\
& - 2\alpha \frac{\beta^2}{\ell} \sum_{a,b} A_{ab} \sum_\mu \langle \xi^\mu \langle M_b (1 - M_b^2) \rangle_z \rangle_\xi \sum_\rho \frac{\delta m_a^{\mu,\rho}}{\ell} \left\{ \sum_{\rho, \rho' \neq \rho} \delta r_b^{\rho\rho'} \right\} \\
& - 2\alpha \frac{\beta^2}{\ell} \sum_{a,b} A_{ab} \sum_\mu \langle \xi^\mu (\langle M_b^3 \rangle_z - \langle M_b \rangle_z \langle M_b^2 \rangle_z) \rangle_\xi \left\{ \sum_\rho \frac{\delta m_a^{\mu,\rho}}{\ell} \right\} \left\{ \sum_{\rho < \rho'} \delta r_b^{\rho\rho'} \right\} \\
& - \alpha^2 \frac{\beta^3}{\ell} \sum_a \langle \langle M_a^4 \rangle_z - \langle M_a^2 \rangle_z^2 \rangle_\xi \left\{ \sum_{\rho < \rho'} \delta r_a^{\rho\rho'} \right\}^2 \\
& - \alpha^2 \frac{\beta^3}{\ell} \sum_a \langle \langle M_a^4 \rangle_z - 2 \langle M_a^2 \rangle_z + 1 \rangle_\xi \sum_{\rho < \rho'} (\delta r_a^{\rho\rho'})^2 \\
& - \alpha^2 \frac{\beta^3}{\ell} \sum_a \langle \langle M_a^2 \rangle_z - \langle M_a^4 \rangle_z \rangle_\xi \sum_\rho \left\{ \sum_{\rho' \neq \rho} \delta r_a^{\rho\rho'} \right\}^2 \\
& - \frac{\alpha}{\beta} \sum_{a,b} 2(\mathcal{R}_{ab})^2 \left\{ \sum_{\rho < \rho'} \delta q_a^{\rho\rho'} \right\} \left\{ \sum_{\sigma < \sigma'} \delta q_b^{\sigma\sigma'} \right\} \\
& - \frac{\alpha}{\beta} \sum_{a,b} \mathcal{R}_{ab} \mathcal{R}'_{ab} \sum_\rho \left\{ \sum_{\rho' \neq \rho} \delta q_a^{\rho\rho'} \right\} \left\{ \sum_{\sigma \neq \rho} \delta q_b^{\rho\sigma} \right\} \\
& - \frac{\alpha}{\beta} \sum_{a,b} (\mathcal{R}'_{ab})^2 \sum_{\rho' \neq \rho} \delta q_a^{\rho\rho'} \delta q_b^{\rho\rho'} \\
& + 2 \frac{\alpha\beta}{\ell} \sum_a \sum_{\rho < \rho'} \delta q_a^{\rho\rho'} \delta r_a^{\rho\rho'}
\end{aligned} \tag{A.39}$$

Esta é a expressão completa para o termo de segunda ordem na expansão da função \mathcal{F}_n em torno dos parâmetros $m_a^{\mu,\rho}$, $q_a^{\rho\rho'}$ e $r_a^{\rho\rho'}$ no ponto de sela.

A.3 Estabilidade das Soluções Homogêneas

Agora devemos substituir na expressão (A.39) o tipo de soluções onde temos:

$$\begin{aligned}
 m_a^\mu &= m_\gamma v_a^\gamma \delta^{\mu 1} \\
 q_a &= q \\
 r_a &= r
 \end{aligned}
 \tag{A.40}$$

Calculando alguns termos separadamente:

$$\langle M_c^p \rangle_z = \int Dz \left\langle \tanh^p \left(\beta \left(v_c^\mu \frac{\lambda_\gamma}{\ell} m_\gamma \xi^1 + \sqrt{\alpha r_\gamma z} \right) \right) \right\rangle_\xi = \langle M^p \rangle_z (v_c^\gamma)^p \tag{A.41}$$

$$\langle \xi^\mu \xi^\nu \langle M_c^2 \rangle_z \rangle_\xi = \delta^{\mu\nu} \langle \langle M_c^2 \rangle_z \rangle_\xi = \delta^{\mu\nu} \langle M^2 \rangle_z \tag{A.42}$$

$$\langle \xi^\mu \xi^\nu (\langle M_c^2 \rangle_z - \langle M_c \rangle_z^2) \rangle_\xi = \delta^{\mu\nu} (\langle M^2 \rangle_z - \langle M \rangle_z^2) \tag{A.43}$$

$$\langle \xi^\mu \langle M_b (1 - M_b^2) \rangle_z \rangle_\xi = \delta^{\mu 1} \langle M (1 - M^2) \rangle_z v_b^\gamma \tag{A.44}$$

$$\langle \langle M_a^4 \rangle_z - \langle M_a^2 \rangle_z^2 \rangle_\xi = \langle M^4 \rangle_z - \langle M^2 \rangle_z^2 \tag{A.45}$$

$$\langle \langle (1 - M_a^2)^2 \rangle_z \rangle_\xi = \langle (1 - M^2)^2 \rangle_z \tag{A.46}$$

$$\begin{aligned}
 &\langle \xi^\mu (\langle M_b^3 \rangle_z - \langle M_b \rangle_z \langle M_b^2 \rangle_z) \rangle_\xi \\
 &= \delta^{\mu 1} (\langle M_b^3 \rangle_z - \langle M_b \rangle_z \langle M_b^2 \rangle_z) v_b^\gamma
 \end{aligned}
 \tag{A.47}$$

e

$$\mathcal{R}_{cd} = \sum_{k=2}^{\infty} \left(\frac{\beta}{\ell} \right)^k [\underline{A}^k]_{cd} q(k-1)(1-q)^{k-2}$$

$$\begin{aligned}
&= -q \frac{\partial}{\partial q} \left[\sum_{k=1}^{\infty} \left(\frac{\beta}{\ell} \right)^k \underline{A}^k q(k-1)(1-q)^{k-2} \right]_{cd} \\
&= \left(\frac{\beta}{\ell} \right)^2 q \left[\frac{\underline{A}^2}{(1 - \beta \frac{\underline{A}}{\ell} (1-q))^2} \right]_{cd} \tag{A.48}
\end{aligned}$$

e

$$\begin{aligned}
\mathcal{R}'_{cd} &= \frac{\beta}{\ell} A_{cd} + \sum_{k=2}^{\infty} \left(\frac{\beta}{\ell} \right)^k [\underline{A}^k]_{cd} (1-q)^{k-2} \\
&= \frac{\beta}{\ell} A_{cd} + \left[\frac{\beta \underline{A}}{\ell} \sum_{k=1}^{\infty} \left(\frac{\beta \underline{A}}{\ell} (1-q) \right)^k \right]_{cd} = \frac{\beta}{\ell} \left[\frac{\underline{A}}{1 - \beta \frac{\underline{A}}{\ell} (1-q)} \right]_{cd} \tag{A.49}
\end{aligned}$$

Substituindo em (A.39) obtemos

$$\begin{aligned}
\delta^2 \mathcal{F}_n &= \sum_{a,b} (A_{ab} - \frac{\beta}{\ell} (1 - \langle M^2 \rangle_z) [\underline{A}^2]_{ab}) \sum_{\mu} \sum_{\rho} \frac{\delta m_a^{\mu,\rho}}{\ell} \frac{\delta m_b^{\mu,\rho}}{\ell} \\
&+ \frac{\beta}{\ell} (\langle M \rangle_z^2 - \langle M^2 \rangle_z) \sum_{a,b} [\underline{A}]_{ab} \sum_{\mu} \left\{ \sum_{\rho} \frac{\delta m_a^{\mu,\rho}}{\ell} \right\} \left\{ \sum_{\rho} \frac{\delta m_b^{\mu,\rho'}}{\ell} \right\} \\
&- 2\alpha \frac{\beta^2}{\ell} \langle M(1-M^2) \rangle_z \sum_{a,b} A_{ab} v_b^{\gamma} \sum_{\rho} \frac{\delta m_a^{1,\rho}}{\ell} \left\{ \sum_{\rho, \rho' \neq \rho} \delta r_b^{\rho\rho'} \right\} \\
&- 2\alpha \frac{\beta^2}{\ell} (\langle M^3 \rangle_z - \langle M \rangle_z \langle M^2 \rangle_z) \sum_{a,b} A_{ab} v_b^{\gamma} \left\{ \sum_{\rho} \frac{\delta m_a^{1,\rho}}{\ell} \right\} \left\{ \sum_{\rho < \rho'} \delta r_b^{\rho\rho'} \right\} \\
&- \alpha^2 \frac{\beta^3}{\ell} (\langle M^4 \rangle_z - \langle M^2 \rangle_z^2) \sum_a \left\{ \sum_{\rho < \rho'} \delta r_a^{\rho\rho'} \right\}^2 \\
&- \alpha^2 \frac{\beta^3}{\ell} (\langle M^4 \rangle_z - 2 \langle M^2 \rangle_z + 1) \sum_a \sum_{\rho < \rho'} (\delta r_a^{\rho\rho'})^2 \\
&- \alpha^2 \frac{\beta^3}{\ell} (\langle M^2 \rangle_z - \langle M^4 \rangle_z) \sum_a \sum_{\rho} \left\{ \sum_{\rho' \neq \rho} \delta r_a^{\rho\rho'} \right\}^2
\end{aligned}$$

$$\begin{aligned}
& - \frac{2\alpha}{\beta} \sum_{a,b} (\mathcal{R}_{ab})^2 \left\{ \sum_{\rho < \rho'} \delta q_a^{\rho\rho'} \right\} \left\{ \sum_{\sigma < \sigma'} \delta q_b^{\sigma\sigma'} \right\} \\
& - \frac{\alpha}{\beta} \sum_{a,b} \mathcal{R}_{ab} \mathcal{R}'_{ab} \sum_{\rho} \left\{ \sum_{\rho' \neq \rho} \delta q_a^{\rho\rho'} \right\} \left\{ \sum_{\sigma \neq \rho} \delta q_b^{\rho\sigma} \right\} \\
& - \frac{\alpha}{\beta} \sum_{a,b} (\mathcal{R}'_{ab})^2 \sum_{\rho' \neq \rho} \delta q_a^{\rho\rho'} \delta q_b^{\rho\rho'} \\
& + 2 \frac{\alpha\beta}{\ell} \sum_a \sum_{\rho < \rho'} \delta q_a^{\rho\rho'} \delta \eta_a^{\rho\rho'}
\end{aligned} \tag{A.50}$$

Para $\delta m_a^{\mu,\rho}$ com $\mu \neq 1$ podemos fazer a seguinte troca de variáveis

$$\delta m_a^{\mu,\rho} = \sum_c V_{ac} \delta \eta_c^{\mu,\rho} \tag{A.51}$$

onde \underline{V} é uma matriz unitária que diagonaliza \underline{A} . Como esta é uma transformação ortogonal o seu jacobiano é 1. Usando isso nos dois primeiros termos de (A.50) resulta

$$\begin{aligned}
& \sum_a \left(\frac{\lambda_a}{\ell} - \left(\frac{\beta}{\ell} \right)^2 (1 - \langle M^2 \rangle_z) \left(\frac{\lambda_a}{\ell} \right)^2 \right) \sum_{\mu} \sum_{\rho} (\delta \eta_a^{\mu,\rho})^2 \\
& + \beta \sum_a \left(\frac{\lambda_a}{\ell} \right)^2 (\langle M \rangle_z^2 - \langle M^2 \rangle_z) \sum_{\mu} \left\{ \sum_{\rho} \delta \eta_a^{\mu,\rho} \right\}^2 \\
& = \sum_a \sum_{\mu} \sum_{\rho, \rho'} (c_a \delta_{\rho\rho'} + c'_a) \delta \eta_a^{\mu,\rho} \delta \eta_a^{\mu,\rho'}
\end{aligned} \tag{A.52}$$

Fazendo outra troca de variáveis, de η para η' no sentido de diagonalizar o espaço de réplicas obtemos

$$\sum_a \sum_{\mu} \sum_{\rho} \Lambda_{\rho,a} (\delta \eta_a^{\mu,\rho})^2 \tag{A.53}$$

com

$$\Lambda_{1,a} = \frac{\lambda_a}{\ell} - \left(\frac{\beta}{\ell}\right)^2 (1 - \langle M^2 \rangle_z) - \beta n \left(\frac{\lambda_a}{\ell}\right)^2 (\langle M \rangle_z^2 - \langle M^2 \rangle_z) \quad (\text{A.54})$$

$$\Lambda_{\rho,a} = \frac{\lambda_a}{\ell} - \left(\frac{\beta}{\ell}\right)^2 (1 - \langle M^2 \rangle_z) \quad (\text{A.55})$$

Vamos agora estudar os termos onde $m_a^{1,\rho}$ acopla com $r_a^{\rho\rho'}$. Fazendo

$$\begin{aligned} \delta m_a^{1,\rho} &= \sum_c V_{ac} \delta \eta_c^{1,\rho} \\ \delta r_a^{\rho\rho'} &= v_a^\gamma \sum_c V_{ca} \delta r_c^{\rho\rho'} \\ \delta q_a^{\rho\rho'} &= v_a^\gamma \sum_c V_{ac} \delta q_c^{\rho\rho'} \end{aligned} \quad (\text{A.56})$$

obtemos de (A.50)

$$\begin{aligned} \delta^2 \mathcal{F}_n &= \sum_a \sum_\mu \sum_\rho \Lambda_a^\rho (\delta \eta_a^{\mu,\rho})^2 \\ &- 2\alpha \frac{\beta^2}{\ell} \langle M(1 - M^2) \rangle_z \sum_a \frac{\lambda_a}{\ell} \sum_\rho \delta \eta_a^{1,\rho} \left\{ \sum_{\rho, \rho' \neq \rho} \delta r_b^{\rho\rho'} \right\} \\ &- 2\alpha \frac{\beta^2}{\ell} (\langle M^3 \rangle_z - \langle M \rangle_z \langle M^2 \rangle_z) \sum_a \frac{\lambda_a}{\ell} \left\{ \sum_\rho \delta \eta_a^{1,\rho} \right\} \left\{ \sum_{\rho < \rho'} \delta r_b^{\rho\rho'} \right\} \\ &- \alpha^2 \frac{\beta^3}{\ell} (\langle M^4 \rangle_z - \langle M^2 \rangle_z^2) \sum_a \left\{ \sum_{\rho < \rho'} \delta r_a^{\rho\rho'} \right\}^2 \\ &- \alpha^2 \frac{\beta^3}{\ell} (\langle M^4 \rangle_z - 2 \langle M^2 \rangle_z + 1) \sum_a \sum_{\rho < \rho'} (\delta r_a^{\rho\rho'})^2 \\ &- \alpha^2 \frac{\beta^3}{\ell} (\langle M^2 \rangle_z - \langle M^4 \rangle_z) \sum_a \sum_\rho \left\{ \sum_{\rho' \neq \rho} \delta r_a^{\rho\rho'} \right\}^2 \end{aligned}$$

$$\begin{aligned}
& - \frac{2\alpha}{\beta} \sum_{a,b} \left(\sum_{cd} V_{ac} (\mathcal{R}_{cd})^2 V_{db} \right) \left\{ \sum_{\rho < \rho'} \delta q_a'^{\rho\rho'} \right\} \left\{ \sum_{\sigma < \sigma'} \delta q_b'^{\sigma\sigma'} \right\} v_a^\gamma v_b^\gamma \\
& - \frac{\alpha}{\beta} \sum_{a,b} \left(\sum_{cd} V_{ac} \mathcal{R}_{cd} \mathcal{R}'_{cd} V_{db} \right) \sum_{\rho} \left\{ \sum_{\rho' \neq \rho} \delta q_a'^{\rho\rho'} \right\} \left\{ \sum_{\sigma \neq \rho} \delta q_b'^{\rho\sigma} \right\} v_a^\gamma v_b^\gamma \\
& - \frac{\alpha}{\beta} \sum_{a,b} \left(\sum_{cd} V_{ac} (\mathcal{R}'_{cd})^2 V_{db} \right) \sum_{\rho' \neq \rho} \delta q_a'^{\rho\rho'} \delta q_b'^{\rho\rho'} v_a^\gamma v_b^\gamma \\
& + 2 \frac{\alpha\beta}{\ell} \sum_a \sum_{\rho < \rho'} \delta q_a'^{\rho\rho'} \delta r_a'^{\rho\rho'}
\end{aligned} \tag{A.57}$$

Devido às características de $\underline{\underline{A}}$ as matrizes de elementos $\{\mathcal{R}_{ab}^2\}$, $\{\mathcal{R}'_{ab}{}^2\}$ e $\{\mathcal{R}_{ab} \mathcal{R}'_{ab}{}^2\}$, que são funções de $\underline{\underline{A}}$, também são diagonalizáveis por $\underline{\underline{V}}$. Assim:

$$\sum_{cd} V_{ac} (\mathcal{R}_{cd})^2 V_{db} = \delta^{ab} \mathcal{R}_a \tag{A.58}$$

$$\sum_{cd} V_{ac} \mathcal{R}_{cd} \mathcal{R}'_{cd} V_{db} = \delta^{ab} \mathcal{R}'_a \tag{A.59}$$

$$\sum_{cd} V_{ac} (\mathcal{R}'_{cd})^2 V_{db} = \delta^{ab} \mathcal{R}''_a \tag{A.60}$$

Onde \mathcal{R}_a , \mathcal{R}'_a e \mathcal{R}''_a são os autovalores destas matrizes.

A partir deste ponto podemos separar as flutuações em diferentes tipos e calcular os autovalores para cada um deles.

A.3.1 Flutuações Fortemente Assimétricas

Esta flutuação é definida como aquela tal que:

$$\delta \eta_a^{1\rho} = 0$$

$$\begin{aligned}\sum_{\rho'} \delta q_a^{\rho\rho'} &= 0 \\ \sum_{\rho'} \delta r_a^{\rho\rho'} &= 0\end{aligned}\tag{A.61}$$

que substituídas em A.57) resulta:

$$\begin{aligned}\delta^2 \mathcal{F}_n |_{F.A.} &= -\frac{\alpha^2 \beta^3}{2\ell} \langle (1 - M^2)^2 \rangle_z \sum_a \sum_{\rho \neq \rho'} (\delta r_a^{\rho\rho'})^2 \\ &+ \frac{\alpha\beta}{\ell} \sum_a \sum_{\rho < \rho'} \delta q_a^{\rho\rho'} \delta r_a^{\rho\rho'} \\ &- \frac{\alpha}{2\beta} \sum_a \mathcal{R}_a'' \sum_{\rho' \neq \rho} (\delta q_a^{\rho\rho'})^2\end{aligned}\tag{A.62}$$

A princípio as somas já são diagonais em réplicas e blocos bastando somente desacoplar $q_a^{\rho\rho'}$ de $r_a^{\rho\rho'}$. Faremos isto completando os quadrados:

$$\delta^2 \mathcal{F}_n = A\delta^2\eta + B\delta\eta\delta r + C\delta^2 r + D\delta r\delta q + E\delta^2 q\tag{A.63}$$

$$= A(\delta\eta k\delta r)^2 + (B - 2kA)\delta\eta\delta r(C - Ak^2)\delta^2 r + D\delta r\delta q + E\delta^2 q\tag{A.64}$$

Escolhendo k tal que $B - 2kA = 0$ e definindo $\delta\xi_\eta = \delta\eta + k\delta r$

$$\delta^2 \mathcal{F}_n = A(\delta\xi_\eta)^2 + (C - \frac{B^2}{4A})\delta^2 r + D\delta r\delta q + E\delta^2 q\tag{A.65}$$

$$\delta^2 \mathcal{F}_n = A(\delta\xi_\eta)^2 + (C - \frac{B^2}{4A})(i\delta\xi_r)^2 + (E - \frac{D^2}{4(C - B^2/4A)})(\delta q)^2\tag{A.66}$$

onde a última expressão foi obtida fazendo $i\delta\xi_r = \delta r + k'\delta q$. Onde, novamente k' tem a função de cancelar o termo cruzado e o número imaginário i foi utilizado para fazer ξ_r ser real.

O resultado final para os autovalores do modo fortemente assimétrico é , então :

$$\Lambda_{\rho,a}^r |_{FA} = \frac{\alpha^2 \beta^3}{2\ell} \langle (1 - m^2)^2 \rangle_z \quad (\text{A.67})$$

$$\Lambda_{\rho,a}^q |_{FA} = \frac{1}{2} \left(\frac{1}{\beta\ell \langle (1 - m^2)^2 \rangle_z} - \alpha\beta^3 \mathcal{R}_a'' \right) \quad (\text{A.68})$$

Onde foi usado (A.65) com $A, B = 0$.

A.3.2 Flutuações Levemente Assimétricas

Este tipo de flutuação é caracterizado por

$$\begin{aligned} \delta q_a^{\rho\rho'} &= \delta q_a^\rho + \delta q_a^{\rho'} \\ \delta r_a^{\rho\rho'} &= \delta r_a^\rho + \delta r_a^{\rho'} \\ \sum_\rho \delta \eta_a^{1\rho} &= \sum_{\rho'} \delta q_a^{\rho\rho'} = \sum_{\rho'} \delta r_a^{\rho\rho'} = 0 \end{aligned} \quad (\text{A.69})$$

De forma que

$$\begin{aligned} \sum_{\rho \neq \rho'} \delta^2 q_a^{\rho\rho'} &= \sum_{\rho \neq \rho'} (\delta^2 q_a^\rho + 2\delta q_a^\rho \delta q_a^{\rho'} + \delta^2 q_a^{\rho'}) \\ &= 2(n-1) \sum_\rho \delta^2 q_a^\rho + 2 \sum_\rho \delta q_a^\rho (\sum_{\rho'} \delta q_a^{\rho'} - \delta q_a^\rho) = 2(n-2) \sum_\rho \delta q_a^\rho \quad (\text{A.70}) \end{aligned}$$

$$\begin{aligned} \sum_{\rho \neq \rho'} \delta q_a^{\rho\rho'} &= \sum_{\rho \neq \rho'} (\delta q_a^\rho + \delta q_a^{\rho'}) \\ &= \sum_\rho \delta q_a^\rho - \delta q_a^{\rho'} + (n-1)\delta q_a^{\rho'} = (n-2)\delta q_a^{\rho'} \quad (\text{A.71}) \end{aligned}$$

Substituindo estas expressões em (A.57) resulta

$$\begin{aligned}
\delta^2 \mathcal{F}_n |_{L.A.} &= \sum_a \sum_\rho \left\{ \left(\frac{\lambda_a}{\ell} - \beta \left(\frac{\lambda_a}{\ell} \right)^2 (1 - \langle M \rangle_z) \right) (\delta \eta_a^{1\rho})^2 \right. \\
&- (n-2) \frac{2\alpha\beta^2}{\ell} \langle M(1-M^2) \rangle_z \frac{\lambda_a}{\ell} \delta \eta_a^{1\rho} \delta r_a^\rho \\
&- (n-2) \frac{\alpha^2\beta^3}{\ell} \langle (1-M^2)^2 \rangle_z (\delta r_a^\rho)^2 \\
&- (n-2)^2 \frac{\alpha^2\beta^3}{\ell} (\langle M^2 \rangle_z - \langle M^4 \rangle_z) (\delta r_a^\rho)^2 \\
&- (n-2) \frac{\alpha}{\beta} (\mathcal{R}''_a + (n-2)\mathcal{R}'_a) (\delta q_a^\rho)^2 \\
&\left. + (n-2) 2 \frac{\alpha\beta}{\ell} \delta q_a^\rho \delta r_a^\rho \right\} \tag{A.72}
\end{aligned}$$

completando os quadrados, como em (A.65), obtemos para os autovalores deste modo:

$$\begin{aligned}
\Lambda_{\rho,a}^\eta |_{LA} &= \frac{\lambda_a}{\ell} - \beta \left(\frac{\lambda_a}{\ell} \right)^2 (1 - \langle M \rangle_z) \\
\Lambda_{\rho,a}^r |_{LA} &= (n-2)^2 \frac{\alpha^2\beta^3}{\ell} \left\{ \frac{\beta}{\ell} \left(\frac{\lambda_a}{\ell} \right)^2 \frac{\langle M(1-M^2) \rangle_z^2}{\Lambda_{\rho,a}^\eta |_{LA}} + \frac{\langle (1-M^2)(1+(n-3)M^2) \rangle_z}{n-2} \right\} \\
\Lambda_{\rho,a}^q |_{LA} &= (n-2)^2 \frac{\alpha^2\beta^2}{\ell^2} \frac{1}{\Lambda_{\rho,a}^\eta |_{LA}} + (2-n) \frac{\alpha}{\beta} (\mathcal{R}''_a + (n-2)\mathcal{R}'_a) \tag{A.73}
\end{aligned}$$

O último tipo de flutuações que estudaremos é o mais importante para a discussão de soluções feita no capítulo 3.

A.3.3 Flutuações Simétricas

$$\begin{aligned}
\delta\eta_a^{1\rho} &= \delta\eta_a^1 \\
\delta q_a^{\rho\rho'} &= \delta q_a \\
\delta r_a^{\rho\rho'} &= \delta r_a
\end{aligned} \tag{A.74}$$

Resultando para $\delta^2\mathcal{F}_n$ em primeira ordem em n :

$$\begin{aligned}
\delta^2\mathcal{F}_n &= n \sum_a \left\{ \left(\frac{\lambda_a}{\ell} - \beta \left(\frac{\lambda_a}{\ell} \right)^2 (1 - \langle M \rangle_z) \right) (\delta\eta_a^1)^2 \right. \\
&+ 2\alpha \frac{\beta^2}{\ell} \langle M(1 - M^2) \rangle_z \frac{\lambda_a}{\ell} \delta\eta_a^1 \delta r'_b \\
&+ \alpha^2 \beta^3 \ell \left(\frac{1}{2} \langle (1 - M^2)^2 \rangle_z - \langle M^2(1 - M^2) \rangle_z \right) (\delta r_a)^2 \\
&- \frac{\alpha\beta}{\ell} \delta q_a \delta r_a - \frac{\alpha}{\beta} (\mathcal{R}'_a - \frac{1}{2} \mathcal{R}''_a) (\delta q_a)^2
\end{aligned} \tag{A.75}$$

Os autovalores do modo simétrico são

$$\begin{aligned}
\Lambda_{\rho,a}^\eta |_S &= \frac{\lambda_a}{\ell} - \beta \left(\frac{\lambda_a}{\ell} \right)^2 (1 - \langle M \rangle_z) \\
\Lambda_{\rho,a}^r |_S &= \frac{\alpha^2 \beta^3}{\ell} \left\{ \frac{\beta}{\ell} \left(\frac{\lambda_a}{\ell} \right)^2 \frac{\langle M(1 - M^2) \rangle_z^2}{\Lambda_{\rho,a}^\eta |_S} - \frac{1}{2} \langle (1 - M^2)(1 - 3)M^2 \rangle_z \right\} \\
\Lambda_{\rho,a}^q |_{LA} &= \frac{\alpha^2 \beta^2}{\ell^2} \frac{1}{\Lambda_{\rho,a}^\eta |_S} - \frac{\alpha}{\beta} (\mathcal{R}'_a - \frac{1}{2} \mathcal{R}''_a)
\end{aligned} \tag{A.76}$$

As condições para a estabilidade das soluções de (3.41), (3.42) e (3.43) saem de exigirmos que estes autovalores sejam positivos. Vemos que ao considerarmos ordem mais baixa em n , para cada autovalor, muitos destes resultam iguais e ao final temos somente quatro condições de estabilidade:

$$\Lambda^n \Rightarrow \frac{\lambda_a}{\ell} - \beta \left(\frac{\lambda_a}{\ell} \right)^2 (1 - \langle M \rangle_z) > 0 \quad (\text{A.77})$$

$$\Lambda^q |_{FA} \Rightarrow 1 - \alpha \beta^4 \mathcal{R}_a'' \ell \langle (1 - M^2)^2 \rangle_z > 0 \quad (\text{A.78})$$

$$\Lambda^r |_{S,LA} \Rightarrow \frac{\beta}{\ell} \left(\frac{\lambda_a}{\ell} \right)^2 \frac{\langle M(1 - M^2) \rangle_z^2}{\Lambda^n} - \frac{1}{2} \langle (1 - M^2)(1 - 3)M^2 \rangle_z > 0 \quad (\text{A.79})$$

$$\Lambda^r |_{S,LA} \Rightarrow \frac{\beta}{\ell} \left(\frac{\lambda_a}{\ell} \right)^2 \frac{\langle M(1 - M^2) \rangle_z^2}{\Lambda^n} - \frac{1}{2} \langle (1 - M^2)(1 - 3)M^2 \rangle_z > 0 \quad (\text{A.80})$$

$$\Lambda^q |_{S,LA} \Rightarrow \frac{\alpha \beta^3}{\ell^2} \frac{1}{\Lambda^r |_{S,LA}} - \mathcal{R}'_a + \frac{1}{2} \mathcal{R}''_a > 0 \quad (\text{A.81})$$

com \mathcal{R}'_a e \mathcal{R}''_a definidos em (A.58).

Apêndice B

Cálculos de Gardner com descendentes

B.1 Cálculo das equações de ponto de sela

As equações de ponto de sela são resultado de igualar a zero as derivadas primeiras de g em relação a q_1 e q_2 .

$$q_c = 4\alpha\epsilon_c^2 \int Dz_1 Dz_2 \frac{1}{\mathcal{W}} \left(-\frac{\partial \mathcal{W}}{\partial q_c} \right) \quad c = 1, 2 \quad (\text{B.1})$$

operando explicitamente as derivadas temos

$$\mathcal{W}_1 = -\frac{\partial \mathcal{W}}{\partial q_1} \quad (\text{B.2})$$

$$= - \int_{C_1}^{\infty} D\lambda \frac{\partial}{\partial q_1} \left\{ H\left(\sqrt{\frac{\epsilon_1}{\epsilon_2}}(-\lambda + C_1 + C_2)\right) - H\left(\sqrt{\frac{\epsilon_1}{\epsilon_2}}(\lambda - C_1 + C_2)\right) \right\} \quad (\text{B.3})$$

$$= \int_{C_1}^{\infty} D\lambda \frac{1}{2\sqrt{\epsilon_1\epsilon_2}} (\lambda + \sqrt{\frac{\epsilon_1}{q_1}} z_1) \left\{ e^{\frac{1}{2} \frac{\epsilon_1}{\epsilon_2} (\lambda - C_1 - C_2)^2} + e^{\frac{1}{2} \frac{\epsilon_1}{\epsilon_2} (\lambda - C_1 + C_2)^2} \right\} \quad (\text{B.4})$$

Tomando somente o primeiro dos termos da soma, observe que o outro nada mais é

que este quando trocamos C_2 por $-C_2$, podemos escrevê-lo como

$$I = \int_0^\infty \frac{dt}{\sqrt{2\pi}} e^{-\frac{1}{2}(t-C_1)^2 - \frac{1}{2} \frac{\epsilon_1 \epsilon_2}{\epsilon_1 + \epsilon_2} (t-C_2)^2} (t + C_1 + \sqrt{\frac{\epsilon_1}{q_1}} z_1) \quad (\text{B.5})$$

$$= e^{\frac{1}{2}z_1^2} \sqrt{\frac{\epsilon_1}{q_1}} \int_0^\infty \frac{dt}{\sqrt{2\pi}} e^{-\frac{1}{2} \frac{\epsilon_1}{\epsilon_2} (t-C_2)^2 - \frac{1}{2}(z_1^2 + t^2 + 2tC_1 + C_1^2)} (z_1 + (t + C_1) \sqrt{\frac{q_1}{\epsilon_1}}) \quad (\text{B.6})$$

$$= -e^{\frac{1}{2}z_1^2} \sqrt{\frac{\epsilon_1}{q_1}} \frac{\partial}{\partial z_1} \left\{ e^{-\frac{1}{2}z_1^2} \int_{C_1}^\infty \frac{d\lambda}{\sqrt{2\pi}} e^{-\frac{1}{2} \frac{\epsilon_1}{\epsilon_2} (t-C_1-C_2)^2} \right\} \quad (\text{B.7})$$

$$= -e^{\frac{1}{2}z_1^2} \sqrt{\frac{\epsilon_1 \epsilon_2}{q_1(\epsilon_1 + \epsilon_2)}} \frac{\partial}{\partial z_1} \left\{ e^{-\frac{1}{2}z_1^2} e^{-\frac{1}{2} \frac{\epsilon_1}{\epsilon_1 + \epsilon_2} (C_1 + C_2)^2} H(\zeta_-) \right\} \quad (\text{B.8})$$

onde $\zeta_\pm = \frac{\epsilon_2 C_1 \pm \epsilon_1 C_2}{\sqrt{\epsilon_2(\epsilon_1 + \epsilon_2)}}$.

Substituindo na expressão inicial podemos integrar por partes a variável z_1 ,

$$\begin{aligned} & \sqrt{\frac{\epsilon_1 \epsilon_2}{q_1(\epsilon_1 + \epsilon_2)}} \int D z_2 \int_{-\infty}^\infty \frac{d z_1}{\sqrt{2\pi}} \frac{1}{\mathcal{W}} \left(\frac{\partial}{\partial z_1} \left\{ e^{-\frac{1}{2}z_1^2} e^{-\frac{1}{2} \frac{\epsilon_1}{\epsilon_1 + \epsilon_2} (C_1 + C_2)^2} H(\zeta_-) \right\} \right)_{z=z_1} \\ &= \sqrt{\frac{\epsilon_1 \epsilon_2}{q_1(\epsilon_1 + \epsilon_2)}} \int D z_1 D z_2 \frac{e^{-\frac{1}{2} \frac{\epsilon_1}{\epsilon_1 + \epsilon_2} (C_1 + C_2)^2} H(\zeta_-)}{\mathcal{W}^2} \left(\frac{\partial \mathcal{W}}{\partial z} \right)_{z=z_1} \end{aligned} \quad (\text{B.9})$$

calculando esta última derivada

$$\frac{\partial \mathcal{W}}{\partial z} = \sqrt{\frac{q_1}{\epsilon_2}} \int_{C_1}^\infty \frac{D \lambda}{\sqrt{2\pi}} \left\{ e^{\frac{1}{2} \frac{\epsilon_1}{\epsilon_2} (\lambda - C_1 - C_2)^2} + e^{\frac{1}{2} \frac{\epsilon_1}{\epsilon_2} (\lambda - C_1 + C_2)^2} \right\} \quad (\text{B.10})$$

$$= \sqrt{\frac{q_1}{2\pi(\epsilon_1 + \epsilon_2)}} \left(e^{-\frac{1}{2} \frac{\epsilon_1}{\epsilon_1 + \epsilon_2} (C_1 + C_2)^2} H(\zeta_-) + e^{-\frac{1}{2} \frac{\epsilon_1}{\epsilon_1 + \epsilon_2} (C_1 + C_2)^2} H(\zeta_+) \right) \quad (\text{B.11})$$

Se agora juntarmos os dois termos em (B.1) obtemos

$$\int Dz_1 Dz_2 \frac{\mathcal{W}_1}{\mathcal{W}} = \frac{1}{4\pi} \frac{1}{\epsilon_1 + \epsilon_2} \int Dz_1 Dz_2 \left(\frac{e^{-\frac{1}{2} \frac{\epsilon_1}{\epsilon_1 + \epsilon_2} (C_1 + C_2)^2} H(\zeta_-) + e^{-\frac{1}{2} \frac{\epsilon_1}{\epsilon_1 + \epsilon_2} (C_1 + C_2)^2} H(\zeta_+)}{\mathcal{W}} \right)^2 \quad (\text{B.12})$$

o cálculo da equação para q_2 é muito semelhante resultando

$$\int Dz_1 Dz_2 \frac{\mathcal{W}_1}{\mathcal{W}} = \frac{1}{4\pi} \frac{1}{\epsilon_1 + \epsilon_2} \int Dz_1 Dz_2 \left(\frac{e^{-\frac{1}{2} \frac{\epsilon_1}{\epsilon_1 + \epsilon_2} (C_1 + C_2)^2} H(\zeta_-) - e^{-\frac{1}{2} \frac{\epsilon_1}{\epsilon_1 + \epsilon_2} (C_1 + C_2)^2} H(\zeta_+)}{\mathcal{W}} \right)^2 \quad (\text{B.13})$$

B.2 Cálculo das equações no limite $\epsilon_1, \epsilon_2 \rightarrow 0$

Primeiramente começamos por escrever a função \mathcal{W} como

$$\mathcal{W} = \Theta(-C_1) \int_{-\infty}^{\infty} D\lambda \left\{ H\left(\sqrt{\frac{\epsilon_1}{\epsilon_2}}(-\lambda + C_1 + C_2)\right) - H\left(\sqrt{\frac{\epsilon_1}{\epsilon_2}}(\lambda - C_1 + C_2)\right) \right\} + \Delta(|C_1|) \quad (\text{B.14})$$

onde

$$\Delta(|C_1|) = \int_{|C_1|}^{\infty} D\lambda \left\{ H\left(\sqrt{\frac{\epsilon_1}{\epsilon_2}}(-\lambda + C_1 + C_2)\right) - H\left(\sqrt{\frac{\epsilon_1}{\epsilon_2}}(\lambda - C_1 + C_2)\right) \right\} \quad (\text{B.15})$$

B.2.1 Cálculo de $\Delta(|C_1|)$

$$\Delta(|C_1|) = \int_{|C_1|}^{\infty} d\{-H(\lambda)\} \left\{ H\left(\sqrt{\frac{\epsilon_1}{\epsilon_2}}(-\lambda + C_1 + C_2)\right) - H\left(\sqrt{\frac{\epsilon_1}{\epsilon_2}}(\lambda - C_1 + C_2)\right) \right\} \quad (\text{B.16})$$

que integrado por partes e considerando que $|C_1| \rightarrow \infty$ nos permite utilizar

$$H(\lambda) = \frac{1}{\sqrt{2\pi}} \frac{e^{\frac{\lambda^2}{2}}}{\lambda} \{1 + o(\lambda^{-2})\} \quad (\text{B.17})$$

e portanto

$$\Delta(|C_1|) = \frac{1}{2\pi} \sqrt{\frac{\epsilon_1}{\epsilon_2}} \{F(C_2) - F(-C_2)\} \quad (\text{B.18})$$

com

$$F(C_2) = \int_{|C_1|}^{\infty} \frac{D\lambda}{\lambda} \exp\left\{-\frac{1}{2} \frac{\epsilon_1}{\epsilon_2} (\lambda - |C_1| - C_2)^2\right\} \quad (\text{B.19})$$

Trocando variáveis $\lambda = t + |C_1|$ obtemos

$$F(C_2) = e^{-\frac{1}{2} \frac{\epsilon_1}{\epsilon_1 + \epsilon_2} (|C_1| + C_2)^2} \{ \Theta(-\bar{\zeta}_-) I_- + \Theta(\bar{\zeta}_-) I_+ \} \quad (\text{B.20})$$

com

$$I_{\pm} = \int_0^{\infty} dt \frac{1}{|C_1| + t} \exp\left\{-\frac{1}{2} \frac{\epsilon_1 + \epsilon_2}{\epsilon_2} \left(t \pm \sqrt{\frac{\epsilon_2}{\epsilon_1 + \epsilon_2}} |\bar{\zeta}_-|\right)^2\right\} \quad (\text{B.21})$$

e

$$\bar{\zeta}_- = \frac{1}{\sqrt{\epsilon_2(\epsilon_1 + \epsilon_2)}} (\epsilon_2 |C_1| - \epsilon_1 C_2) \quad (\text{B.22})$$

No caso de I_- a parte dominante na integral é em torno de $t \equiv \frac{\epsilon_2}{\epsilon_1 + \epsilon_2} |\bar{\zeta}_-|$ assim podemos considerar

$$\frac{1}{|C_1| + t} = \frac{1}{|C_1| + \frac{\epsilon_2}{\epsilon_1 + \epsilon_2} |\bar{\zeta}_-| + t - \frac{\epsilon_2}{\epsilon_1 + \epsilon_2} |\bar{\zeta}_-|} \quad (\text{B.23})$$

$$= \frac{1}{|C_1| + \frac{\epsilon_2}{\epsilon_1 + \epsilon_2} |\bar{\zeta}_-|} \left\{ 1 - \frac{t - \frac{\epsilon_2}{\epsilon_1 + \epsilon_2} |\bar{\zeta}_-|}{|C_1| + \frac{\epsilon_2}{\epsilon_1 + \epsilon_2} |\bar{\zeta}_-|} + \dots \right\} \quad (\text{B.24})$$

tomando em primeira ordem

$$I_- = \frac{\sqrt{2\pi}}{|C_1| + \frac{\epsilon_2}{\epsilon_1 + \epsilon_2} |\bar{\zeta}_-|} \sqrt{\frac{\epsilon_2}{\epsilon_1 + \epsilon_2}} H(-|\bar{\zeta}_-|) \quad (\text{B.25})$$

Já para I_+ a parte dominante está em $t \equiv 0$ logo

$$I_+ = \frac{\sqrt{2\pi}}{|C_1|} \sqrt{\frac{\epsilon_2}{\epsilon_1 + \epsilon_2}} H(|\bar{\zeta}_-|) \quad (\text{B.26})$$

Voltando a $F(C_2)$:

$$F(C_2) = \sqrt{\frac{2\pi\epsilon_2}{\epsilon_1 + \epsilon_2}} e^{-\frac{1}{2} \frac{\epsilon_2}{\epsilon_1 + \epsilon_2} (|C_1| + C_2)^2} H(\bar{\zeta}_-) \left\{ \frac{\Theta(-\bar{\zeta}_-)}{|C_1| + \sqrt{\frac{\epsilon_2}{\epsilon_1 + \epsilon_2}} \bar{\zeta}_-} + \frac{\Theta(\bar{\zeta}_-)}{|C_1|} \right\} \quad (\text{B.27})$$

e a $\Delta(|C_1|)$

$$\Delta(|C_1|) = \sqrt{\frac{\epsilon_1}{2\pi(\epsilon_1 + \epsilon_2)}} \left\{ e^{-\frac{1}{2} \frac{\epsilon_2}{\epsilon_1 + \epsilon_2} (|C_1| + C_2)^2} H(\bar{\zeta}_-) \left\{ \frac{\Theta(-\bar{\zeta}_-)}{|C_1| + \sqrt{\frac{\epsilon_2}{\epsilon_1 + \epsilon_2}} \bar{\zeta}_-} + \frac{\Theta(\bar{\zeta}_-)}{|C_1|} \right\} + \right.$$

$$e^{-\frac{1}{2} \frac{\epsilon_2}{\epsilon_1 + \epsilon_2} (|C_1| - C_2)^2} H(\bar{\zeta}_+) \left\{ \frac{\Theta(-\bar{\zeta}_+)}{|C_1| + \sqrt{\frac{\epsilon_2}{\epsilon_1 + \epsilon_2}} \zeta_+} + \frac{\Theta(\bar{\zeta}_+)}{|C_1|} \right\} \quad (B.28)$$

Nos interessa determinar

$$K^\pm = \frac{e^{-\frac{1}{2} \frac{\epsilon_2}{\epsilon_1 + \epsilon_2} (C_1 + C_2)^2} H(\zeta_-) \pm e^{-\frac{1}{2} \frac{\epsilon_2}{\epsilon_1 + \epsilon_2} (C_1 - C_2)^2} H(\zeta_+)}{\mathcal{W}^2} \quad (B.29)$$

no limite em que $\epsilon_1, \epsilon_2 \rightarrow \infty$ com $\epsilon_1/\epsilon_2 = \text{constante}$. Primeiramente temos que

$$\begin{aligned} \frac{1}{\mathcal{W}^2} &= \Theta(-C_1) \frac{1}{\left\{ H(\sqrt{\frac{\epsilon_1}{\epsilon_1 + \epsilon_2}} (C_1 + C_2)) - H(\sqrt{\frac{\epsilon_1}{\epsilon_1 + \epsilon_2}} (-C_1 + C_2)) + \Delta(|C_1|) \right\}^2} \\ &+ \Theta(C_1) \frac{1}{\{\Delta(|C_1|)\}^2} \end{aligned} \quad (B.30)$$

logo

$$K^\pm = (K_1^\pm)^2 + (K_2^\pm)^2 \quad (B.31)$$

com

$$\begin{aligned} K_1^\pm &= \Theta(-C_1) \left\{ \frac{e^{-\frac{1}{2} \frac{\epsilon_2}{\epsilon_1 + \epsilon_2} (|C_1| + C_2)^2} H(-\bar{\zeta}_-) \pm e^{-\frac{1}{2} \frac{\epsilon_2}{\epsilon_1 + \epsilon_2} (|C_1| - C_2)^2} H(-\bar{\zeta}_+)}{H(\sqrt{\frac{\epsilon_1}{\epsilon_1 + \epsilon_2}} (C_1 + C_2)) - H(\sqrt{\frac{\epsilon_1}{\epsilon_1 + \epsilon_2}} (-C_1 + C_2)) + \Delta(|C_1|)} \right\} \\ K_2^\pm &= \Theta(C_1) \left\{ \frac{e^{-\frac{1}{2} \frac{\epsilon_2}{\epsilon_1 + \epsilon_2} (|C_1| + C_2)^2} H(\bar{\zeta}_-) \pm e^{-\frac{1}{2} \frac{\epsilon_2}{\epsilon_1 + \epsilon_2} (|C_1| - C_2)^2} H(\bar{\zeta}_+)}{\Delta(|C_1|)} \right\} \end{aligned} \quad (B.32)$$

expandindo a função $H(x)$ para grandes argumentos

$$H(x) \simeq \Theta(-x) + \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi x}}$$

e substituindo em K_1^\pm resulta

$$K_1^\pm = \Theta(-C_1)[\Theta(-|C_1|+C_2) + \Theta(-|C_1|-C_2)] \times \left(\frac{e^{-\frac{1}{2}\frac{\epsilon_2}{\epsilon_1+\epsilon_2}(|C_1|-C_2)^2}\Theta(\bar{\zeta}_+) \pm e^{-\frac{1}{2}\frac{\epsilon_2}{\epsilon_1+\epsilon_2}(|C_1|+C_2)^2}\Theta(\bar{\zeta}_-)}{e^{-\frac{1}{2}\frac{\epsilon_2}{\epsilon_1+\epsilon_2}(|C_1|-C_2)^2} - e^{-\frac{1}{2}\frac{\epsilon_2}{\epsilon_1+\epsilon_2}(|C_1|+C_2)^2} + \Delta(|C_1|)} \right) \quad (\text{B.33})$$

observamos que

$$\Theta(-|C_1|+C_2)\Theta(\bar{\zeta}_+) = \Theta(-|C_1|+C_2)\Theta(\epsilon_2|C_1|+C_2) = \Theta(-|C_1|+C_2)$$

$$\Theta(-|C_1|-C_2)\Theta(\bar{\zeta}_+) = \Theta(-|C_1|-C_2)\Theta(\epsilon_2|C_1|+C_2) = 0$$

$$\Theta(-|C_1|+C_2)\Theta(\bar{\zeta}_-) = \Theta(-|C_1|+C_2)\Theta(\epsilon_2|C_1|-C_2) = 0$$

$$\Theta(-|C_1|-C_2)\Theta(\bar{\zeta}_-) = \Theta(-|C_1|+C_2)\Theta(\epsilon_2|C_1|-C_2) = \Theta(-|C_1|-C_2) \quad (\text{B.34})$$

substituindo a expressão de $\Delta(|C_1|)$

$$K_1^\pm = \frac{\Theta(-C_1)\Theta(-|C_1|+C_2)\Theta(-\bar{\zeta}_-)\Theta(\bar{\zeta}_+)e^{-\frac{1}{2}\bar{x}^2}}{-\frac{e^{-\frac{1}{2}\bar{x}^2}}{\sqrt{2\pi\bar{x}_-}} - \frac{e^{-\frac{1}{2}\bar{x}_+^2}}{\sqrt{2\pi\bar{x}_+}} + \sqrt{\frac{\epsilon_1}{2\pi(\epsilon_1+\epsilon_2)}} \left\{ \frac{e^{-\frac{1}{2}\bar{x}^2}H(\bar{\zeta}_+)}{|C_1|} + \frac{e^{-\frac{1}{2}\bar{x}_+^2}H(\bar{\zeta}_-)}{|C_1|-\sqrt{\frac{\epsilon_2}{\epsilon_1+\epsilon_2}}\bar{\zeta}_-} \right\}} \pm$$

$$\frac{\Theta(-C_1)\Theta(-|C_1|-C_2)\Theta(\bar{\zeta}_-)\Theta(-\bar{\zeta}_+)e^{-\frac{1}{2}\bar{x}_+^2}}{-\frac{e^{-\frac{1}{2}\bar{x}_-^2}}{\sqrt{2\pi\bar{x}_-}} - \frac{e^{-\frac{1}{2}\bar{x}_+^2}}{\sqrt{2\pi\bar{x}_+}} + \sqrt{\frac{\epsilon_1}{2\pi(\epsilon_1+\epsilon_2)}} \left\{ \frac{e^{-\frac{1}{2}\bar{x}_-^2}H(\bar{\zeta}_+)}{|C_1|-\sqrt{\frac{\epsilon_2}{\epsilon_1+\epsilon_2}}\bar{\zeta}_+} + \frac{e^{-\frac{1}{2}\bar{x}_+^2}H(\bar{\zeta}_-)}{|C_1|} \right\}} \quad (\text{B.35})$$

- No primeiro termo : $C_2 > 0 \implies ||C_1| + C_2| > ||C_1| - C_2| \implies e^{-\frac{1}{2}x_+^2}/e^{-\frac{1}{2}x_-^2} \rightarrow 0$
 os termos em $e^{-\frac{1}{2}x_-^2}$ são dominantes e $e^{-\frac{1}{2}x_-^2} \gg e^{-\frac{1}{2}x_-^2}H(\bar{x}_+)$ já que $\bar{x}_+ > 0$.

- No segundo termo : $C_2 < 0 \implies ||C_1| + C_2| > ||C_1| - C_2| \implies e^{-\frac{1}{2}x_-^2}/e^{-\frac{1}{2}x_+^2} \rightarrow 0$
 e $e^{-\frac{1}{2}x_+^2}$ são dominantes e $e^{-\frac{1}{2}x_+^2} \gg e^{-\frac{1}{2}x_+^2}H(\bar{x}_-)$.

Logo

$$K_1^\pm = \Theta(-C_1)\Theta(-|C_1|+C_2)\sqrt{\frac{2\pi\epsilon_1}{\epsilon_1+\epsilon_2}}(-|C_1|+C_2) \quad (\text{B.36})$$

$$\pm\Theta(-C_1)\Theta(-|C_1|-C_2)\sqrt{\frac{2\pi\epsilon_1}{\epsilon_1+\epsilon_2}}(-|C_1|-C_2) \quad (\text{B.37})$$

A obtenção de K_2^\pm segue a mesma rotina, de modo que somente apresentamos o resultado final:

$$\begin{aligned} K_2^+ &= \Theta(C_1)\Theta(-\epsilon_2|C_1|+\epsilon_1C_2)\sqrt{\frac{2\pi\epsilon_1}{\epsilon_1+\epsilon_2}}(|C_1|+C_2) \\ &+ \Theta(C_1)\Theta(-\epsilon_2|C_1|-\epsilon_1C_2) + \sqrt{\frac{2\pi\epsilon_1}{\epsilon_1+\epsilon_2}}(|C_1|-C_2) \\ &+ \Theta(C_1)\Theta(\epsilon_2|C_1|+\epsilon_1C_2)\Theta(\epsilon_2|C_1|-\epsilon_1C_2)\sqrt{\frac{2\pi(\epsilon_1+\epsilon_2)}{\epsilon_1}}|C_1| \quad (\text{B.38}) \end{aligned}$$

$$\begin{aligned} K_2^- &= \Theta(C_1)\Theta(-\epsilon_2|C_1|+\epsilon_1C_2)\sqrt{\frac{2\pi\epsilon_1}{\epsilon_1+\epsilon_2}}(|C_1|+C_2) \\ &+ \Theta(C_1)\Theta(-\epsilon_2|C_1|-\epsilon_1C_2) + \sqrt{\frac{2\pi\epsilon_1}{\epsilon_1+\epsilon_2}}(|C_1|-C_2) \end{aligned}$$

$$+ \Theta(C_1) \Theta(\epsilon_2 | C_1 | +\epsilon_1 C_2) \Theta(\epsilon_2 | C_1 | -\epsilon_1 C_2) \sqrt{\frac{2\pi(\epsilon_1 + \epsilon_2)}{e_2}} C_2 \quad (\text{B.39})$$

Assim voltando finalmente a K^\pm temos

$$\begin{aligned} K^+ &= \frac{2\pi\epsilon_1}{\epsilon_1 + \epsilon_2} \left\{ \Theta(-C_1) \Theta(-|C_1| + C_2) (|C_1| - C_2)^2 \right. \\ &\quad + \Theta(-C_1) \Theta(-|C_1| - C_2) (|C_1| + C_2)^2 \\ &\quad + \Theta(C_1) \Theta\left(-\frac{\epsilon_2}{\epsilon_1} | C_1 | + C_2\right) (| C_1 | + C_2)^2 \\ &\quad \left. + \Theta(C_1) \Theta\left(-\frac{\epsilon_2}{\epsilon_1} | C_1 | - C_2\right) (| C_1 | - C_2)^2 \right\} \\ &\quad + \frac{2\pi(\epsilon_1 + \epsilon_2)}{e_1} \Theta(C_1) \Theta\left(\frac{\epsilon_2}{\epsilon_1} | C_1 | + C_2\right) \Theta\left(\frac{\epsilon_2}{\epsilon_1} | C_1 | - C_2\right) (C_1)^2 \quad (\text{B.40}) \end{aligned}$$

$$\begin{aligned} K^- &= \frac{2\pi\epsilon_1}{\epsilon_1 + \epsilon_2} \left\{ \Theta(-C_1) \Theta(-|C_1| + C_2) (|C_1| - C_2)^2 \right. \\ &\quad + \Theta(-C_1) \Theta(-|C_1| - C_2) (|C_1| + C_2)^2 \\ &\quad + \Theta(C_1) \Theta\left(-\frac{\epsilon_2}{\epsilon_1} | C_1 | + C_2\right) (| C_1 | + C_2)^2 \\ &\quad \left. + \Theta(C_1) \Theta\left(-\frac{\epsilon_2}{\epsilon_1} | C_1 | - C_2\right) (| C_1 | - C_2)^2 \right\} \\ &\quad + \frac{2\pi\epsilon_1(\epsilon_1 + \epsilon_2)}{e_2^2} \Theta(C_1) \Theta\left(\frac{\epsilon_2}{\epsilon_1} | C_1 | + C_2\right) \Theta\left(\frac{\epsilon_2}{\epsilon_1} | C_1 | - C_2\right) (C_1)^2 \quad (\text{B.41}) \end{aligned}$$

Voltando às equações de ponto de sela:

$$A_1^2 = \frac{\epsilon_1^2}{\epsilon_1 + \epsilon_2} \frac{\alpha}{\pi} \int Dz_1 Dz_2 K^+ \quad (\text{B.42})$$

$$A_2^2 = \frac{\epsilon_2^2}{\epsilon_1 + \epsilon_2} \frac{\alpha}{\pi} \int Dz_1 Dz_2 K^- \quad (\text{B.43})$$

$$(\text{B.44})$$

Relembrando que

$$C_1 = \frac{1}{\sqrt{\epsilon_1}} \left(\frac{\kappa_1 + \kappa_2}{\sqrt{2}} + \sqrt{A_1^2} z_1 \right) \quad (\text{B.45})$$

$$C_2 = \frac{1}{\sqrt{\epsilon_1}} \left(\frac{\kappa_1 - \kappa_2}{\sqrt{2}} + \sqrt{A_1^2} z_1 \right) \quad (\text{B.46})$$

e supondo que $\kappa_1 = \kappa_2 = 0$ podemos escrever (B.42) como:

$$A_1^2 = \frac{2\epsilon_1^2}{\epsilon_1 + \epsilon_2} \alpha \left\{ \frac{2}{\epsilon_1 + \epsilon_2} (S_1 + S_2) + \frac{\epsilon_1 + \epsilon_2}{\epsilon_1^2} S_3^+ \right\} \quad (\text{B.47})$$

$$A_2^2 = \frac{2\epsilon_2^2}{\epsilon_1 + \epsilon_2} \alpha \left\{ \frac{2}{\epsilon_1 + \epsilon_2} (S_1 + S_2) + \frac{\epsilon_1 + \epsilon_2}{\epsilon_1^2} S_3^- \right\} \quad (\text{B.48})$$

com

$$S_1 = A_1^2 \int Dz_1 Dz_2 \Theta(-z_1) \Theta(z_1 + \frac{A_2}{A_1} z_2) (z_1 + \frac{A_2}{A_1} z_2)^2, \quad (\text{B.49})$$

$$S_2 = A_1^2 \int Dz_1 Dz_2 \Theta(z_1) \Theta(-z_1 + \frac{A_2 \epsilon_1}{A_1 \epsilon_2} z_2) (z_1 + \frac{A_2}{A_1} z_2)^2, \quad (\text{B.50})$$

$$S_3^+ = A_1^2 \int Dz_1 Dz_2 \Theta(z_1) \Theta(z_1 - \frac{A_2 \epsilon_1}{A_1 \epsilon_2} z_2) \Theta(z_1 + \frac{A_2 \epsilon_1}{A_1 \epsilon_2} z_2) (z_1)^2, \quad (\text{B.51})$$

$$S_3^- = A_2^2 \int Dz_1 Dz_2 \Theta(z_1) \Theta(z_1 - \frac{A_2 \epsilon_1}{A_1 \epsilon_2} z_2) \Theta(z_1 + \frac{A_2 \epsilon_1}{A_1 \epsilon_2} z_2) (z_2)^2. \quad (\text{B.52})$$

Se trocarmos de variável $z_2 = z_1 t$ e integrarmos em z_1 resulta

$$\begin{aligned} S_1 &= A_1^2 \int_{-\infty}^{\infty} \frac{dt}{\sqrt{2\pi}} \Theta\left(\frac{A_2}{A_1}t - 1\right) \left(\frac{A_2}{A_1}t - 1\right)^2 \int_{-\infty}^{\infty} Dz_1 e^{-\frac{1}{2}z_1^2 t^2} z_1^3 \\ &= \frac{A_2^2}{\pi} \int_{\frac{A_1}{A_2}}^{\infty} dt \left(\frac{t - A_1/A_2}{t^2 + 1}\right)^2, \end{aligned} \quad (\text{B.53})$$

$$S_2 = \frac{A_2^2}{\pi} \int_{\frac{A_1 \epsilon_2}{A_2 \epsilon_1}}^{\infty} dt \left(\frac{t + A_1/A_2}{t^2 + 1}\right)^2, \quad (\text{B.54})$$

$$S_3^+ = \frac{2A_1^2}{\pi} \int_0^{\frac{A_1 \epsilon_2}{A_2 \epsilon_1}} dt \left(\frac{1}{t^2 + 1}\right)^2, \quad (\text{B.55})$$

$$S_3^- = \frac{2A_1^2 \epsilon_1^2}{\pi \epsilon_2^2} \int_0^{\frac{A_1 \epsilon_2}{A_2 \epsilon_1}} dt \left(\frac{t}{t^2 + 1}\right)^2. \quad (\text{B.56})$$

Reagrupando as integrais em (B.47) obtemos

$$\begin{aligned} A_1^2 &= \frac{2\epsilon_1^2}{\epsilon_1 + \epsilon_2} \alpha \left\{ \frac{2}{\epsilon_1 + \epsilon_2} \left(\frac{A_2^2}{\pi} \int_{-\frac{A_2}{A_1}}^{\frac{A_2 \epsilon_1}{A_1 \epsilon_2}} dt \left(\frac{t - A_2/A_1}{t^2 + 1}\right)^2 \right) \right. \\ &\quad \left. + \frac{\epsilon_1 + \epsilon_2}{\epsilon_1^2} \frac{2A_1^2}{\pi} \int_0^{\frac{A_1 \epsilon_2}{A_2 \epsilon_1}} dt \left(\frac{1}{t^2 + 1}\right)^2 \right\} \end{aligned} \quad (\text{B.57})$$

$$\begin{aligned} A_2^2 &= \frac{2\epsilon_2^2}{\epsilon_1 + \epsilon_2} \alpha \left\{ \frac{2}{\epsilon_1 + \epsilon_2} \left(\frac{A_2^2}{\pi} \int_{-\frac{A_2}{A_1}}^{\frac{A_2 \epsilon_1}{A_1 \epsilon_2}} dt \left(\frac{t - A_2/A_1}{t^2 + 1}\right)^2 \right) \right. \\ &\quad \left. + \frac{\epsilon_1 + \epsilon_2}{\epsilon_1^2} \frac{2A_1^2}{\pi} \int_0^{\frac{A_1 \epsilon_2}{A_2 \epsilon_1}} dt \left(\frac{t}{t^2 + 1}\right)^2 \right\} \end{aligned} \quad (\text{B.58})$$

Referências Bibliográficas

- [1] *Bases Biológicas do Comportamento*, Ed. Steven Rose, Ed UnB (1973)1981.
- [2] TOULOUSE, G., *J. Phys. A: Math. Gen.*, **A22**, p. 1959-1968, 1989.
- [3] McCULLOCH, W. S. e PITTS, W. A., *Bull. Math. Biophys.*, **5**, p. 115, 1943.
- [4] TOULOUSE, G., *Bull. Soc. Fr. Phys.*, p. 8-11, janeiro 1989.
- [5] DOMANY, E., *J. Stat. Phys.*, **51**.(5/6), p. 743-775, 1988.
- [6] Gyorgyi,G. e Tishby, N., *Proceed. STATPHYS 17 Workshop on Neural Networks and Spin Glass*,Ed W. K. Theumann e R. Koberle, World Scientific, 1990.
- [7] KINZEL, W., *Z. Phys. B*, **60**, p. 205-213, 1985.
- [8] HOPFIELD, J. J., *Proc. Natl. Acad. Sci. USA* **79**, p. 2254-2558, 1982.
- [9] HEBB, D. O., *The Organization of Behavior*, Wiley, New York, 1949.
- [10] MEZARD, M., PARISI, G. e VIRASORO, M. A., *Spin Glass Theory and Beyond - World Scientific Notes in Physics V.9*, World Scientific Publishing Co.,Singapura, 1987.
- [11] MEZARD, M., TOULOUSE, G. *La Recherche*, **22**, p. 616-623, 1991.
- [12] FONTANARI, J. F., KOBERLE, R., *Phys. Rev. A*, **36**(5), p.2475-2477, 1987.
- [13] LITTLE, W. A., *Math. Biosciences*. **19**, p. 101-120, 1974.

- [14] AMIT, D., GUTFREUND, H. e SOMPOLINSKY, H. , *Ann. Phys.*, **173**, p. 30-67, 1987.
- [15] AMIT, D., GUTFREUND, H. e SOMPOLINSKY, H., *Phys. Rev. A*, **32**(2), p. 1007-1018, 1985.
- [16] FONTANARI, J. F., THEUMANN, W. K., *J. Physique*, **51**, p.375, 1990.
- [17] PERSONNAZ, L., GUYON, I. e DREYFUS, G., *J. Physique Lett.*, **46**, p. L359-L365, 1985.
- [18] PARGA, N. e VIRASORO, M. A., *J. Physique*, **47**, p. 1857-1864, 1986.
- [19] ANDERSON, J. A., *IEEE Trans. Sys., Man, Cyb.*, **SMC-13**(5), p. 799-815, 1983.
- [20] SHERRINGTON, D., KIRKPATRICK, S., *Phy. Rev. Lett.*, **35** (26), p.1792-1796, 1975.
- [21] DOTSENKO, V., *Physica A*, **140**, p. 410-415, 1986.
- [22] GUTFREUND, H., *Physica A*, **163**, p. 373-385, 1990.
- [23] MEZARD, M., VIRASORO, M. A. *J. Physique*, **46**, p. 1293-1307, 1985.
- [24] FEIGELMAN, M. V. e IOFFE, L. B., *Int. J. Mod. Phys. B*, **1**, p. 51-68, 1987.
- [25] GUTFREUND, H., *Phys. Rev. A*, **37**, p. 570-577, 1988.
- [26] AMIT, D., GUTFREUND, H. e SOMPOLINSKY, H., *Phys. Rev. A*, **35**(5), p. 2293-2303, 1987.
- [27] DERRIDA, B., GARDNER, E. e ZIPPELIUS, A., *Europhys. Lett.*, **4**(2), p. 167-173, 1987.
- [28] CANNING, A., GARDNER, E., *J. Phys. A : Math. Gen.*, **21**, p. 3275-3284, 1988.

- [29] SUTTON, J. P., BEIS, J. S., e TRAINOR, L. E. H., *J. Phys. A : Math. Gen.*, **21**, p. 4443-4454, 1988.
- [30] SZENTÁGOTHAJ, J. *Proc. R. Soc. Lond. B*, **201**, p.219-248, 1978.
- [31] NOEST, A. J., *Phys. Rev. Lett.*, **63** (13), p. 1739-1742, 1989.
- [32] AMIT, D., *Modelling Brain Function*. Cambridge University Press, 1989.
- [33] ANDERSON, P. W., *Physica A*, **140**, p. 405-409, 1986.
- [34] DOTSENKO, V., *J. Phys. C : Sol. St. Phys.*, **18**, p.L1017-L1022, 1985.
- [35] DYSON, F. *Commun. Math. Phys.*, **12**, p. 91-107, 1969.
- [36] NADAL, J. P., TOULOUSE, G., CHANGEUX, J. P. e DEHAENE, S., *Europhys. Lett.*, **1** (10), p. 535-542, 1986.
- [37] RAMMAL, R., TOULOUSE, G. e VIRASORO, M. A., *Rev. of Mod. Phys.*, **58**(3), p. 765-788, 1986.
- [38] PARISI, G., *J. Phys. A : Math. Gen.*, **13**, p. 1101-1112, 1980.
- [39] AMIT, D., *Field Theory, The Renormalization Group, and The Critical Phenomena*. McGraw - Hill, 1978.
- [40] COOLEN, A. C. C. *preprint*, Statistical Mech. of Neural Networks - Lecture Notes in Physics Ed L. Garrido, Springer Verlag, 1990.
- [41] VIANA, L. *J. Physique*, **49**, p.167-174, 1988.
- [42] LAUTRUP, B., *preprint*, The Niels Bohr Institute, NBI-HE-88-06, 1988.
- [43] ABRIKOSOV, A. A., GORKOV L. P., DZIALOSHINSKI, I. E., *Methods of Quantum Field Theory in Statistical Physics*. Englewood Cliffs, Prentice Hall, 1964.
- [44] GARDNER, E., *J. Phys. A : Math. Gen.*, **21**, p. 257-270, 1988.

- [45] GARDNER, E., DERRIDA, B. *J. Phys. A : Math. Gen.*, **21**, p. 217-284, 1988.
- [46] KEPLER, T. B., e ABBOTT, L. F. *J. Phys. France*, **49**, p. 1657-1662, 1988.
- [47] MEZARD, M., NADAL, J. P. e TOULOUSE, G., *J. Physique*, **47**, p. 1457-1462, 1986.
- [48] DE ALMEIDA, J. R. L. e THOULESS, D. J. *J. Phys. A : Math. Gen.*, **11** (5), p. 983-990, 1978.
- [49] IDIART, M. A. P. e THEUMANN, A., *Proceed. STATPHYS 17 Workshop on Neural Networks and Spin Glass*, Ed W. K. Theumann e R. Koberle, World Scientific, 1990.
- [50] IDIART, M. A. P. e THEUMANN, A., *J. Phys. A : Math. Gen.*, **24**, p. L649-L658, 1991.
- [51] IDIART, M. A. P. e THEUMANN, A., a ser publicado no *J. Phys. A : Math. Gen.*.
- [52] IDIART, M. A. P. e THEUMANN, A., em preparação.