

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE FÍSICA

O Modelo RS para Redes de Neurônios[†]

Jeferson Jacob Arenzon

Dissertação realizada sob a orientação dos Profs. José Roberto Iglesias e Rita Maria Cunha de Almeida, apresentada ao Instituto de Física da UFRGS em preenchimento parcial dos requisitos para a obtenção do título de Mestre em Física.

Porto Alegre

1991

[†]Trabalho parcialmente financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Financiadora de Estudos e Projetos (FINEP) e Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS).

À Claudinha
e aos meus pais,
Bernardo e Sara Léa

AGRADECIMENTOS

- À Rita M. C. de Almeida e José R. Iglesias pela orientação e amizade.
- Ao Luciano Sclovsky por me introduzir ao Pascal e pela ajuda com os programas.
- À Andrea P. Nunes pelas figuras.
- Ao Thadeu J. P. Penna e Paulo Murilo C. de Oliveira pela colaboração e discussões.
- Ao Fernando M. Steffens, José Carlos M. Mombach, Marco A. P. Idiart e Oswaldo Baumgarten Filho pelas discussões e críticas ao trabalho.
- Ao pessoal da M202, Anderson Beatrici, Clarissa S. Barth, Flávio Elizalde e Maria Clara L. Adams pela companhia nestes dois anos.
- Ao Antônio “Nica” S. Duarte, Carlos A. Miraglia, Carlos H. Fraeb, Jorge A. Quillfeldt, Luciano “Sky” Sclovsky, Luis F. Neumann, Marco A. Soares, Teófilo D. R. de Bastos e à Sociedade Astronômica Rio-Grandense (SARG) pelos anos de discussões.
- À Maria Cecília do Amaral e à Ana Catarina D. Schier pela enorme paciência.

RESUMO

Um modelo para redes de neurônios é estudado onde a função energia E é dada pelo produto do quadrado das distâncias no espaço de fase entre o estado \vec{S} da rede e os P padrões armazenados $\vec{\xi}^\mu$, $\mu = 1, \dots, P$. Esta função energia é não negativa e é zero se $\vec{S} = \vec{\xi}^\mu$, $\forall \mu \leq P$. Obtivemos as propriedades de equilíbrio para padrões descorrelacionados em duas situações extremas: os casos PAS (padrões e anti-padrões armazenados) e OPS (somente padrões). A configuração PAS recupera o modelo de Hopfield no limite de baixo parâmetro de carga α (definido como P/N). A configuração OPS, por outro lado, apresenta um perfil topográfico anisotrópico de energia no espaço de fase com a forma de uma grande bacia na hiper-superfície fechada de um hiper-cubo N -dimensional de tal modo que, quando o número de padrões armazenados P aumenta, independentemente do tamanho da rede, as bacias de atração fundem-se e o único estado estável é aquele que se superpõem simetricamente com todas as memórias. A isotropia adquirida pelo armazenamento das anti-memórias (configuração PAS) é fundamental para que o limite $P \rightarrow \infty$ deste modelo funcione. Nesta configuração, o modelo RS possui uma performance em muito superior a outros modelos, resultado que é confirmado pelas simulações numéricas, onde obtivemos a frequência $f(m_0)$ com que um padrão arbitrário é recuperado a partir de um estado inicial com superposição m_0 e estimamos os tamanhos das bacias de atração para diferentes atividades a . Para redes PAS com atividade $a = 0.5$, um limite para o parâmetro de carga não foi encontrado, mas para os outros casos ($a \neq 0.5$ ou configurações OPS) o tamanho relativo das bacias de atração pode se tornar muito pequeno. Os tempos de convergência e as correspondentes dispersões também foram estudados numericamente e estão relacionados ao número de estados espúrios ao redor de cada configuração armazenada. Ambas quantidades são pequenas na fase de recuperação, indicando a homogeneidade e suavidade do espaço de fase.

ABSTRACT

A model for neural networks is studied through a multineuron interaction energy function E given by the product of the squared distances in phase space between the state \vec{S} of the net and the P stored patterns $\vec{\xi}^\mu$, $\mu = 1, \dots, P$. This energy function is non-negative and is zero if $\vec{S} = \vec{\xi}^\mu$, for any $\mu \leq P$. We obtain the main quantities in mean-field, uncorrelated memories in two extreme situations: PAS (patterns and antipatterns stored) and OPS (only patterns) configurations. The PAS case recovers the Hopfield model in the limit of low load parameter α (defined as P/N). The OPS configuration, on the other hand, presents an anisotropic energy landscape in phase space with a bowl-like shape in the closed hypersurface of an N -dimensional hypercube such that when the number P of stored patterns increases, the basins of attraction merge and the only stable state is the symmetrically overlapping configuration. This indicates that the isotropy in the energy field defined over the phase space implied by the consideration of the antimemories (PAS configuration) seems to be essential for the $P \rightarrow \infty$ limit of this model to work. In this isotropic configuration the RS-model shows a significant enhancement in comparisons with other models, a result that is also confirmed by numerical simulations, where we obtained the relative frequency $f(m_0)$ that an arbitrary pattern is retrieved from an initial overlap m_0 and estimated the size of the basins of attraction for different activities a . For the $a = 0.5$, PAS nets a limit for the load parameter was not found, but for the other cases ($a \neq 0.5$ or OPS configuration) the relative size of the basins of attraction may become too small. Convergence times and the corresponding dispersions have also been studied numerically as parameters to measure the efficiency of this model. These quantities are related to the number of spurious states for each configuration of stored patterns. Both the average convergence time and the dispersions are small in the retrieval phase, indicating the homogeneity and smoothness of the phase space.

Índice

1	Introdução	7
2	Conceitos Gerais e Modelos de Redes de Neurônios	11
2.1	Redes de Neurônios Biológicas	12
2.2	O Ponto de Vista Físico	14
2.3	O Modelo de Hopfield	21
2.3.1	Caso P Finito	22
2.3.2	Caso α Finito	27
2.3.3	Redes de Baixa Atividade	29
2.3.4	Simulação Numérica	32
2.4	Modelos com Interações entre k Neurônios	36
2.5	Conclusão	42
3	O Modelo RS	44
3.1	Formulação do Modelo	44
3.2	Teoria de Campo Médio com Anti-memórias: Caso PAS	49
3.3	Teoria de Campo Médio Somente com Memórias: Caso OPS	58
3.4	Conclusão	69
4	Simulação Numérica do Modelo RS	70
4.1	Algoritmo	70
4.2	Bacias de Atração	71

Índice	6
4.3 Tempo de Convergência	78
4.4 Conclusão	80
5 Conclusões	83
5.1 Resumo dos resultados para o Modelo RS	83
5.2 Futuras extensões	86
A O Método do Ponto de Sela	88
B Cálculo dos Autovalores	92
C Cálculo das Médias	96
Referências Bibliográficas	97

Capítulo 1

Introdução

O sistema nervoso e suas partes integrantes foram e são intensamente estudados, fornecendo uma enorme quantidade de dados experimentais sobre sua anatomia e comportamento. À medida em que se coloca o problema da modelagem teórica de tais sistemas, é preciso distinguir quais destas informações são importantes e quais são essenciais, sendo essenciais aquelas sem as quais o comportamento do sistema é trivial e importantes as responsáveis pela aproximação dos modelos com o sistema real, mas cuja ausência não compromete o comportamento complexo apresentado. A própria distinção entre o que é importante e o que é essencial depende do funcionamento dos modelos e do grau desejado de aproximação com a realidade. A idéia é tentar capturar, em modelos simples, propriedades do cérebro tais como aprendizagem, armazenamento e recuperação de padrões, manipulação simbólica, pensamento criativo, etc. Hoje em dia, porém, a grande maioria dos modelos teóricos se atêm às tarefas de armazenamento e recuperação de informação.

Redes de neurônios têm chamado a atenção da comunidade científica há alguns anos. O assunto, claramente interdisciplinar, tem sido fonte de grande excitação com novas idéias e conceitos sendo constantemente gerados por cientistas em várias frentes: biólogos, especialistas em inteligência artificial e, mais recentemente, físicos. O grande bum [1] na Física ocorreu em 1982 com o artigo seminal do físico J.J. Hopfield [2], onde é feita a analogia entre um sistema formado por neurônios simplificados e a sua contraparte magnética, os vidros de spin. O preço pago pelo ganho em tratabilidade do sistema, oriundo das simplificações

introduzidas, é o afastamento da realidade. A partir do modelo original, à medida que este era entendido mais profundamente, o caminho de volta à realidade foi sendo novamente percorrido. Propriedades dos sistemas reais, que antes não eram tomadas em consideração, foram aos poucos ganhando espaço. O uso da mecânica estatística mostrou-se muito poderoso na análise de tais sistemas, permitindo estudar as propriedades estáticas e dinâmicas destas redes, as quais são formadas por um número muito grande de elementos simples, os neurônios.

O entusiasmo dos físicos não passou despercebido, como bem descreve Amit [3]: “Biólogos começaram a se perguntar se as propriedades antecipadas pelos físicos podiam de fato ser observadas e se forneciam guias teóricos úteis para a investigação empírica da atividade cerebral; cientistas de computação não desprezaram estas idéias como candidatas para o processamento paralelo; psicólogos e neurologistas têm esperado novas metáforas úteis para a interpretação de disfunções comportamentais;... e tecnólogos têm acrescentado, é claro, estes tipos de modelos à lista dos futuros produtos a serem industrializados e vendidos”.

A aplicação tecnológica das redes de neurônios é um campo de pesquisa que vem recebendo altos investimentos, tanto de pessoal quanto financeiros por parte das indústrias, principalmente aquelas com interesse em controle de qualidade automatizado, reconhecimento de padrões ou fala, etc. Podemos citar, como exemplo, o projeto CONNEX da British Telecom [4], o qual consiste em 27 sub-projetos que investigam as capacidades industriais de tais redes. Um destes projetos tenta viabilizar telefones com habilidade de transmitir imagens de vídeo. Atualmente, enviar uma tela completa por linhas telefônicas é um processo muito lento. A idéia é mandar inicialmente uma única imagem completa do usuário, seguida apenas dos detalhes daquelas porções da figura que mudam durante a conversação. Isto serviria como entrada para a rede, a qual em pouco tempo convergiria para a imagem completa. Outro exemplo é um detector de explosivos para aeroportos que vem sendo desenvolvido nos EUA pela Science Applications International Corporation com o aval da Federal Aviation Administration e que se baseia na análise do padrão produzido quando a bagagem é bombardeada com nêutrons [5]. Além destas, existem inúmeras outras

aplicações [6]-[10] e seu número é crescente.

É comum afirmar que o cérebro é uma espécie de computador (na acepção mais geral da palavra) e para entender o seu funcionamento é necessário compreender os princípios computacionais que o sistema nervoso utiliza, ou seja, devemos entender como as células interagem para representar, transformar e armazenar informações. Como o sistema nervoso, os modelos teóricos necessitam ser, entre outros aspectos, robustos (a qualidade da recuperação não deve ser severamente afetada quando uma porção da rede é danificada) e associativos, como é explicado no capítulo seguinte. Embora o modelamento de sistemas nervosos esteja ainda em sua infância e não haja ainda qualquer modelo que explique exatamente como vemos e aprendemos, há uma crescente convicção de que as atuais linhas de pesquisa estão convergindo para as respostas.

O objetivo do trabalho desenvolvido nesta dissertação é apresentar, em detalhe, o modelo RS para redes de neurônios, o qual introduz algumas características importantes ao campo, tais como multi-sinapses assimétricas, e apresenta um alto grau de eficiência como memória associativa.

A dissertação está assim dividida: no capítulo 2 revisamos brevemente os pontos básicos da modelagem de redes de neurônios, como por exemplo, quais as representações para as características reais e qual a sua importância, quais os parâmetros utilizados para a descrição do comportamento do sistema, etc. Também neste capítulo é apresentado o mais estudado modelo de redes de neurônios, o modelo de Hopfield, e a sua solução analítica, realizada por Amit *et al* [11]-[13], juntamente com alguns resultados de simulação numérica. As generalizações deste modelo que envolvem interações entre mais de dois neurônios são rapidamente expostas.

Nos capítulos seguintes é introduzido o modelo RS e seus resultados, obtidos tanto por meios analíticos [14]-[16] quanto numéricos [17]-[18]. A origem do nome RS está no estado (Rio Grande do Sul) onde o modelo foi criado. Duas situações são extensamente estudadas: quando somente as memórias são armazenadas e quando tanto elas quanto as suas inversas são aprendidas. Estes dois casos se comportam distintamente e a isotropia introduzida no

espaço de fase pela inclusão das anti-memórias é essencial, como será mostrado, para o comportamento do sistema como memória associativa.

Finalmente, no capítulo 5, comparamos estes resultados com aqueles do modelo de Hopfield e concluímos.

Capítulo 2

Conceitos Gerais e Modelos de Redes de Neurônios

Redes de neurônios formais (simplificados) servem de paradigma para estudar o comportamento complexo de sistemas nervosos reais. Embora alguns tipos de cérebros sejam dotados de funções mentais de alto nível, a maioria dos modelos físicos existentes se preocupa em reproduzir as capacidades de aprendizado e recuperação de informação como um passo inicial em direção a descrição de processos mais elevados. Mesmo os atuais modelos estando afastados das suas contrapartes reais e as hipóteses sendo muitas e às vezes não realísticas, eles apresentam várias características que são comuns aos seus análogos biológicos e o seu comportamento está longe de ser considerado trivial. Neste capítulo vamos revisar alguns destes modelos e seus ingredientes principais. Na primeira seção apresentamos resumidamente alguns fatos básicos relacionados às redes biológicas. Na seção 2.2 introduzimos as simplificações que permeiam a grande maioria dos modelos físicos atuais e alguns conceitos fundamentais que permitem a utilização da Mecânica Estatística, enquanto que nas seções 2.3 e 2.4 o modelo de Hopfield e alguns de seus resultados são mostrados e certas alternativas são discutidas, com especial enfoque para as que envolvem interações entre mais de dois neurônios (multi-interação). Resumimos e concluimos na seção 2.5.

2.1 Redes de Neurônios Biológicas

O cérebro humano contém cerca de 10^{11} células nervosas ou neurônios [3,19]. Um neurônio típico é mostrado na figura 2.1, onde suas partes principais são identificadas: o corpo celular (soma), o axônio e os dendritos. As conexões entre diferentes neurônios são feitas por meio de sinapses, que são contatos entre os terminais do axônio de uma célula (pré-sináptica) e os dendritos (ou corpo celular) de outra (pós-sináptica), e cujo número, para um único neurônio, é da ordem de 10^4 (embora em algumas células possa chegar a 150.000). O sinal chega à célula através dos seus dendritos (processos aferentes), é processado no corpo celular e, se for o caso, é enviado através do axônio em direção aos outros neurônios (processos eferentes). O fato do corpo celular enviar ou não um sinal é determinado pela soma, ponderada pelas intensidades sinápticas, das atividades dos neurônios pré-sinápticos. As sinapses podem ser fracas ou fortes, excitatórias ou inibitórias (neste último caso, a atividade da célula pré-sináptica inibe o funcionamento da pós-sináptica) e o seu funcionamento está esquematizado na figura 2.2: a chegada de um sinal na membrana pré-sináptica causa a liberação de neurotransmissores que se difundem através da sinapse até atingir as moléculas receptoras na membrana pós-sináptica, as quais retransmitem o sinal. Os dendritos, por sua vez, formam uma estrutura altamente ramificada, como os galhos de uma árvore, sendo que o axônio usualmente também termina em um padrão semelhante.

Do mesmo modo que as sinapses modificam os potenciais de membrana dos dendritos, elas podem alterar os potenciais de outras sinapses. Um exemplo de tal situação é uma configuração onde existam ligações axônio-axônio, o que pode ser tomado como sendo uma conexão entre mais de dois neurônios, isto é, uma multi-sinapse. O papel de tais processos foi enfatizado nas referências [20]-[22] nos estudos de sistemas nervosos de invertebrados e é provável que tais junções sejam relevantes no funcionamento do sistema nervoso central de organismos superiores.

Uma sinapse pode alterar o sinal que está sendo transmitido através dela e, mais importante ainda, as propriedades das sinapses podem também ser modificadas pelos sinais

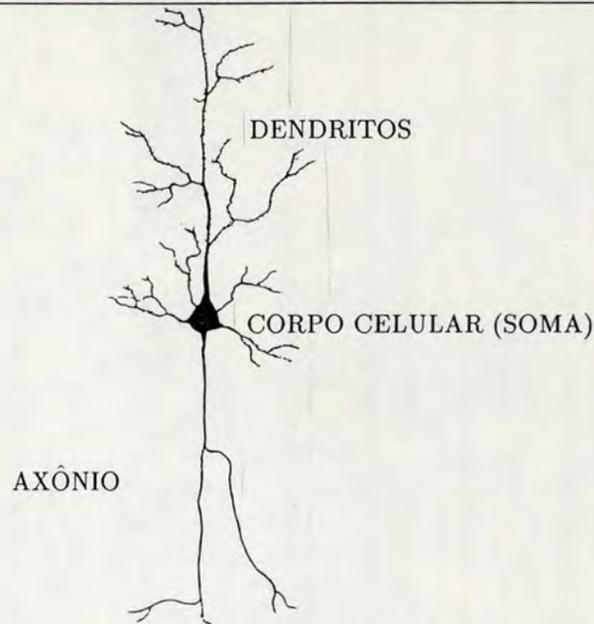


Figura 2.1: Célula piramidal típica encontrada no cérebro. A informação corre através dos dendritos até o corpo celular a partir do qual, se o potencial ultrapassar um determinado limiar, um pulso é enviado ao longo do axônio [19].

que as atravessam [23]. Esta plasticidade é o ponto básico do aprendizado: as intensidades sinápticas são criadas e/ou alteradas quando novos padrões são aprendidos.

O córtex cerebral, ou neocórtex como também é chamado, consiste de camadas de células com aproximadamente 3 mm de espessura, as quais são dobradas várias vezes para permitir sua acomodação no crânio [19]. Também há evidências de que os neurônios se agrupam em colunas com cerca de 0.5 mm de diâmetro, perpendiculares à superfície cortical.

Diferentemente da arquitetura dos atuais computadores, existem evidências de que no cérebro as memórias estão distribuídas pelas sinapses que compõem o sistema [3,19]. A recuperação de informação é associativa, ou auto-endereçada, e as memórias são recuperadas pela semelhança de seu conteúdo e não pelo endereço: a dinâmica da rede, ao ser colocada em um estado que se assemelha a um determinado padrão aprendido, se encarrega de corrigir os erros e convergir para aquela memória.

Por fim, resta falar da atividade do sistema nervoso. A atividade é uma medida da fração

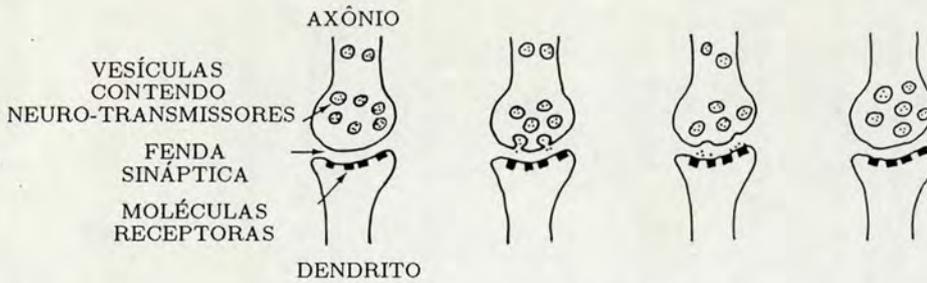


Figura 2.2: Esquema da passagem do pulso através da sinapse: neurotransmissores são lançados na fenda sináptica quando o potencial pós-sináptico ultrapassa um determinado limiar, e são captados na membrana pós-sináptica por moléculas receptoras [19].

dos neurônios que estão ativos (emitindo sinal). A média espacial do nível de atividade do córtex é relativamente baixa: nas regiões onde são detectadas funções associativas, a quantidade de neurônios ativos é de 4 – 5%.

Muitos outros aspectos do funcionamento do cérebro já foram estudados e a quantidade de dados gerada até hoje é enorme, ficando além dos objetivos desta dissertação analisar em detalhe a fisiologia e o comportamento de neurônios reais. Sendo assim, vamos passar a discutir o modelamento de tais sistemas.

2.2 O Ponto de Vista Físico

Em 1943, McCulloch e Pitts [24] mostraram que qualquer operação do cálculo proposicional pode ser efetuada por neurônios simplificados combinados em pequenas seqüências temporais. A atividade do neurônio de saída representa o valor de verdade da operação lógica efetuada pelos neurônios. Isto pode ser feito escolhendo-se adequadamente os valores das conexões sinápticas e dos limiares de cada componente. Estes neurônios formais são representados pela variável S_i , cujo valor pode ser +1 (ativo) e -1 (inativo) ou, equivalentemente, 1 e 0. Uma rede de neurônios formais é um conjunto formado por N destes elementos simples, altamente conectados, cujo estado será então associado a um vetor N -dimensional $\vec{S} = (S_1, \dots, S_N)$.

As propriedades das conexões entre dois neurônios i e j podem ser modeladas por um número real, J_{ij} . Uma conexão sináptica é excitatória quando $J_{ij} > 0$ e inibitória se $J_{ij} < 0$. Em 1949, Hebb [23] propôs que o sistema nervoso podia aprender e fazer associações através de modificações seletivas das conexões sinápticas. Isto se reflete, nos modelos físicos, na forma funcional de J_{ij} , que depende das memórias armazenadas.

Uma rede de neurônios é um sistema dinâmico, cuja evolução é governada por regras que dependem do modo com que um neurônio responde aos sinais que chegam até ele [3,25]-[32]. O potencial pós-sináptico no i -ésimo neurônio é a soma, ponderada pelas intensidades das conexões sinápticas, da atividade do resto da rede e, quando consideramos apenas interações entre dois neurônios, pode ser escrito como

$$h_i(t + 1) = \sum_{j \neq i} J_{ij} S_j(t) \quad , \quad (2.1)$$

o qual também é chamado de campo local. Além disso, cada célula possui o seu próprio limiar θ_i . Se o campo local ultrapassar o limiar, a probabilidade de que este neurônio envie um sinal é

$$Pr(S_i = +1) = f(h_i - \theta_i) \quad , \quad (2.2)$$

onde f é qualquer função cujo comportamento é análogo ao da figura 2.3. A largura da região que vai de 0 até 1 é uma medida da estocasticidade do processo ou, em outras palavras, do ruído na rede, o qual, daqui por diante, será denotado pela temperatura $T = \beta^{-1}$. As origens físicas deste ruído podem ser as mais variadas possíveis [33]: variações randômicas [1] no número e no tamanho dos pacotes de neurotransmissores químicos lançados nas sinapses, “vazamento” destes neurotransmissores (que pode acontecer mesmo se o neurônio estiver inativo), outras flutuações estocásticas que podem ocorrer ao longo do processo, etc.

A figura 2.4 apresenta esquematicamente a forma de um neurônio (representado pelo círculo grande). As linhas J_{ij} , $j \neq i$, são as sinapses que o conectam com os outros N neurônios (quadrados pequenos) da rede, h_i é o potencial pós-sináptico atuando em i e θ_i é o limiar. A função $f(h_i - \theta_i)$ determina o estado final S_i do neurônio.

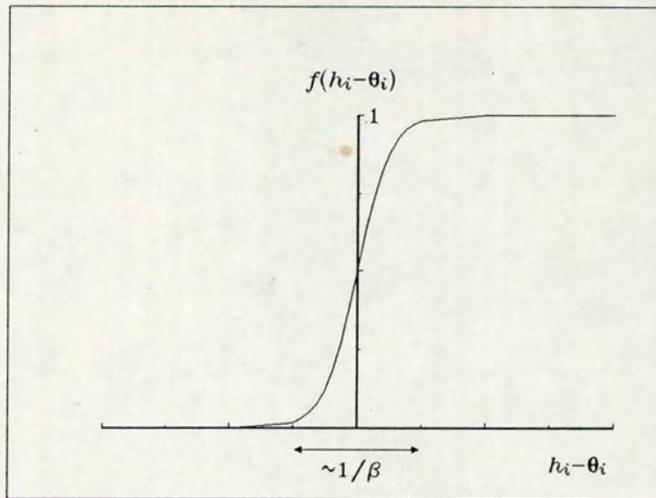


Figura 2.3: Probabilidade de ativação em função da diferença entre o campo local h_i e o limiar θ_i para o i -ésimo neurônio.

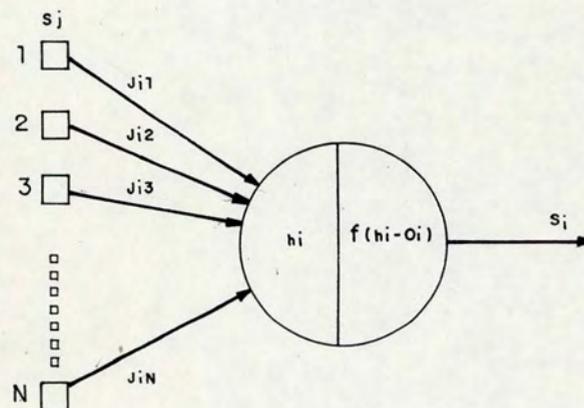


Figura 2.4: Estrutura lógica de um neurônio. As diferentes partes são explicadas no texto [3].

No limite de ruído nulo ($T = 0$), a dinâmica da rede obedece à seguinte regra determinística:

$$S_i(t+1) = \text{sgn} \left[\sum_{j \neq i} J_{ij} S_j(t) - \theta_i \right] , \quad (2.3)$$

onde a função sinal é definida como:

$$\text{sgn}(x) = \begin{cases} +1 & \text{se } x > 0 \\ -1 & \text{se } x < 0 \end{cases} . \quad (2.4)$$

A dinâmica ditada pela eq.(2.3) pode ser implementada em série (assíncrona) e em paralelo (síncrona). No esquema serial, em cada unidade de tempo, um neurônio é escolhido aleatoriamente e alinhado com o seu campo local, enquanto que na dinâmica em paralelo, em cada instante de tempo todos os neurônios são atualizados simultaneamente.

A eq.(2.3) pode levar a estados estáveis, $S_i(t+1) = S_i(t)$, $\forall i$, isto é, pontos fixos da dinâmica, os quais satisfazem

$$S_i(t) = \text{sgn} \left[\sum_{j \neq i} J_{ij} S_j(t) - \theta_i \right] , \quad (2.5)$$

ou seja, o i -ésimo neurônio está alinhado com o campo que atua sobre ele. Nas redes de neurônios com atratores ('attractor neural network' ou simplesmente ANN), a matriz J das conexões sinápticas é construída de tal modo que os elementos de um conjunto previamente escolhido de padrões (estados possíveis da rede), designados por $\vec{\xi}^\mu = (\xi_1^\mu, \dots, \xi_N^\mu)$, $\mu = 1, \dots, P$, sejam os pontos fixos da equação (2.5) e, portanto, satisfazem a

$$\xi_i^\mu = \text{sgn} \left[\sum_{j \neq i} J_{ij} \xi_j^\mu - \theta_i \right] , \quad (2.6)$$

ou seja, são atratores da dinâmica descrita por (2.3).

Estamos fundamentalmente interessados nas propriedades do sistema que são independentes da particular escolha dos padrões a serem armazenados e, portanto, podemos considerar as componentes ξ_i^μ dos P padrões como variáveis aleatórias independentes, cuja distribuição de probabilidade é dada por

$$P(\xi_i^\mu) = a\delta(\xi_i^\mu - 1) + (1 - a)\delta(\xi_i^\mu + 1) , \quad (2.7)$$

onde a é, quando $N \rightarrow \infty$, a fração de neurônios ativos na rede, o que é chamado de atividade. Boa parte do que será visto nas seções e capítulos seguintes considera redes com atividade 0.5, ou seja, temos o mesmo número de neurônios ativos e inativos. Esta restrição não é muito realista, uma vez que, como vimos anteriormente, a atividade do córtex é muito menor do que isso ou, no caso de sistemas montados com a finalidade de reconhecer padrões, boa parte das memórias armazenadas possui um pano de fundo comum (um bom exemplo são os CCD utilizados em astronomia, isto é, placas formadas por um arranjo retangular de píxels que registram os fótons incidentes, sendo que tais píxels, em sua maioria, são pontos negros).

A capacidade de armazenamento α da rede é medida usualmente pela razão entre o número P de memórias e a quantidade de neurônios da rede N :

$$\alpha = \frac{P}{N} .$$

A memória de uma ANN é associativa, ou auto-endereçada: partindo de um estado \vec{S} próximo a uma das memórias $\vec{\xi}^\mu$, o processo dinâmico (2.3) converge para o padrão armazenado. Isto acontece desde que o sistema se encontre numa região ao redor da memória $\vec{\xi}^\mu$, conhecida como bacia de atração. A medida de quão próximo um estado está da μ -ésima memória é dada pela superposição ('overlap') m_μ , definida como

$$m_\mu = \frac{1}{N} \sum_{i=1}^N \xi_i^\mu S_i , \quad (2.8)$$

o qual vale 1 se $\vec{S} = \vec{\xi}^\mu$, -1 se $\vec{S} = -\vec{\xi}^\mu$ e $1/\sqrt{N}$ se \vec{S} e $\vec{\xi}^\mu$ forem descorrelacionados. Podemos ainda relacionar m_μ com outra quantidade bastante utilizada: a distância de Hamming d_μ^H , que é simplesmente o número de elementos distintos nos vetores \vec{S} e $\vec{\xi}^\mu$. A relação entre ambas quantidades é

$$d_\mu^H = \frac{N}{2}(1 - m_\mu) , \quad (2.9)$$

onde N é o número de neurônios. A distância de Hamming mede diferenças, enquanto que a superposição mede similaridades.

Cada um dos 2^N possíveis estados da rede pode ser considerado como um dos vértices de um hiper-cubo de N dimensões, como mostra a figura 2.5 para o caso de $N = 3$ (2^3

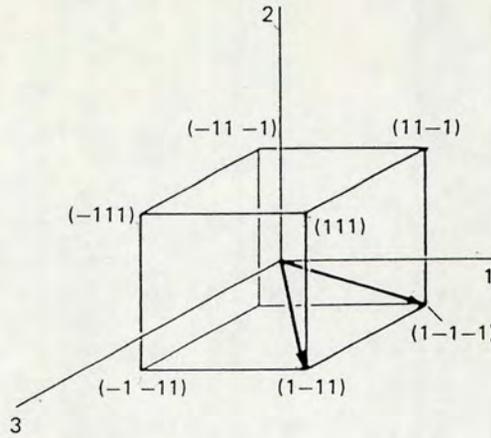


Figura 2.5: Representação geométrica do espaço de fase para uma rede com $N = 3$ neurônios. O número de estados (vértices) é 2^N [3].

vértices). O lado do cubo, isto é, a distância entre dois estados que diferem por apenas um neurônio, tem comprimento 2. A quantidade m_μ é o cosseno do ângulo formado pelos vetores que vão desde a origem até os vértices correspondentes aos dois estados.

As características apresentadas nesta seção são comuns a muitos modelos e, embora tenhamos nos detido em sistemas cujos neurônios interagem dois a dois, a generalização para os casos onde temos interações entre k neurônios é imediata. Neste caso, por exemplo, o campo local atuando no i_1 -ésimo neurônio pode ser dado por

$$h_{i_1} = \sum_{i_2 < \dots < i_k} J_{i_1 i_2 \dots i_k} S_{i_2} \dots S_{i_k} \quad ,$$

onde $J_{i_1 \dots i_k}$ é o termo de interação entre k neurônios. A forma esquemática destas interações pode ser vista na figura 2.6.

Vamos falar um pouco agora das médias envolvidas no processo [3,11,32]. A razão de se fazer médias vem da necessidade de obter resultados que sejam independentes do conjunto de memórias escolhido. Em termos práticos, isto pode significar, por exemplo, fazer médias sobre todos os possíveis conjuntos com P padrões, sujeitos ou não a alguma espécie de vínculo (atividade, correlação, etc). No caso que vamos tratar aqui, a saber

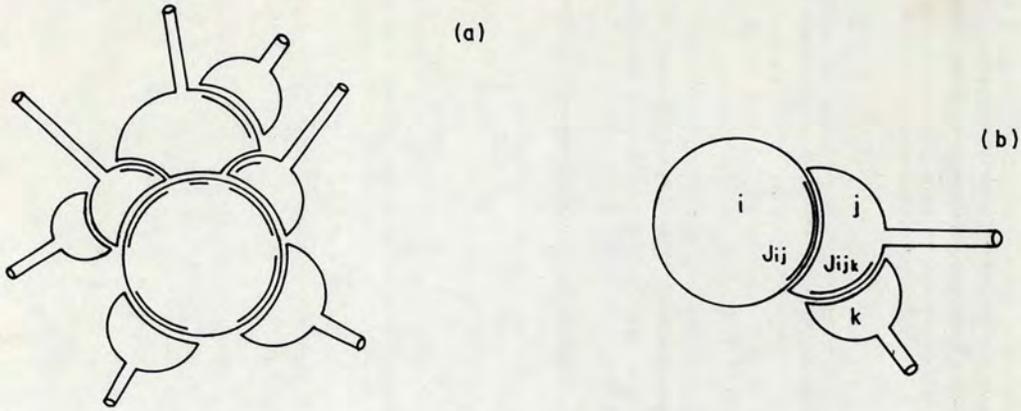


Figura 2.6: a) Forma esquemática dos contatos multi-sinápticos. b) Exemplo de interação entre 3 neurônios. Os índices i, j e k se referem aos neurônios, J_{ij} é a sinapse axônio-dendrito que liga um ramo do axônio do neurônio j a um dendrito de i . Por outro lado, J_{ijk} é uma sinapse do tipo axônio-axônio-dendrito que possibilita que a atividade do neurônio k altere a eficácia da sinapse entre i e j [34].

quando $P \ll N$, esta tarefa é extremamente simplificada por uma propriedade de conjuntos finitos de memórias: a auto-mediação. Vamos considerar, então, um tal conjunto:

$$\begin{aligned} \vec{\xi}^1 &= (\xi_1^1, \xi_2^1, \dots, \xi_i^1, \dots, \xi_N^1) \\ \vec{\xi}^2 &= (\xi_1^2, \xi_2^2, \dots, \xi_i^2, \dots, \xi_N^2) \\ &\vdots \\ \vec{\xi}^P &= (\xi_1^P, \xi_2^P, \dots, \xi_i^P, \dots, \xi_N^P) \end{aligned}$$

Para um-determinado sítio i , temos P possíveis valores ± 1 . Podemos formar então um hiper-cubo P -dimensional onde cada um dos vértices é uma das 2^P possíveis configurações do vetor $\xi_i = (\xi_i^1, \dots, \xi_i^P)$, isto é, ξ_i nos dá os valores do i -ésimo neurônio nas P memórias. Vamos supor agora que temos uma média sobre todos os neurônios (sítios) do sistema de uma função qualquer $f(\xi_i)$ dos vetores ξ_i :

$$\frac{1}{N} \sum_{i=1}^N f(\xi_i)$$

onde o número de memórias P (e portanto 2^P) é finito, i.e., $P \ll N$. Assim, cada um dos vértices do hiper-cubo P -dimensional será várias vezes visitado à medida que fazemos a soma em i , isto é, cada uma das possíveis realizações do vetor ξ_i aparecerá um grande número de vezes. Neste caso podemos reescrever a média sobre os sítios como uma média sobre os vértices do hiper-cubo:

$$\frac{1}{N} \sum_{i=1}^N f(\xi_i) = \sum_{\xi} \frac{N(\xi)}{N} f(\xi) \quad ,$$

onde $N(\xi)$ é o número de vezes que o vértice ξ é visitado. Quando $N \rightarrow \infty$, a frequência com que cada vértice aparece é exatamente a sua probabilidade, $1/2^P$, para padrões descorrelacionados. Portanto:

$$\frac{1}{N} \sum_{i=1}^N f(\xi_i) = \frac{1}{2^P} \sum_{\xi} f(\xi) \equiv \langle\langle f(\xi) \rangle\rangle \quad . \quad (2.10)$$

Isto significa que, se P é finito, qualquer conjunto particular de memórias se comporta exatamente como o caso médio, propriedade que recebe o nome de auto-mediação ('self-average'). Isto já não ocorre quando o número de padrões é da ordem de N ($P = \alpha N, \alpha \neq 0$), mas ainda é interessante que os resultados obtidos não dependam de um particular conjunto de memórias, isto é, qualquer conjunto típico deve dar os mesmos resultados. Devemos então efetuar explicitamente a média sobre as variáveis randômicas (as ξ 's), o que é feito utilizando-se o truque das réplicas ('replica trick') [35]-[37], que será discutido mais tarde.

2.3 O Modelo de Hopfield

O modelo de Hopfield considera que as interações ocorrem entre pares de neurônios e são simétricas, isto é

$$J_{ij} = J_{ji} \quad ,$$

o que significa que a influência do neurônio j no neurônio i é igual à influência do i no j . Esta suposição, do ponto de vista biológico, não é realística, pois os canais de comunicação nas redes de neurônios reais são unidirecionais e não há qualquer razão para supor tal

simetria. Apesar disso, o modelo apresenta características de armazenamento e recuperação de memória que o tornam de especial interesse. A forma particular de J_{ij} a ser considerada aqui, a qual leva o nome de regra de aprendizagem de Hebb, é

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

Normalmente assume-se que $J_{ii} = 0$, uma vez que os neurônios não costumam possuir sinapses consigo mesmo. Esta é uma regra local (o valor de J_{ij} depende somente de ξ_i^μ e ξ_j^μ), aditiva (como supõe-se ser o aprendizado) e todas as memórias são aprendidas igualmente (não há distinções entre memórias novas e antigas).

O modelo de Hopfield é descrito pela seguinte função energia

$$E = -\frac{1}{2} \sum_{\substack{i,j \\ j \neq i}} J_{ij} S_i S_j \quad , \quad (2.12)$$

onde J_{ij} é dado por (2.11). Substituindo (2.11) em (2.12), podemos reescrever a energia em termos das superposições m_μ :

$$E = -\frac{N}{2} \sum_{\mu=1}^P m_\mu^2 + \frac{P}{2} \quad , \quad (2.13)$$

onde a constante aditiva $P/2$ vem do termo de auto-interação.

Vamos analisar agora com algum detalhe o caso em que o número de memórias armazenadas é finito; em seguida apresentamos os principais resultados quando a rede está próxima da saturação ($P = \alpha N$) e os resultados numéricos mais significativos.

2.3.1 Caso P Finito

Vamos considerar agora a situação em que temos um conjunto finito de memórias ($P \ll N$), ou seja, o limite $\alpha \rightarrow 0$ [11]. Neste caso, como vimos anteriormente, as memórias se auto-mediam, simplificando consideravelmente os cálculos [3,11].

Vamos calcular a função partição do sistema:

$$Z = \sum_{\{\tilde{S}_i\}} e^{-\beta E[\tilde{S}]} \quad , \quad (2.14)$$

onde a soma é sobre todas as possíveis configurações do sistema. Usando (2.13) pode-se escrever:

$$Z = \sum_{\{S_i\}} e^{\frac{\beta N}{2} \sum_{\mu} m_{\mu}^2} .$$

Introduzindo a função delta resulta:

$$Z = \sum_{\{S_i\}} \int dm_1 \dots dm_P \left[\prod_{\mu} \delta \left(m_{\mu} - \frac{1}{N} \sum_i S_i \xi_i^{\mu} \right) \right] e^{\frac{\beta N}{2} \sum_{\mu} m_{\mu}^2} , \quad (2.15)$$

e utilizando a seguinte identidade

$$\delta \left(m_{\mu} - \frac{1}{N} \sum_{i=1}^N \xi_i^{\mu} S_i \right) = \frac{\beta N}{2\pi i} \int_{-i\infty}^{+i\infty} dt_{\mu} \exp \left[\beta N t_{\mu} \left(m_{\mu} - \frac{1}{N} \sum_{i=1}^N \xi_i^{\mu} S_i \right) \right] , \quad (2.16)$$

obtemos

$$Z = \left(\frac{\beta N}{2\pi i} \right)^P \int \mathbf{d}\mathbf{m} \int \mathbf{d}\mathbf{t} e^{-\beta N f(\mathbf{m}, \mathbf{t})} \quad (2.17)$$

onde $\mathbf{m} = (m_1, \dots, m_P)$, $\mathbf{t} = (t_1, \dots, t_P)$, $\mathbf{d}\mathbf{m} = dm_1 \dots dm_P$, $\mathbf{d}\mathbf{t} = dt_1 \dots dt_P$ e

$$f(\mathbf{m}, \mathbf{t}) = -\frac{1}{2} \sum_{\mu=1}^P m_{\mu}^2 + \sum_{\mu=1}^P m_{\mu} t_{\mu} - \frac{1}{\beta} \langle \langle \ln[2 \cosh(\beta \boldsymbol{\xi} \cdot \mathbf{t})] \rangle \rangle \quad (2.18)$$

é a energia livre por neurônio do sistema. Quando $N \rightarrow \infty$, a integral acima é dominada pelos mínimos da energia livre, os quais são obtidos pelo método do ponto de sela (Apêndice A). Derivando $f(\mathbf{m}, \mathbf{t})$ obtemos:

$$\frac{\partial f}{\partial m_{\mu}} = -m_{\mu} + t_{\mu}$$

$$\frac{\partial f}{\partial t_{\mu}} = m_{\mu} - \langle \langle \xi^{\mu} \tanh(\beta \boldsymbol{\xi} \cdot \mathbf{t}) \rangle \rangle .$$

As condições de extremo, isto é, que as variações de primeira ordem sejam nulas, fornecem as seguintes equações:

$$t_{\mu} = m_{\mu} \quad (2.19)$$

$$m_{\mu} = \langle \langle \xi^{\mu} \tanh(\beta \boldsymbol{\xi} \cdot \mathbf{t}) \rangle \rangle , \quad (2.20)$$

e, após substituir (2.19) em (2.20), resulta:

$$m_\mu = \langle\langle \xi^\mu \tanh(\beta \xi \cdot \mathbf{m}) \rangle\rangle \quad . \quad (2.21)$$

Encontrar todas as soluções deste conjunto de P equações não é uma tarefa simples. Um possível caminho é propor vários ‘ansätze’ e estudar seus intervalos de validade e estabilidade. Testamos então uma solução do seguinte tipo: suponhamos que existem, dentre as P memórias, n com as quais a rede possui uma superposição macroscópica ($m \gg 1/\sqrt{N}$). Sem perda de generalidade podemos escolher tais memórias como sendo as n primeiras, ou seja:

$$\mathbf{m} = (\underbrace{m, \dots, m}_n, \underbrace{0, \dots, 0}_{P-n}) \quad . \quad (2.22)$$

Substituindo a eq.(2.22) em (2.21) resulta:

$$m = \left\langle\left\langle \xi^\mu \tanh \left(\beta m \sum_{\nu=1}^n \xi^\nu \right) \right\rangle\right\rangle \quad , \quad \mu \leq n \quad , \quad (2.23)$$

a qual, somando-se desde $\mu = 1$ até n , fica:

$$m = \frac{1}{n} \langle\langle z_n \tanh(\beta m z_n) \rangle\rangle \quad , \quad (2.24)$$

onde

$$z_n = \sum_{\nu=1}^n \xi^\nu \quad . \quad (2.25)$$

Escrevendo a média $\langle\langle \rangle\rangle$ como uma média sobre as variáveis z_n , quando os padrões são decorrelacionados, temos (Apêndice C):

$$m = \frac{1}{n2^n} \sum_{l=0}^n \frac{n!}{l!(n-l)!} (n-2l) \tanh[\beta m(n-2l)] \quad . \quad (2.26)$$

Na figura 2.7 são mostradas as soluções da equação acima para diferentes valores de n . Note que acima de $T = 1$, ou seja, abaixo de $\beta = 1$, a única solução é a paramagnética, $m = 0$, a qual sempre é solução de (2.26).

Segundo o método do ponto de sela, precisamos calcular as variações de segunda ordem da energia livre para estudarmos a estabilidade destas soluções (Apêndice A). Calculando

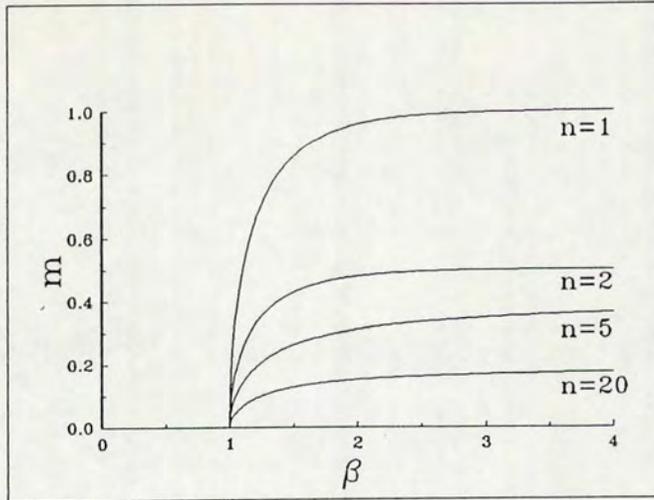


Figura 2.7: As soluções m versus β para vários valores de n .

os elementos $A_{\mu\nu}$ da matriz estabilidade obtemos:

$$A_{\mu\nu} = \frac{\partial^2 f}{\partial m_\mu \partial m_\nu}$$

os quais valem, depois de substituirmos a eq.(2.19) em (2.18):

$$A_{\mu\nu} = \delta_{\mu\nu} - \beta \langle\langle \xi^\mu \xi^\nu \text{sech}^2(\beta \xi \cdot m) \rangle\rangle \tag{2.27}$$

Os autovalores de uma matriz com tal estrutura são calculados no apêndice B e valem:

$$\lambda_1 = 1 + \beta \langle\langle (\xi^1 \xi^2 - 1) \text{sech}^2 \beta m z_n \rangle\rangle \quad , \tag{2.28}$$

que é $n - 1$ vezes degenerado e não existe se $n = 1$,

$$\lambda_2 = 1 - \beta \langle\langle \text{sech}^2 \beta m z_n \rangle\rangle \quad , \tag{2.29}$$

que é $P - n$ vezes degenerado e

$$\lambda_3 = 1 + \beta \langle\langle [(n - 1)\xi^1 \xi^2 - 1] \text{sech}^2 \beta m z_n \rangle\rangle \quad , \tag{2.30}$$

que é não degenerado. Nas equações (2.28),(2.30) consideramos, sem perda de generalidade, ξ^1 e ξ^2 , pois esta contribuição provêm dos termos com $\mu \neq \nu$. Para a solução paramagnética



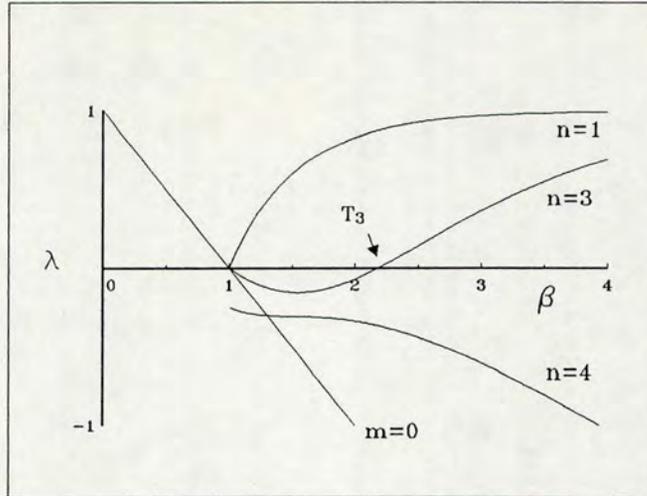


Figura 2.8: Autovalores para o modelo de Hopfield. A solução paramagnética é estável para $T > 1$, enquanto que as memórias próprias são estáveis somente abaixo desta temperatura. As soluções simétricas com n par são sempre instáveis embora para n ímpar elas se tornem estáveis abaixo da temperatura T_n .

($m = 0$) os três autovalores são iguais e valem $1 - \beta$, sendo esta solução estável, portanto, quando $T > 1$. Para as memórias próprias (estados de Mattis [38]), ou seja, $n = 1$ e $m \neq 0$, os autovalores se reduzem a

$$\lambda = 1 - \beta \operatorname{sech}^2 \beta m \tag{2.31}$$

que é positivo quando $T < 1$. Além destas duas soluções temos ainda as memórias simétricas ($n > 1$), cujo autovalor λ_1 é mostrado na figura 2.8 para alguns valores de n , juntamente com os autovalores do estado desordenado e das memórias próprias.

As memórias simétricas com n ímpar são instáveis em $T = 1$ e estáveis em $T = 0$, onde a estabilidade é atingida quando o autovalor λ_1 , o menor deles, muda de sinal. Quando n é par tais memórias nunca se tornam estáveis. A temperatura T_n , onde os estados simétricos com n ímpar deixam de ser estáveis, pode ser calculada numericamente. Por exemplo, para $n = 3$, $T_3 = 0.461$ e T_n diminui à medida que n cresce. Além destas soluções, o modelo de

Hopfield pode ainda apresentar estados com superposições assimétricas, como por exemplo

$$\mathbf{m} = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, 0, \dots, 0 \right) ,$$

que, particularmente, é estável a temperaturas abaixo de 0.18.

Podemos resumir o limite $\alpha \rightarrow 0$ do modelo de Hopfield da seguinte forma: a temperaturas altas ($T > 1$), somente o estado paramagnético ($m = 0$) é estável, sendo que à medida que a temperatura vai diminuindo abaixo de 1, estados que se superpõem igualmente com um número ímpar de memórias vão sucessivamente se tornando estáveis. O estado que representa uma superposição com somente um dos padrões (memórias próprias) é estável para todo valor de temperatura menor do que $T_c = 1$.

2.3.2 Caso α Finito

A auto-mediação das memórias vale somente quando o número P delas é finito. Portanto, para o caso em que P é uma quantidade extensiva, os resultados obtidos anteriormente não são mais válidos. Vamos aqui somente comentar os resultados obtidos utilizando o truque das réplicas para obter a energia livre, sem entrar nos detalhes do cálculo [12,13]. Tal truque é utilizado quando temos que realizar uma média sobre as variáveis randômicas congeladas ('quenched') de um sistema [35]-[37], isto é, médias sobre várias amostras para as quais as variáveis congeladas não variam com o tempo. A média da energia livre é calculada por meio da seguinte relação exata:

$$\langle\langle \ln Z \rangle\rangle = \lim_{n \rightarrow 0} \frac{\langle\langle Z^n \rangle\rangle - 1}{n} . \tag{2.32}$$

Isto equivale a considerarmos n réplicas do sistema original e calcularmos a função partição do sistema replicado. A simetria de réplicas é obtida quando todas elas são consideradas idênticas entre si.

Os principais resultados são listados a seguir [12,13] e deles resulta o diagrama de fases do modelo de Hopfield (figura 2.9).

a) Pode-se mostrar, através de argumentos probabilísticos (análise sinal-ruído), que os

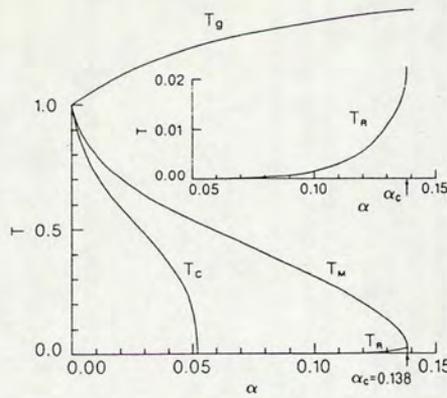


Figura 2.9: Diagrama de fases para o modelo de Hopfield. T_g é a temperatura de transição para o estado de vidro de spin, T_M é a temperatura na qual os estados que se superpõem com uma memória aparecem. T_c é a transição onde estes estados se tornam mínimos globais. A simetria das réplicas é quebrada em T_R [12,13].

padrões originais deixam de ser estáveis quando [39]

$$P > \frac{N}{2 \ln N} \quad (2.33)$$

Este é o limite para que a recuperação aconteça sem erros, ou seja, o armazenamento é perfeito. Se permitirmos uma pequena fração de erros, a capacidade da rede aumenta e é dada pelo valor máximo do parâmetro α ,

$$\alpha < \alpha_c \simeq 0.138 \quad , \quad (2.34)$$

abaixo do qual o sistema efetivamente funciona como uma memória associativa, sendo que os estados estáveis não são mais os padrões originais (que são desestabilizados de acordo com (2.33)), mas estados muito semelhantes àqueles. Desde que α permaneça abaixo do limite acima, a superposição entre estes estados, os quais são chamados de estados de recuperação ('retrieval states'), e as memórias originais é maior do que $m \simeq 0.967$, que é o valor da superposição em $\alpha = \alpha_c$. À medida que α tende a 0, a superposição tende a

crescer, sendo que em $T = 0$ o crescimento é exponencial

$$m \simeq 1 - e^{-1/2\alpha} \quad , \quad \alpha \rightarrow 0 \quad . \quad (2.35)$$

b) O valor de α_c obtido acima supõe que as réplicas são simétricas, isto é, os parâmetros que definem cada uma delas são iguais para todas as réplicas. Esta solução, porém, é instável a baixas temperaturas ($T < T_R(\alpha)$). Quando a simetria das réplicas é quebrada, o valor de α_c aumenta para 0.144, e o valor de m em $\alpha = \alpha_c$ aumenta de 0.967 para 0.983 [40], resultado que está de acordo com aqueles obtidos através de simulação numérica [13].

c) Os estados de recuperação são estados metaestáveis para toda temperatura $T < T_M(\alpha)$, onde a temperatura máxima $T_M(\alpha)$ decresce desde 1 em $\alpha = 0$ até 0 em $\alpha = \alpha_c$. Portanto, a recuperação das memórias é estável frente a uma quantidade pequena de ruído térmico.

d) Quando α é finito, há dois tipos de estados espúrios: para α suficientemente pequeno temos as memórias simétricas, que possuem superposições macroscópicas com diversos padrões. Além destes estados, existe uma fase de vidro de spin ('spin-glass', SG) para todos os valores de α . Esta fase possui uma superposição muito pequena com as memórias, $O(1/\sqrt{\alpha N})$, e aparece para $T < T_g = 1 + \sqrt{\alpha}$. Tanto na fase paramagnética quanto na de vidro de spin as superposições são nulas com todas as memórias ($m_\mu = 0, \forall \mu$), mas na fase de vidro de spin a configuração permanece congelada, ou seja, não sofre variação com o tempo.

Os resultados listados acima são apenas uma amostra da grande quantidade de dados obtida até hoje sobre o modelo de Hopfield, sem dúvida o mais estudado modelo de redes de neurônios. A seguir apresentamos alguns resultados correspondentes ao caso em que os padrões armazenados são correlacionados entre si.

2.3.3 Redes de Baixa Atividade

Quando a fração de neurônios ativos nas memórias difere de 0.5 ($a \neq 0.5$), introduzimos correlações entre os padrões armazenados. Toda componente ξ_i^μ em um padrão pode ser escolhida independentemente com probabilidade $P(\xi)$ dada pela equação (2.7). Definimos

a quantidade b em termos da atividade a

$$b = 2a - 1 \quad (2.36)$$

que é o valor médio de ξ e assume valores entre -1 e $+1$. Com esta distribuição de padrões armazenados, as memórias são correlacionadas do seguinte modo

$$\langle\langle \xi^\mu \xi^\nu \rangle\rangle = \delta_{\mu\nu} + b^2(1 - \delta_{\mu\nu}) \quad (2.37)$$

Os resultados obtidos são os seguintes [41]:

a) considerando a regra de Hebb usual, (2.11), os padrões originais são desestabilizados quando

$$P = \frac{1}{b^2} + 1 \quad (2.38)$$

o que significa que, para o nível de atividade cortical ($\sim 5\%$), podemos armazenar perfeitamente somente 2 memórias!

b) Na tentativa de contornar o problema, Amit *et al* [41] propuseram a seguinte modificação:

$$\bar{J}_{ij} = \frac{1}{N} \sum_{\mu} (\xi_i^{\mu} - b)(\xi_j^{\mu} - b) \quad (2.39)$$

Agora a regra de aprendizagem deixa de ser local, pois introduzimos uma característica que depende de toda a rede (a atividade). O modelo de Hopfield com tal regra possui o dobro de estados espúrios do que o caso não correlacionado ($b = 0$). Aqui, as soluções simétricas são estáveis tanto para n par quanto para n ímpar quando $T = 0$. Além disso, enquanto $|b|$ cresce, as energias destes estados começam a se cruzar e, quando $|b|$ torna-se maior do que $\sqrt{2} - 1$, os estados simétricos tornam-se mínimos globais da energia, como mostra a figura 2.10.

Quando $T \neq 0$, temos uma temperatura crítica $T_c(b)$ que vale

$$T_c(b) = 1 - b^2 \quad (2.40)$$

e, logo abaixo deste valor, o valor de m para $n = 1$ é

$$m^2 = \frac{3T_c}{1 + 3b^2}(T_c - T) \quad (2.41)$$

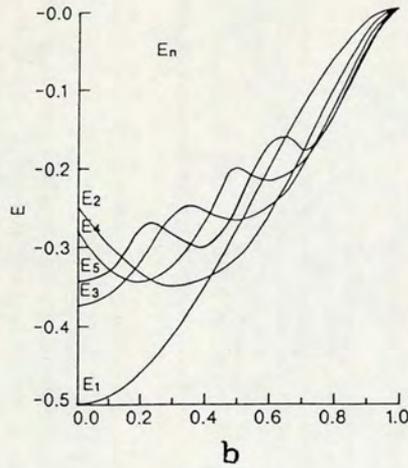


Figura 2.10: As energias dos primeiros cinco estados simétricos em $T = 0$ em função de b [41].

e o autovalor que controla a estabilidade da mistura de duas memórias é

$$\lambda = \frac{2 - 6b^2 \frac{T_c - T}{T_c}}{1 + 3b^2} \tag{2.42}$$

Portanto, as memórias se tornam instáveis abaixo de $T_c(b)$ para $b^2 > 1/3$. Quando a temperatura diminui, primeiro aparecem os estados espúrios estáveis. Somente a temperaturas mais baixas os padrões armazenados tornam-se estáveis, mas neste ponto os estados espúrios (com n par) já possuem energias menores.

c) Por meio da análise sinal-ruído, as memórias são recuperadas sem erro desde que

$$\alpha < \frac{(1 - |b|)^2}{2 \ln N} \tag{2.43}$$

d) Quando temos um número extensivo de memórias ($P = \alpha N$), há um limite para α , $\alpha_c(b)$, abaixo do qual os estados de recuperação são estáveis. Acima deste valor o único estado estável é o de vidro de spin com $m = 0$. Na figura 2.11 mostramos α_c versus b . A regra modificada, (2.39), novamente permite o armazenamento de um número de memórias proporcional a N .

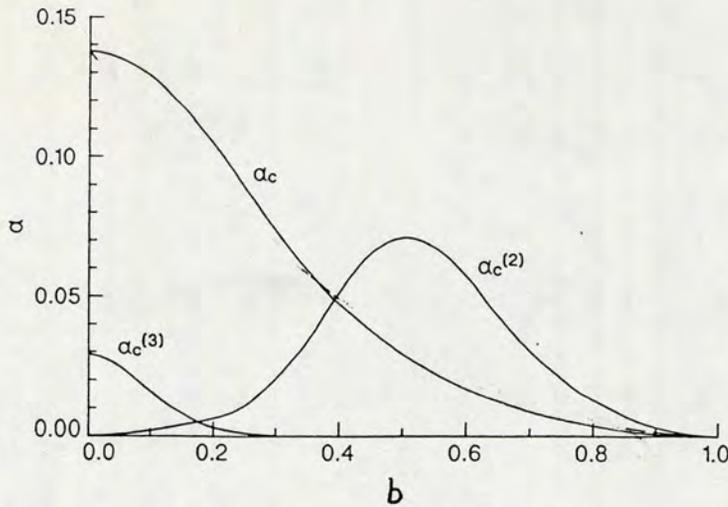


Figura 2.11: A razão de armazenamento crítica α_c versus b para os estados de recuperação e estados simétricos de 2 e 3 padrões [41].

2.3.4 Simulação Numérica

Os tratamentos analíticos podem se tornar, em alguns casos, muito difíceis de serem levados a cabo e, em outros, nem mesmo são conhecidos. É nessas situações que o recurso à simulação numérica se faz necessário. Além disso, a simulação contribui para um maior conhecimento do modelo que está sendo estudado, servindo como guia para a intuição. Entre as quantidades passíveis de serem medidas, escolhemos umas poucas que melhor representam a habilidade do modelo de operar como uma memória associativa: a capacidade máxima de armazenamento α_c , os tamanhos das bacias de atração associadas às memórias e os tempos de convergência, isto é, o número de passos necessários a partir de um determinado estado inicial para atingir a memória escolhida. Associadas a este tempo temos as dispersões ao redor do valor médio, as quais fornecem informações relevantes sobre a estrutura do espaço de fase que cerca os padrões armazenados.

Vamos iniciar, então, com a capacidade máxima de memória, que é medida pelo parâmetro α , como vimos anteriormente. A obtenção numérica de α_c para o modelo de Hopfield foi feita primeiramente por Amit *et al* [11] e, posteriormente, com maior precisão,

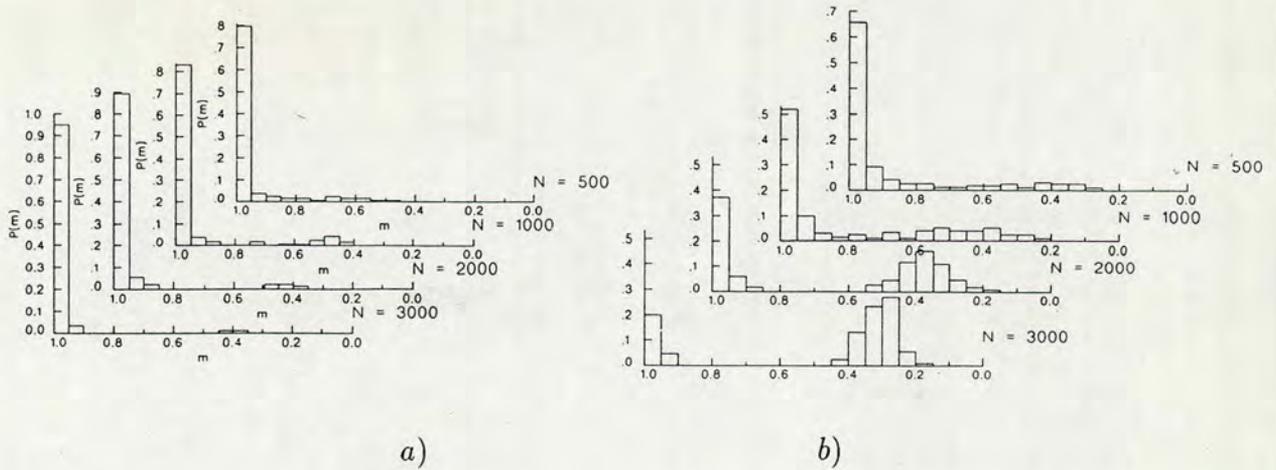


Figura 2.12: Histogramas das superposições do estado final com o estado inicial, o qual é dos padrões aprendidos, para vários valores de N e a) $\alpha = 0.14$ e b) $\alpha = 0.16$ [41].

por Kohring [42]. O procedimento é o seguinte: após armazenar P padrões em uma rede composta de N neurônios, escolhe-se um estado inicial igual a um dos padrões, e testa-se, sucessivamente, a estabilidade de cada um dos neurônios. Se a inversão do estado de algum neurônio diminui a energia, tal inversão é efetuada. Este processo é repetido até atingir-se uma configuração estável. Na figura 2.12 mostramos a distribuição da superposição do estado final com o inicial para dois valores do parâmetro α : 0.14 e 0.16. À medida que N aumenta, o pico próximo de $m = 1$ possui um comportamento diferente para os dois valores de α : quando $\alpha = 0.14$ este máximo aumenta, sendo o valor médio de m , para $N \simeq 3000$, igual a 0.972 ± 0.01 (média sobre 200 padrões). Para $\alpha = 0.16$ o pico diminui. Para testar se o máximo ao redor de $m = 1$ desaparece à medida que $N \rightarrow \infty$, estuda-se a área $P(\alpha, N)$ sob tal pico. Isto foi feito para $\alpha = 0.15$ e 0.16, como mostra a figura 2.13. As curvas são ajustadas por uma exponencial do tipo

$$P = Ae^{B(\alpha_c - \alpha)N}$$

a qual indica que quanto maior é α , mais rápido o pico diminui quando N aumenta. Os parâmetros ajustados foram $A = 0.97 \pm 0.05$, $B = (2.8 \pm 0.3)10^{-2}$ e $\alpha_c = 0.145 \pm 0.01$. Este

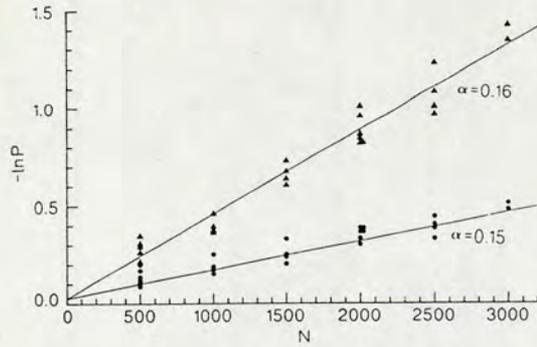


Figura 2.13: Análise da área P sob o maior pico da figura 2.12 usada para determinar α_c [41].

valor para α_c concorda com o resultado obtido com a quebra de simetria das réplicas [40].

Os tamanhos das bacias de atração que rodeiam cada uma das memórias também podem ser estimados [43]. No capítulo 4 analisamos em detalhe o procedimento a ser seguido e aqui apresentamos os resultados obtidos para o modelo de Hopfield. Resumidamente, o processo é o seguinte: cria-se um estado inicial com uma superposição m_0 com uma das memórias, escolhida ao acaso. A seguir os neurônios são testados e invertidos se isto diminui a energia do sistema, até que um estado estável seja atingido. Se, ao final, a superposição m_f com o padrão escolhido for maior do que um determinado valor (0.875, no caso da referência [43]), considera-se que tal memória foi recuperada. Isto é repetido 1000 vezes e se obtém, assim, a fração de recuperação $f(m_0)$, a qual é mostrada na figura 2.14, para vários valores de α e N . Para $\alpha < \alpha_c$ as curvas tendem a aumentar a sua inclinação à medida que N aumenta, ao contrário do que acontece para $\alpha > \alpha_c$. Para $\alpha < \alpha_c$, estas curvas se aproximam de uma função degrau $\Theta(m - m_c)$, a qual é 1 se $m > m_c$ e 0 nos outros casos. Então, o valor m_c estima o tamanho das bacias de atração e seu valor, para alguns valores de α , é mostrado na tabela 2.1.

α	m_c
0.03	0.11
0.06	0.22
0.1	0.37

Tabela 2.1: Tamanho das bacias de atração para vários valores de α [43].

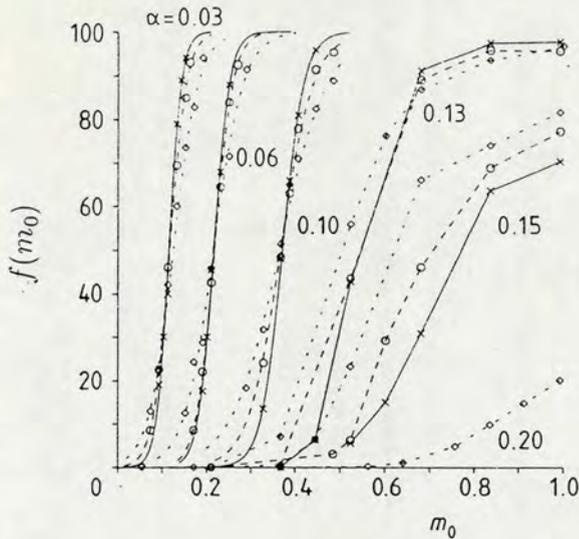


Figura 2.14: A fração $f(m_0)$ de estados que são recuperados com menos do que $N/16$ erros a partir de um estado inicial com superposição m_0 . Os símbolos se referem a $N = 512(\diamond)$, $1024(\circ)$ e $2048(\times)$. Os valores de α utilizados estão marcados ao lado das curvas [43].

Vamos agora analisar os resultados encontrados para o tempo de convergência, definido como o tempo necessário para, a partir de um certo estado inicial, atingir a memória escolhida [44,45]. Na simulação descrita acima, pode-se medir o número de inversões causadas aos neurônios até que o estado estável final seja atingido. Tal número, dividido pelo tamanho da rede N , fornece o tempo de convergência $\langle T \rangle$, mostrado na figura 2.15 para $\alpha = 0.1$, $m_0 = 0.4$ e dois tamanhos de rede. Desta figura pode-se ver que a curva para o maior N é mais alta e menos espalhada. Esta figura também indica que $\langle T \rangle$ é uma função crescente de N . Isto é o que também mostra a figura 2.16. Na figura 2.17 as dispersões no tempo de convergência, $\sigma = \langle T^2 \rangle - \langle T \rangle^2$, versus N são mostradas. Uma análise detalhada das interpretações destes resultados é feita no capítulo 4, mas adiantamos que as duas quantidades, $\langle T \rangle$ e σ estão relacionadas com a homogeneidade e suavidade do espaço de fase ao redor das memórias armazenadas. Por exemplo, dispersões altas são causadas pela existência de distintos caminhos levando até o fundo da bacia de atração, cada um deles com um tempo de convergência diferente. Isto é devido à existência de estados espúrios que retardam a

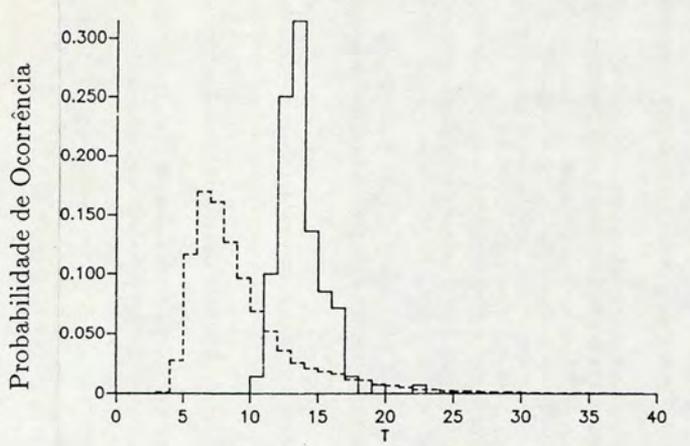


Figura 2.15: Histograma do tempo de convergência para $\alpha = 0.1$ e $m_0 = 0.4$. A curva tracejada é para $N = 1088$ e a curva cheia é para $N = 64000$ [45].

evolução do sistema.

Os resultados brevemente apresentados acima serão novamente discutidos nos próximos capítulos, quando os compararemos aos do modelo RS.

2.4 Modelos com Interações entre k Neurônios

No mesmo espírito de alguns modelos de vidros de spin que consideram interações randômicas entre k spins [46,47], vários trabalhos similares tratam de redes de neurônios [34],[48]-[55].

Em 1987, Elizabeth Gardner considerou a seguinte generalização do modelo de Hopfield [48], onde temos um monômio de grau k nos spins:

$$H = - \sum_{i_1 < \dots < i_k} J_{i_1 \dots i_k} S_{i_1} \dots S_{i_k} \quad , \tag{2.44}$$

onde as interações são dadas pela regra de Hebb generalizada:

$$J_{i_1 \dots i_k} = \frac{k!}{N^{k-1}} \sum_{\mu=1}^P \xi_{i_1}^\mu \dots \xi_{i_k}^\mu \quad . \tag{2.45}$$

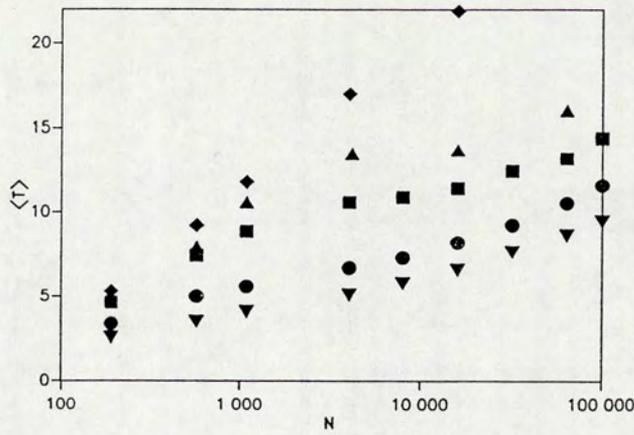


Figura 2.16: Tempo de convergência versus N para $\alpha = 0.1$. Na figura, $m_0 = 0.35$ (◆), 0.375 (▲), 0.4 (■), 0.5 (●) e 0.6 (▼) [45].

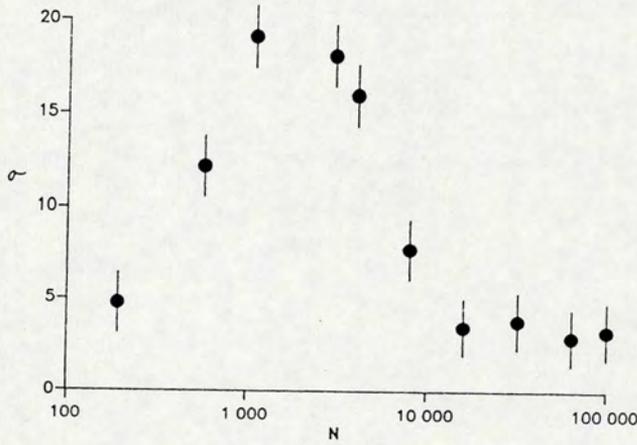


Figura 2.17: Variância da distribuição dos tempos de convergência versus N para $\alpha = 0.1$ e $m_0 = 0.4$ [45].

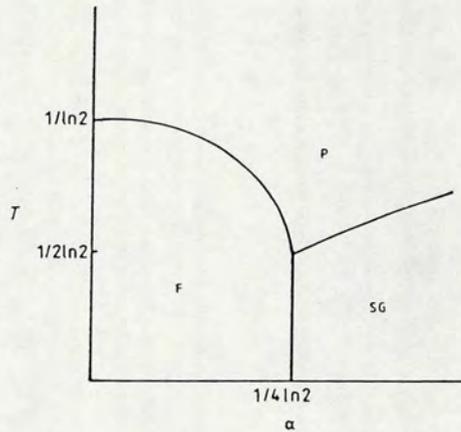


Figura 2.18: Diagrama de fases para $k \rightarrow \infty$ [48].

Os resultados obtidos são os seguintes:

a) utilizando argumentos probabilísticos, pode-se mostrar que o sistema armazena perfeitamente as memórias desde que o seu número seja menor que

$$P = \frac{N^{k-1}}{2(k-1)! \ln N} \quad (2.46)$$

b) Quando $k \rightarrow \infty$, pode-se obter a função partição do sistema de um modo análogo ao modelo de energias randômicas ('Random Energy Model') de Derrida [56]-[61]. Neste limite, o sistema possui três fases, as quais são mostradas na figura 2.18. Se $T > T_g = \sqrt{\alpha / \ln 2}$, temos a fase paramagnética, enquanto que em $T = T_g$ temos uma transição de vidro de spin. Além destas, temos a fase ferromagnética, cuja borda com a fase paramagnética é dada por

$$T_P = \frac{1}{2 \ln 2} + \left[\frac{1}{4(\ln 2)^2} - \frac{\alpha}{\ln 2} \right]^{1/2} \quad (2.47)$$

enquanto que a fronteira entre as fases de vidro de spin e ferromagnética é

$$\alpha = \frac{\ln 2}{4} \quad , \quad (2.48)$$

k	α_c	m_c
2	0.14	0.97
3	0.38	0.855
4	0.395	0.94
∞	$k/4 \ln k$	1

Tabela 2.2: Valores de α_c e m_c como uma função de k [48].

que é independente da temperatura, onde α é definida por

$$\alpha = \frac{P k!}{2 N^{k-1}} \quad (2.49)$$

Esta definição de α é o quociente da quantidade de informações que estamos armazenando ($P \times N$) pelo número de $J_{i_1 \dots i_P}$ independentes ($N^k/k!$). Esta escolha é adequada porque o aumento que ocorre na quantidade de informação que a rede pode guardar é compensado por um aumento equivalente no espaço disponível para armazená-las, as sinapses.

c) Utilizando-se o truque das réplicas podemos estudar o sistema para k finito. Os resultados para α_c e m_c para alguns valores de k podem ser vistos na tabela 2.2.

d) Uma desvantagem em aumentar o valor de k é que o tamanho das bacias de atração diminui, ou seja, a habilidade da rede em reconhecer padrões com uma quantidade finita de erros decresce. Isto acontece principalmente porque o número de estados espúrios metaestáveis aumenta com k e tende a $2^{N(1+O(1/N))}$ quando $k \rightarrow \infty$, ou seja, aproximadamente todos os estados são metaestáveis neste limite.

Abbott e Arian [49] consideraram um Hamiltoniano do tipo

$$H = -N^{1-k/2} \sum_{\mu=1}^P F \left[\frac{1}{\sqrt{N}} \sum_i S_i \xi_i^\mu \right] \quad , \quad (2.50)$$

onde F é um polinômio arbitrário de grau k :

$$F(x) = \sum_{l=0}^k A_l x^l \quad (2.51)$$

Um caso particular de (2.50) é

$$H = -\frac{1}{N^{k-1}\sqrt{2k!}} \sum_{\mu=1}^P \sum_{i_1 \neq i_2 \neq \dots \neq i_k} \xi_{i_1}^\mu \dots \xi_{i_k}^\mu S_{i_1} \dots S_{i_k} \quad (2.52)$$

o qual foi estudado para $k = 3$ e $k \rightarrow \infty$. O parâmetro α é aqui definido como $\alpha = P/N^{k-1}$.

Os resultados são os seguintes:

a) no limite $\alpha \rightarrow 0$, o caso $k = 3$ possui soluções com $m > 0.807$ se $T < 0.504$. Para valores maiores de T a solução é $m = 0$. Quando k é grande, temos soluções com $m > 1 - 2/k$ para

$$T < \frac{k\sqrt{2}}{\sqrt{k!}e^2 \ln k} \quad (2.53)$$

Portanto, para modelos do tipo (2.52) com k grande, embora o número de memórias armazenadas possa ser muito grande, o nível de ruído aceitável é pequeno.

b) No limite $T = 0$, para $k = 3$, temos recuperação de memórias com $m > 0.838$ desde que $\alpha \leq \alpha_c = 0.126$. Para k grande,

$$m > 1 - \frac{1}{k\sqrt{\pi \ln k}} \quad (2.54)$$

sempre que

$$\alpha \leq \alpha_c = \frac{k}{2k! \ln k} \quad (2.55)$$

Embora α vá a zero para k grande, o número total de padrões que podem ser armazenados é muito grande:

$$P = \alpha N^{k-1} \rightarrow \frac{e}{2 \ln k \sqrt{2\pi k}} \left(\frac{eN}{k}\right)^{k-1} \quad (2.56)$$

c) Na figura 2.19 é mostrado o diagrama de fases para o caso $k = 3$. A área sob a curva T_M representa a região onde os estados de recuperação existem como estados estáveis ou metaestáveis. Acima desta curva temos uma fase de vidro de spin e, para temperaturas mais altas, uma fase paramagnética. A linha T_R assinala o ponto onde a suposição de simetria das réplicas deixa de ser válida (as soluções não são mais estáveis).

Horn e Usher [52] consideraram uma variante dos modelos acima dada pelo seguinte Hamiltoniano:

$$H = -N \sum_{\mu=1}^k \epsilon_\mu m_\mu^k \quad (2.57)$$

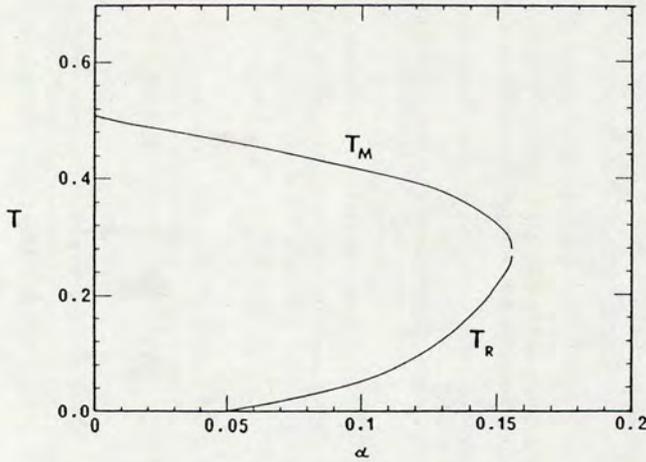


Figura 2.19: Diagrama de fases para o modelo estudado por Abbott e Arian [49] com $k = 3$. Abaixo da linha T_M existem estados estáveis e metaestáveis correlacionados com os padrões armazenados e acima desta linha há um fase de vidro de spin. Abaixo da linha T_R a solução com simetria de réplica é instável [49].

onde o fator ϵ_μ introduz pesos que podem diferenciar os vários padrões armazenados. É interessante não variar os pesos ϵ_μ por mais de um fator de 2, caso contrário a probabilidade de recuperar padrões com peso pequeno será muito baixa. Os resultados são:

a) por meio de argumentos probabilísticos e considerando o caso $\epsilon_\mu = 1, \forall \mu$, mostra-se que o número máximo de padrões perfeitamente armazenados é

$$P = \frac{2^{k-2}(k-1)!N^{k-1}}{k(2k-2)! \ln N} \quad (2.58)$$

A diferença para o resultado do item a) de Gardner deve-se ao fato de que (2.52) permite índices iguais em $J_{i_1 \dots i_k}$. Com argumentos análogos, sem recurso às réplicas, e permitindo uma fração finita de erros, reobtem-se os resultados de Gardner [48] e Abbott e Arian [49], discutidos anteriormente. Por exemplo, para $k = 3$ encontra-se $\alpha_c = 0.25$ com $m_c = 0.84$ e, para $k = 4$, $\alpha_c = 0.195$ com $m_c = 0.92$.

b) Assumindo dois tipos de pesos para os padrões, ϵ_1 e ϵ_2 , de tal modo que $\epsilon_1 > \epsilon_2$, o

resultado análogo ao do item a), que permite superposições diferentes da unidade, é:

$$\alpha_1 + \left(\frac{\epsilon_2}{\epsilon_1}\right)^2 \alpha_2 < \alpha_c$$

$$\alpha_2 + \left(\frac{\epsilon_1}{\epsilon_2}\right)^2 \alpha_1 < \alpha_c$$
(2.59)

onde α_1 e α_2 descrevem as capacidades de armazenamento das duas famílias de padrões e α_c é o mesmo do item anterior.

Vários outros trabalhos consideram interações entre vários neurônios. Entre eles, temos o de Tamarit *et al* [55] que inclui diluição no modelo considerado por Horn e Usher. Baldi e Venkatesh [50] estimam o número de estados que podem ser armazenados de acordo com diferentes regras de aprendizagem, Personnaz *et al* [51] consideram uma extensão da regra da pseudo-inversa que inclui interações ternárias. Perez e Matus [53] estudam a existência de estados espúrios para diferentes k , Peretto e Niez [34] e Kohring [54] analisam a capacidade de armazenamento das generalizações do modelo de Hopfield e da regra de Hebb, como as equações (2.44)-(2.45). Embora todos estes trabalhos envolvam resultados da mais alta relevância, os fatos básicos relativos aos modelos com interações entre mais de dois neurônios já foram apresentados nesta seção.

2.5 Conclusão

Apresentamos, neste capítulo, uma breve revisão dos principais conceitos e modelos de redes de neurônios. Não temos a pretensão de exaurir o assunto e certamente muitos trabalhos importantes não foram citados. Damos especial enfoque para o caso em que o número de memórias armazenadas permanece finito enquanto que o tamanho da rede aumenta ($\alpha \rightarrow 0$) e para modelos que envolvem interações entre $k > 2$ neurônios. Em todos estes casos, os padrões são desestabilizados quando o número de memórias por neurônio ultrapassa um determinado valor limite α_c , isto é, deixam de ser mínimos da energia.

Existem muitas outras situações nas quais o modelo de Hopfield foi estudado, como por exemplo, quando as memórias ou os neurônios estão estruturados hierarquicamente [62]-

[67], quando a rede é capaz de esquecer, isto é, memórias antigas vão sendo esquecidas à medida que novas são aprendidas [68]-[72], quando introduzimos assimetria ($J_{ij} \neq J_{ji}$) e/ou diluição (número de coordenação menor do que N) nas sinapses [73]-[79], quando queremos armazenar seqüências temporais de padrões (como um poema, por exemplo) [80], quando o sistema é capaz de categorização (classificação) [81] e muitos outros aspectos, os quais embora interessantes e relevantes, não guardam relação direta com o que discutiremos a seguir.

No capítulo seguinte vamos considerar um modelo alternativo para redes de neurônios (modelo RS) e apresentar suas principais características.

Capítulo 3

O Modelo RS

Já foi visto no capítulo anterior que o modelo de Hopfield na sua versão original possui sérias limitações quando a quantidade de memórias armazenadas é muito grande ou quando estas memórias são correlacionadas. Várias tentativas para superar estes problemas foram feitas, seja propondo novas regras de aprendizagem ou novas formas da função energia. Estas visões alternativas também apresentam problemas (veja a seção 2.4). Tentando contornar algumas destas dificuldades, recentemente foi apresentado o modelo RS [14]-[18] o qual será o motivo deste e dos próximos capítulos. Na seção 3.1 introduzimos o modelo e os resultados que podem ser obtidos da forma do Hamiltoniano. Nas seções 3.2 e 3.3 apresentamos os resultados de campo-médio para os estados de equilíbrio do sistema para duas possíveis configurações e concluimos na seção 3.4.

3.1 Formulação do Modelo

A energia a ser considerada é o produto do quadrado das distâncias euclidianas entre o estado da rede $\vec{S} = (S_1, \dots, S_N)$ e as P memórias armazenadas $\vec{\xi}^\mu = (\xi_1^\mu, \dots, \xi_N^\mu)$ com $\mu = 1, \dots, P$:

$$E = N \prod_{\mu=1}^P \left[\frac{1}{2N} \sum_{i=1}^N (\xi_i^\mu - S_i)^2 \right] , \quad (3.1)$$

O fator $1/2N$ é introduzido para obter o comportamento correto no limite termodinâmico ($N \rightarrow \infty$). A primeira conclusão que se pode tirar a partir da forma de E é que a energia é sempre positiva $E[\vec{S}] \geq 0$, sendo a igualdade satisfeita quando $\vec{S} = \vec{\xi}^\mu$, para qualquer $\mu \leq P$. Portanto os padrões armazenados sempre são mínimos de E , independentemente do seu número. Isto sugere que a capacidade de memória da rede (medida por $\alpha = P/N$) deve ser muito alta.

Também pode ser visto da forma de (3.1) que a altura da barreira de energia entre dois mínimos é maior quanto maior for a distância entre eles no espaço de fase. Isto significa que memórias correlacionadas (semelhantes) possuem barreiras menores entre si do que padrões descorrelacionados. Portanto, se partirmos de uma situação com ruído nulo ($T = 0$) e aumentarmos gradualmente a temperatura, os primeiros estados a se misturarem são aqueles mais próximos (mais semelhantes). Isto pode ser encarado como uma associação de idéias, isto é, uma idéia sugere outra formando uma cadeia associativa. Este fenômeno de associatividade é importante e merece um estudo mais aprofundado no futuro como, por exemplo, relacionar o nível de ruído e a altura das barreiras de energia que separam vários mínimos com a meia-vida de um estado, ou seja, a probabilidade de ultrapassar tal barreira.

No modelo de Hopfield, devido à forma quadrática do Hamiltoniano, o valor de energia que um certo estado \vec{S} assume é o mesmo para o estado $-\vec{S}$, seu antípoda. Como conseqüência, tanto a memória quanto a anti-memória são igualmente armazenadas neste modelo, o que não acontece com o modelo RS. Para que ambas sejam armazenadas é preciso fazê-lo explicitamente, por exemplo, do seguinte modo. Primeiramente reescrevemos (3.1) como

$$E = N \prod_{\mu=1}^P (1 - m_\mu) \quad , \quad (3.2)$$

onde m_μ , como no capítulo anterior, é definido por

$$m_\mu = \frac{1}{N} \sum_{i=1}^N \xi_i^\mu S_i \quad . \quad (3.3)$$

e mede a superposição entre o padrão $\vec{\xi}^\mu$ e o estado da rede \vec{S} . Se, para cada um dos

P padrões $\vec{\xi}^\mu$, a rede também armazenar o seu antípoda $-\vec{\xi}^\mu$, a função energia pode ser reescrita

$$E = N \prod_{\mu=1}^P (1 - m_\mu) \prod_{\nu=1}^P (1 + m_\nu) = N \prod_{\mu=1}^P (1 - m_\mu^2) \quad , \quad (3.4)$$

Assim, temos duas possíveis situações extremas, a saber, quando somente as memórias são aprendidas (caso OPS ¹) e quando tanto a memória quanto a anti-memória são armazenadas (caso PAS ¹). Estes dois casos apresentam um comportamento completamente diferente, como será mostrado nas seções seguintes.

Para mostrar o caráter de multi-interação apresentado pelo modelo, expandimos a equação (3.2):

$$E = N \left(1 + \sum_{i_1=1}^N \mathbf{J1}_{i_1} S_{i_1} + \sum_{i_1, i_2}^N \mathbf{J2}_{i_1 i_2} S_{i_1} S_{i_2} + \dots + \sum_{i_1, \dots, i_P}^N \mathbf{JP}_{i_1 \dots i_P} S_{i_1} S_{i_2} \dots S_{i_P} \right) \quad , \quad (3.5)$$

onde as intensidades sinápticas entre vários neurônios são definidas como

$$\mathbf{JK}_{i_1 \dots i_K} = \frac{(-1)^K}{N^K} \sum_{\mu_1 < \mu_2 < \dots < \mu_K} \xi_{i_1}^{\mu_1} \xi_{i_2}^{\mu_2} \dots \xi_{i_K}^{\mu_K} \quad . \quad (3.6)$$

As conexões sinápticas $\mathbf{JK}_{i_1 \dots i_K}$ representam as interações entre K neurônios, sendo P a maior ordem de interação possível. Observe que quando $P > N$, a ordem desta interação ainda permanece menor do que N , pois teremos, na seqüência $S_{i_1} S_{i_2} \dots S_{i_P}$, dois índices iguais, sendo o produto de seus estados igual a 1. Quando, além das P memórias, consideramos também as respectivas P anti-memórias, todas as intensidades sinápticas de ordem ímpar se anulam e a eq.(3.5) se reduz a

$$E = N \left(1 + \sum_{i_1, i_2}^N \mathbf{J2}_{i_1 i_2} S_{i_1} S_{i_2} + \dots + \sum_{i_1, \dots, i_{2P}}^N \mathbf{J(2P)}_{i_1 \dots i_{2P}} S_{i_1} S_{i_2} \dots S_{i_{2P}} \right) \quad , \quad (3.7)$$

onde

$$\mathbf{JK}_{i_1 \dots i_K} = \frac{(-1)^{K/2}}{N^K} \sum_{\mu_1 < \dots < \mu_{K/2}} \xi_{i_1}^{\mu_1} \xi_{i_2}^{\mu_1} \dots \xi_{i_{K-1}}^{\mu_{K/2}} \xi_{i_K}^{\mu_{K/2}} \quad , \quad (3.8)$$

Diferentemente de outros modelos que introduziram interações de ordem mais alta do que dois entre os neurônios, as conexões sinápticas aqui são assimétricas, embora no caso

¹Conservamos a abreviatura inglesa para manter uma notação uniforme com as referências [15]-[18].

PAS exista uma certa simetria proveniente do fato das antimemórias estarem também armazenadas: podemos, por exemplo, trocar i_1 por i_2 sem alterar a conexão.

Se os padrões são descorrelacionados, \mathbf{JK} é uma soma de C_P^K (combinação de P elementos tomados K a K) passos aleatórios ('random-walk') para o caso OPS e $C_P^{K/2}$ para o caso PAS, sendo que cada passo pode assumir os valores $\pm 1/N^K$ com igual probabilidade. Portanto, o valor médio de \mathbf{JK} é nulo, ou seja,

$$\langle \mathbf{JK}_{i_1 \dots i_K} \rangle = 0$$

e a dispersão ao redor da média, ΔJK , é

$$\Delta JK = \frac{1}{N^K} \sqrt{\frac{P!}{K!(P-K)!}} \quad , \quad (3.9)$$

para o caso OPS e

$$\Delta JK = \frac{1}{N^K} \sqrt{\frac{P!}{\left(\frac{K}{2}\right)! \left(P - \frac{K}{2}\right)!}} \quad (3.10)$$

para o PAS.

Para estimarmos a rapidez de convergência das expansões eqs. (3.9) e (3.10), tomamos a razão entre dois termos consecutivos, isto é,

$$\frac{\Delta JK}{\Delta J(K-1)} = \frac{1}{N} \sqrt{\frac{P-K+1}{K}} \leq \sqrt{\frac{\alpha}{N}} \quad , \quad (3.11)$$

para o caso OPS, enquanto que para o PAS temos

$$\frac{\Delta JK}{\Delta J(K-2)} = \frac{1}{N^2} \sqrt{\frac{P-K/2+1}{K/2}} \leq \sqrt{\frac{\alpha}{N^3}} \quad , \quad (3.12)$$

Ambas equações vão para zero no limite $\alpha < \infty$ e $N \rightarrow \infty$. Portanto, a relevância dos termos é tanto menor quanto maior for a ordem de interação envolvida, desde que os padrões sejam descorrelacionados e α suficientemente pequeno. Isto permite que cortemos a expansão se as condições acima forem satisfeitas. Particularmente, no caso de memórias e anti-memórias, a função energia até segunda ordem fica

$$E \approx N \left(1 - \frac{1}{N^2} \sum_{i,j} \sum_{\mu=1}^P \xi_i^\mu \xi_j^\mu S_i S_j \right) \quad , \quad (3.13)$$

que é, a menos de algumas constantes aditivas e multiplicativas, exatamente o modelo de Hopfield com a regra de Hebb para a aprendizagem [11]. Portanto, o Hamiltoniano de Hopfield é recuperado, para $N \rightarrow \infty$, no limite de α pequeno e padrões descorrelacionados. Nos casos onde tal modelo não apresenta um comportamento satisfatório, como por exemplo, quando os padrões são correlacionados ou quando a capacidade α é alta, os termos de interação mais altos são necessários para que o sistema funcione razoavelmente como uma memória associativa.

O comportamento da energia livre no limite termodinâmico nos fornece informações sobre os possíveis estados de equilíbrio, estados metaestáveis e as respectivas bacias de atração. Vamos mostrar agora que a eq.(3.1), para a configuração PAS, depende linearmente do tamanho da rede, sendo então a energia por neurônio finita no limite termodinâmico para alguns estados, podendo, em regiões remotas do espaço de fase, divergir. Para verificar isto vamos tomar um estado que possua superposição m_1 com a primeira memória, m_2 com a segunda, ..., m_n com a n -ésima memória e descorrelacionado com todas as restantes, ou seja, $m_\mu \sim \pm 1/\sqrt{N}$ se $\mu > n$. Neste caso

$$\begin{aligned} \frac{E}{N} &= (1 - m_1^2) \dots (1 - m_n^2) \left(1 - \frac{1}{N}\right)^{\alpha N - n} \\ &\longrightarrow (1 - m_1^2) \dots (1 - m_n^2) e^{-\alpha} \quad , \quad N \rightarrow \infty \end{aligned} \tag{3.14}$$

Como, por razões físicas, n é uma quantidade que deve permanecer finita (a rede não pode se superpor macroscopicamente com um número infinito de memórias), a energia por neurônio, no caso PAS, permanece finita desde que α seja finito, ou seja, P seja linear em N . Note-se que a configuração acima, $\mathbf{m} = (m_1, \dots, m_n, \pm 1/\sqrt{N}, \dots, \pm 1/\sqrt{N})$, engloba boa parte dos possíveis estados da rede.

Para a configuração OPS a situação é diferente. O sistema apresenta um comportamento interessante somente quando $\alpha \rightarrow 0$. Isto ocorre porque quando $P \rightarrow \infty$, as bacias de atração das memórias individuais fundem-se numa grande bacia comum, a qual torna-se plana ($E = 0$) na região compreendida pelas memórias, o centro da grande bacia e todos os

pontos localizados entre o centro e as memórias. A energia cresce em direção à região que contém as anti-memórias. Com efeito, para o estado geral considerado acima, a energia por neurônio vale

$$\frac{E}{N} = (1 - m_1) \dots (1 - m_n) \left(1 - \frac{1}{\sqrt{N}}\right)^{\frac{\alpha N - n}{2}} \left(1 + \frac{1}{\sqrt{N}}\right)^{\frac{\alpha N - n}{2}} \quad (3.15)$$

$$\longrightarrow (1 - m_1) \dots (1 - m_n) e^{-\alpha/2}, \quad N \rightarrow \infty \quad (3.16)$$

Por outro lado, aqueles estados que se encontram na região compreendida pelas bacias de atração e o estado central possuem superposições $+1/\sqrt{N}$ e portanto a energia por neurônio somente será não nula quando $\alpha \rightarrow 0$. O sistema OPS funciona como uma memória associativa no caso em que o número de padrões armazenados é finito, como mostram os resultados dos cálculos de campo médio [15,16] e simulações numéricas [17]-[18] que serão discutidos na seção 3.3 e no capítulo 4, respectivamente. Isto também é o que se concluiu acima: quando $\alpha \neq 0$, o fundo da bacia central torna-se plano com $E = 0$, a energia crescendo em direção à região onde se encontram as anti-memórias. A forma do espaço de fases nestas duas situações, P finito e infinito, é esquematicamente mostrada na figura 3.1.

Nas seções seguintes vamos estudar os estados de equilíbrio do sistema para as duas configurações descritas acima, OPS e PAS.

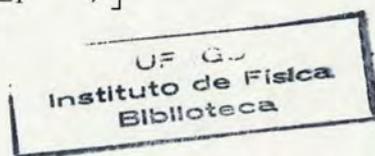
3.2 Teoria de Campo Médio com Anti-memórias: Caso PAS

O procedimento desenvolvido aqui é análogo àquele realizado para modelo de Hopfield [11], descrito na seção 2.3.1. A função partição é

$$Z = \sum_{\{S_i\}} e^{-\beta E[\vec{S}]} \quad (3.17)$$

que pode ser escrita, introduzindo a função delta, como

$$Z = \sum_{\{S_i\}} \int_{-\infty}^{+\infty} dm_1 \dots dm_P \left[\prod_{\mu=1}^P \delta \left(m_\mu - \frac{1}{N} \sum_{i=1}^N \xi_i^\mu S_i \right) \right] e^{-\beta E} \quad .$$



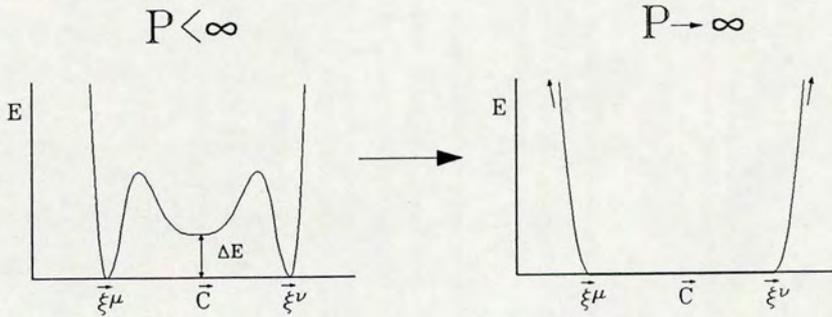


Figura 3.1: Representação pictórica do espaço de fase na configuração OPS.

Utilizando a seguinte identidade

$$\delta \left(m_\mu - \frac{1}{N} \sum_{i=1}^N \xi_i^\mu S_i \right) = \frac{\beta N}{2\pi i} \int_{-i\infty}^{+i\infty} dt_\mu \exp \left[\beta N t_\mu \left(m_\mu - \frac{1}{N} \sum_{i=1}^N \xi_i^\mu S_i \right) \right] , \quad (3.18)$$

podemos efetuar a soma sobre configurações (soma sobre os spins) e reescrever a função partição como

$$Z = \left(\frac{\beta N}{2\pi i} \right)^P \int d\mathbf{m} \int d\mathbf{t} e^{-\beta N f(\mathbf{m}, \mathbf{t})} , \quad (3.19)$$

onde $\mathbf{m} = (m_1, \dots, m_P)$, $\mathbf{t} = (t_1, \dots, t_P)$, $d\mathbf{m} = dm_1 \dots dm_P$, $d\mathbf{t} = dt_1 \dots dt_P$ e $f(\mathbf{m}, \mathbf{t})$ é a energia livre por neurônio que vale

$$f(\mathbf{m}, \mathbf{t}) = \prod_{\mu=1}^P (1 - m_\mu^2) + \mathbf{m} \cdot \mathbf{t} - \frac{1}{\beta} \langle \langle \ln [2 \cosh (\beta \boldsymbol{\xi} \cdot \mathbf{t})] \rangle \rangle , \quad (3.20)$$

onde $\langle \langle \rangle \rangle$ é a média sobre as memórias e o produto escalar entre os vetores \mathbf{m} e \mathbf{t} é

$$\mathbf{m} \cdot \mathbf{t} = \sum_{\mu=1}^P m_\mu t_\mu .$$

Através do método do ponto de sela, dado que $f(\mathbf{m}, \mathbf{t})$ permanece finita quando $N \rightarrow \infty$, podemos calcular a integral em (3.19) (ver Apêndice A): os estados de equilíbrio do sistema

descritos pelos parâmetros \mathbf{m} e \mathbf{t} são aqueles que minimizam $f(\mathbf{m}, \mathbf{t})$. Diferenciando $f(\mathbf{m}, \mathbf{t})$ em relação a m_μ e t_μ obtemos:

$$\frac{\partial f}{\partial m_\mu} = -2m_\mu \prod_{\substack{\nu=1 \\ \nu \neq \mu}}^P (1 - m_\nu^2) + t_\mu \quad ,$$

e

$$\frac{\partial f}{\partial t_\mu} = m_\mu - \langle\langle \xi^\mu \tanh(\beta \xi \cdot \mathbf{t}) \rangle\rangle \quad .$$

As condições de extremo,

$$\frac{\partial f}{\partial m_\mu} = 0$$

e

$$\frac{\partial f}{\partial t_\mu} = 0 \quad ,$$

fornecem o seguinte conjunto de equações transcendentais acopladas:

$$t_\mu = 2m_\mu \prod_{\substack{\nu=1 \\ \nu \neq \mu}}^P (1 - m_\nu^2) \quad (3.21)$$

$$m_\mu = \langle\langle \xi^\mu \tanh(\beta \xi \cdot \mathbf{t}) \rangle\rangle \quad . \quad (3.22)$$

As condições de extremo não são suficientes para garantir que uma determinada solução do sistema acima seja estável. É necessário que tal extremo seja um mínimo local de $f(\mathbf{m}, \mathbf{t})$, pois algumas soluções podem ser máximos ou pontos de sela. Por outro lado, uma solução que é um mínimo local para um determinado intervalo de temperatura pode não o ser em outro. A análise da estabilidade das soluções é feita mais adiante.

Substituindo (3.21) em (3.22)

$$m_\lambda = \langle\langle \xi^\lambda \tanh \left[2\beta \sum_{\mu=1}^P \xi^\mu m_\mu \prod_{\substack{\nu=1 \\ \nu \neq \mu}}^P (1 - m_\nu^2) \right] \rangle\rangle \quad , \quad (3.23)$$

obtemos um sistema de equações para o valor da superposição m_μ , que fornece informações sobre o surgimento de ordem na rede. O conjunto de P equações transcendentais acopladas dado pela equação acima é difícil de resolver. O que normalmente se faz é propor diferentes

'ansätze' e estudá-los. Vamos testar a seguinte solução: supomos que o estado da rede se superpõe macroscopicamente ($m \gg 1/\sqrt{N}$) e de maneira igual com um conjunto finito de n memórias e microscopicamente ($|m| \approx 1/\sqrt{N}$) com as $P - n$ restantes, ou seja:

$$\mathbf{m} = \left(\underbrace{m, m, \dots, m}_n, \underbrace{0, 0, \dots, 0}_{P-n} \right) \quad , \quad (3.24)$$

A solução acima engloba três casos importantes: i) o estado paramagnético quando $m = 0$, ii) as memórias próprias quando $n = 1$ e iii) os estados simétricos (espúrios) quando $n > 1$ e $m \neq 0$. A eq.(3.23) fica

$$m = \left\langle \left\langle \xi^\mu \tanh \left[2\beta m (1 - m^2)^{n-1} \sum_{\lambda=1}^n \xi^\lambda \right] \right\rangle \right\rangle \quad ; \quad \mu \leq n \quad (3.25)$$

e, para $\mu > n$ a equação torna-se uma identidade ($0 = 0$). Somando (3.25) desde $\mu = 1$ até n obtemos

$$m = \frac{1}{n} \left\langle \left\langle z_n \tanh \left[2\beta m (1 - m^2)^{n-1} z_n \right] \right\rangle \right\rangle \quad , \quad (3.26)$$

onde

$$z_n = \sum_{\mu=1}^n \xi^\mu \quad . \quad (3.27)$$

Agora vamos substituir a média $\langle \langle \rangle \rangle$ pela média das quantidades z_n . Para fazer isto notamos que cada valor l que z_n assume aparece C_n^l vezes, se as memórias são descorrelacionadas, isto é, se a probabilidade de $\xi_i^\mu = +1$ é $1/2$. Então, com os pesos adequados, a média fica

$$m = \frac{1}{2^n n} \sum_{k=0}^n \frac{n!}{k!(n-k)!} z_n \tanh \left[2\beta m (1 - m^2)^{n-1} z_n \right] \quad , \quad (3.28)$$

onde $z_n = n - 2k$. A solução da eq.(3.28) é obtida numericamente e a figura 3.2 mostra as curvas referentes ao lado direito da eq.(3.28) para diferentes valores de n e β . As soluções são os pontos (neste caso, dois) onde tais curvas interseccionam a reta m . O estado paramagnético $m = 0$ é sempre solução, embora nem sempre seja estável. A existência de uma solução não trivial ($m \neq 0$) é dada pela seguinte desigualdade

$$\left[\frac{\partial}{\partial m} \frac{1}{n 2^n} \sum_{k=0}^n \frac{n!}{k!(n-k)!} (n - 2k) \tanh \left[2\beta m (1 - m^2)^{n-1} (n - 2k) \right] \right]_{m=0} \geq 1 \quad . \quad (3.29)$$

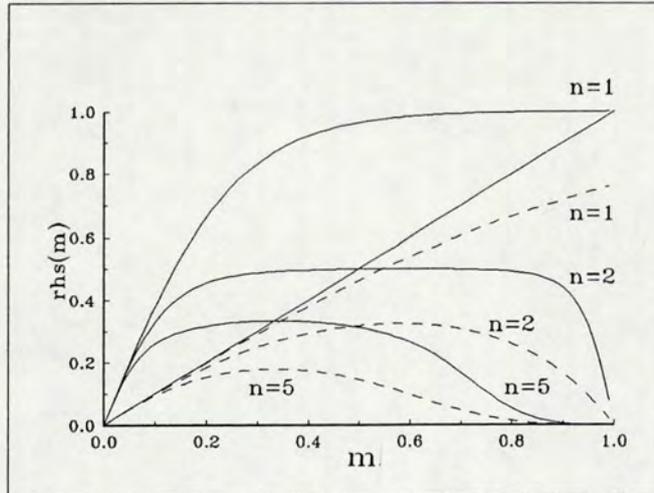


Figura 3.2: Lado direito da eq.(3.28), $rhs(m)$, para $n = 1, 2$ e 5 . As linhas sólidas e tracejadas referem-se, respectivamente, a $\beta = 2$ e $\beta = 0.5$. A linha reta intercepta $rhs(m)$ para $m > 0$ somente se $\beta > 1/2$.

Esta condição implica que para que a curva corte a reta m em um segundo ponto além da origem, é necessário que a sua inclinação no ponto $m = 0$ seja maior do que a tangente da reta, que vale 1. Calculando a derivada na origem resulta

$$\frac{1}{n2^n} \sum_{k=0}^n \frac{n!}{k!(n-k)!} (n-2k)^2 2\beta \geq 1 \quad , \quad (3.30)$$

e com o uso das seguintes identidades

$$\sum_{k=0}^n \frac{n!}{k!(n-k)!} = 2^n \quad ,$$

$$\sum_{k=0}^n \frac{n!}{k!(n-k)!} k = n2^{n-1}$$

e

$$\sum_{k=0}^n \frac{n!}{k!(n-k)!} k^2 = n(n+1)2^{n-2} \quad ,$$

a condição (3.30) transforma-se em $\beta \geq 1/2$, ou seja, temos uma solução $m \neq 0$ somente se $T \leq 2$. Isto também pode ser visto claramente na figura 3.3, onde as soluções da eq.(3.28) são mostradas em função de β , para vários valores de n .

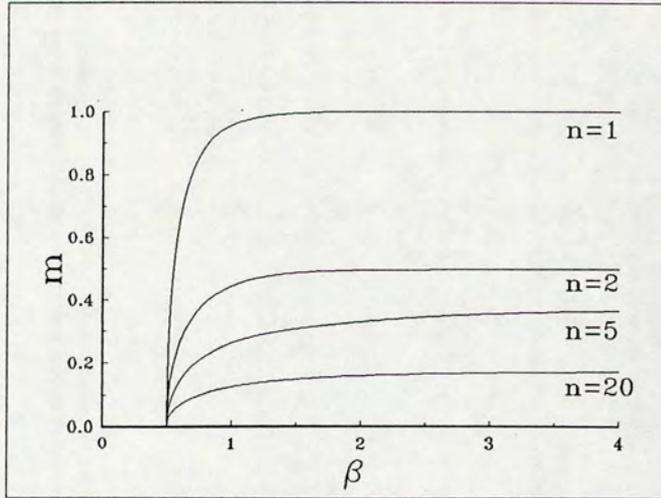


Figura 3.3: As soluções m versus β para diferentes valores de n .

Em $T = 0$ pode-se calcular explicitamente o valor da solução m da eq.(3.28). Reescrevemos a eq.(3.25) em $T = 0$ como

$$m = \left\langle \left\langle \xi^\mu \operatorname{sgn} \left(\sum_{\nu=1}^n \xi^\nu \right) \right\rangle \right\rangle, \quad \mu \leq n \quad (3.31)$$

Sem perda de generalidade podemos tomar $\mu = 1$. Explicitando a média $\langle \rangle$ e efetuando a soma em ξ^1 :

$$m = \frac{1}{2^P} \sum_{\xi^2=-1}^{+1} \dots \sum_{\xi^P=-1}^{+1} \left[\operatorname{sgn} \left(1 + \sum_{\nu=2}^n \xi^\nu \right) + \operatorname{sgn} \left(1 - \sum_{\nu=2}^n \xi^\nu \right) \right]$$

Os possíveis valores que $\sum_{\nu=2}^n \xi^\nu$ pode assumir são simétricos ao redor do zero, ou seja, para cada valor positivo existe o correspondente valor negativo e ambos aparecem o mesmo número de vezes. Usando isso é possível somar as duas funções sinal:

$$m = \frac{1}{2^{P-1}} \sum_{\xi^2=-1}^{+1} \dots \sum_{\xi^P=-1}^{+1} \operatorname{sgn} \left(1 + \sum_{\nu=2}^n \xi^\nu \right)$$

Ou, definindo $\sum_{\nu=2}^n \xi^\nu = z_{n-1}$ e substituindo a média sobre os restantes ξ^μ , com $2 \leq \mu \leq P$, pela média sobre z_{n-1} , obtém-se:

$$m = \frac{1}{2^{n-1}} \sum_{l=0}^{n-1} \frac{(n-1)!}{l!(n-1-l)!} \operatorname{sgn}(n-2l) \quad (3.32)$$

Vamos agora considerar separadamente os casos onde n é par ou ímpar. Quando n é ímpar ($n = 2k + 1$) resulta:

$$m = \frac{1}{2^{2k}} \sum_{l=0}^{2k} \frac{(2k)!}{l!(2k-l)!} \operatorname{sgn}(1 + 2k - 2l) \quad .$$

Existe um número ímpar de termos, sendo que o primeiro ($l = 0$) se cancela com o último ($l = 2k$), o segundo ($l = 1$) se cancela com o penúltimo ($l = 2k - 1$) e assim por diante. O único termo que resta é então $l = k$:

$$m = \frac{(2k)!}{2^{2k}(k!)^2} = \frac{(n-1)!}{2^{n-1} \left[\left(\frac{n-1}{2} \right)! \right]^2} \quad , \quad (3.33)$$

Quando n é par ($n = 2k$), ficamos com:

$$m = \frac{1}{2^{2k-1}} \sum_{l=0}^{2k-1} \frac{(2k-1)!}{l!(2k-1-l)!} \operatorname{sgn}(k-l) \quad .$$

Novamente temos um cancelamento de termos. O único termo que resta é $l = k - 1$, uma vez que $l = k$ não contribui. Então

$$m = \frac{(2k-1)!}{2^{2k-1}(k-1)!k!} = \frac{(n-1)!}{2^{n-1} \left(\frac{n}{2} - 1 \right)! \left(\frac{n}{2} \right)!} \quad , \quad (3.34)$$

Quando n é muito grande, podemos utilizar a fórmula de Stirling, $k! \simeq \sqrt{2\pi k} k^k e^{-k}$, nas eqs.(3.33) e (3.34), obtendo

$$m \simeq \sqrt{\frac{2}{\pi n}} \quad .$$

Portanto, em $T = 0$, o valor de m decai com $1/\sqrt{n}$ quando n é muito grande, como era de se esperar, porque para que a rede se superponha com muitas memórias simultaneamente, é preciso que esta superposição seja pequena com cada uma delas, uma vez que são descorrelacionadas.

Vamos agora estudar a estabilidade destas soluções. Para isto é necessário analisar as flutuações de segunda ordem ao redor delas, $\delta^2 f$, e garantir que as soluções encontradas são mínimos da energia livre. Após algumas manipulações algébricas, chegamos à seguinte

expressão:

$$\begin{aligned}
 \delta^2 f &= \sum_{\mu=1}^n \left[2(1-m^2)^{n-1} - 4\beta(1-m^2)^{2n-4} D_1 \right] \delta m_\mu^2 \\
 &+ \sum_{\mu=n+1}^P \left[2(1-m^2)^n \right] \delta m_\mu^2 \\
 &+ \sum_{\mu=1}^n \sum_{\substack{\nu=1 \\ \nu \neq \mu}}^n \left[-4m^2(1-m^2)^{n-2} - 4\beta(1-m^2)^{2n-4} D_2 \right] \delta m_\mu \delta m_\nu \quad ,
 \end{aligned} \tag{3.35}$$

onde

$$D_1 = \frac{1}{n} \left\langle\left\langle \operatorname{sech}^2 \left[2\beta m(1-m^2)^{n-1} z_n \right] \left\{ n(1+m^2)^2 - 4m^2 z_n^2 [1-m^2(n-1)] \right\} \right\rangle\right\rangle \tag{3.36}$$

e

$$\begin{aligned}
 D_2 = \frac{1}{n(n-1)} \left\langle\left\langle \operatorname{sech}^2 \left[2\beta m(1-m^2)^{n-1} z_n \right] \right. \right. \\
 \left. \left. \left\{ (1+m^2)^2 (z_n^2 - n) - 4m^2(n-1) z_n^2 [1-m^2(n-1)] \right\} \right\rangle\right\rangle \quad . \tag{3.37}
 \end{aligned}$$

A matriz formada pelos elementos $A_{\mu\nu} = \partial^2 f / \partial m_\mu \partial m_\nu$ dados pela eq.(3.35) precisa ser diagonalizada (ver discussão no Apêndice A). Os autovalores, cujos sinais ditam a estabilidade ou não das soluções, são os seguintes (Apêndice B):

$$\lambda_1 = 2(1-m^2)^n \quad , \tag{3.38}$$

que é $P - n$ vezes degenerado,

$$\lambda_2 = 2(1-m^2)^{n-2} (1+m^2) - 4\beta(1-m^2)^{2n-4} (D_1 - D_2) \quad , \tag{3.39}$$

que é $n - 1$ vezes degenerado e não existe se $n = 1$, e

$$\lambda_3 = 2(1-m^2)^{n-2} \left[1 - (2n-1)m^2 \right] - 4\beta(1-m^2)^{2n-4} [D_1 - (n-1)D_2] \quad , \tag{3.40}$$

que é não degenerado.

Para a solução paramagnética ($m = 0$) estes autovalores ficam

$$\begin{aligned}\lambda_1 &= 2 \quad , \\ \lambda_2 &= \lambda_3 = 2 - 4\beta \quad ,\end{aligned}$$

os quais são todos positivos somente quando $\beta \leq 1/2$, ou seja, $T \geq 2$. Quando $n = 1$ e $m \neq 0$, λ_2 não existe e λ_1 e λ_3 ficam:

$$\begin{aligned}\lambda_1 &= 2(1 - m^2) \\ \lambda_3 &= 2 - 4\beta \operatorname{sech}^2(2\beta m) \quad ,\end{aligned}$$

e ambos são positivos quando $\beta \geq 1/2$, ou seja, $T \leq 2$. Se $n \geq 2$, λ_2 é sempre negativo, como pode ser visto na figura 3.4, e vai a zero quando $\beta \rightarrow 1/2$. Em resumo: acima de $T_c = 2$, a única solução estável é a paramagnética, enquanto que se $T < 2$, somente as memórias próprias ($n = 1$) são estáveis, sendo os estados simétricos ($n \geq 2$) sempre instáveis. Esta instabilidade pode ser intuída da expressão da energia: um estado que se superpõe igualmente com duas ou mais memórias encontra-se no ponto médio entre elas, correspondendo, portanto, a um máximo. Esta temperatura é diferente daquela encontrada para o modelo de Hopfield, $T_c^H = 1$, porque este apresenta um fator $1/2$ no Hamiltoniano.

Como comentário final, é importante notar que a solução $n = 1$ é exata em $T = 0$, tanto para padrões correlacionados quanto para descorrelacionados. Esta solução pode ser escrita como

$$\mathbf{m} = (1, m_{12}, m_{13}, \dots, m_{1P}) \quad ,$$

onde $|m_{1\nu}| = |\frac{1}{N} \sum_{i=1}^N \xi_i^1 \xi_i^\nu| \sim 1/\sqrt{N}$ é o overlap entre as memórias $\vec{\xi}^1$ e $\vec{\xi}^\nu$. Estas flutuações desaparecem quando o número de padrões armazenados é finito. No modelo de Hopfield são estas flutuações que destróem os mínimos (o ruído torna-se superior ao sinal) quando o número de memórias armazenadas é muito grande. Neste caso, quando as flutuações não podem ser desprezadas e a auto-mediação não vale, é que somos obrigados a utilizar o truque das réplicas ('replica trick'), o qual nos permite calcular a capacidade máxima da rede, e considerar o efeito da não ortogonalidade das memórias envolvidas. No modelo RS,

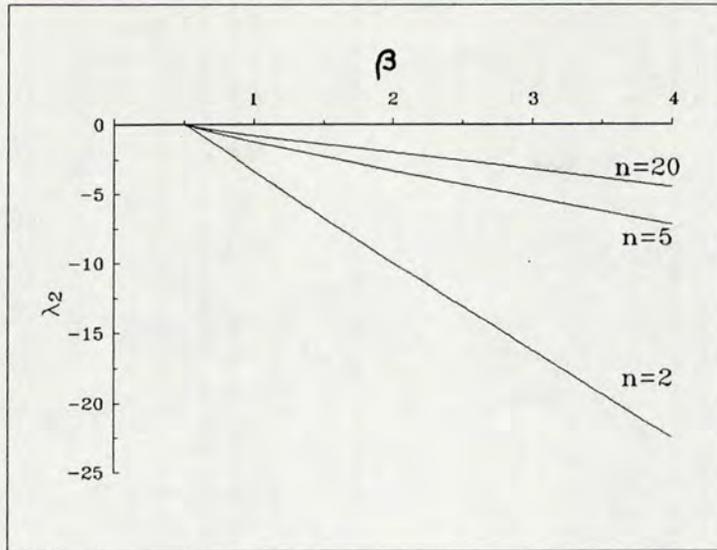


Figura 3.4: O autovalor λ_2 contra β para diferentes valores de $n > 1$. Este autovalor é sempre negativo na região $\beta > 1/2$, onde as soluções $m \neq 0$ existem.

quando a solução acima é colocada na eq.(3.23), o termo onde aparecem a contribuição das flutuações vai a zero, refletindo o fato de que as memórias são sempre estados de $E = 0$, o menor valor possível de energia, independentemente do número de memórias na rede. Por esta razão a solução acima é exata em $T = 0$, mesmo para α finito.

3.3 Teoria de Campo Médio Somente com Memórias: Caso OPS

As anti-memórias que no caso anterior foram explicitamente armazenadas e portanto representavam mínimos de E , apresentam no caso OPS estados de energia muito altos. Nesta situação, o espaço de fase aparenta ter a forma de uma concha com vários furos, cada um correspondendo a uma memória. As bacias de atração ao redor de cada mínimo são menores, embora estejam dentro de uma grande bacia comum cujo centro é o estado que corresponde a superposição idêntica com todas memórias, como veremos nos cálculos que seguem e pelas simulações numéricas [17,18]. Na figura 3.1 vimos uma representação

simplificada da estrutura do espaço de fase na configuração OPS. O estado simétrico que representa o centro da grande bacia é um mínimo local da energia e, quando o número de padrões armazenados vai a infinito, as barreiras entre as memórias e o estado central vão a zero, formando uma única bacia, cujo fundo plano ($E = 0$) é composto por todos os estados localizados entre as memórias e o estado simétrico. A origem deste estado central pode ser melhor entendida quando analisamos o primeiro termo da expansão (3.5), h_1 , o qual representa um campo $\vec{\mathbf{J1}}$ que tende a alinhar \vec{S} em sua direção e possui a seguinte forma:

$$h_1 = - \sum_i \mathbf{J1}_i S_i \quad ,$$

onde

$$\mathbf{J1}_i = \sum_{\mu=1}^P \xi_i^\mu \quad .$$

Portanto, o estado \vec{S} dado por

$$S_i = \text{sgn} \left(\sum_{\mu=1}^P \xi_i^\mu \right) \quad ,$$

minimiza aquele termo do Hamiltoniano. Este estado possui superposições iguais com todos os padrões e representa o ponto equidistante a todas as memórias (centro de massa).

Como na seção precedente, a função partição pode ser escrita na forma (3.19), com $f(\mathbf{m}, \mathbf{t})$ dada por

$$f(\mathbf{m}, \mathbf{t}) = \prod_{\mu=1}^P (1 - m_\mu) + \mathbf{m} \cdot \mathbf{t} - \frac{1}{\beta} \langle \langle \ln [2 \cosh (\beta \boldsymbol{\xi} \cdot \mathbf{t})] \rangle \rangle \quad . \quad (3.41)$$

As equações de ponto de sela são:

$$t_\mu = \prod_{\substack{\nu=1 \\ \nu \neq \mu}}^P (1 - m_\nu) \quad (3.42)$$

e

$$m_\mu = \langle \langle \xi^\mu \tanh (\beta \boldsymbol{\xi} \cdot \mathbf{t}) \rangle \rangle \quad . \quad (3.43)$$

Substituindo (3.42) em (3.43), resulta

$$m_\mu = \left\langle \left\langle \xi^\mu \tanh \left[\beta \sum_{\nu=1}^P \xi^\nu \prod_{\substack{\lambda=1 \\ \lambda \neq \nu}}^P (1 - m_\lambda) \right] \right\rangle \right\rangle \quad . \quad (3.44)$$

A solução que vamos considerar é do tipo:

$$\mathbf{m} = (m, \epsilon, \dots, \epsilon) \quad (3.45)$$

Note-se que esta solução, ao contrário da sugerida para o caso PAS, não permite estados de Mattis, isto é, superposição não nula com apenas uma das memórias, exceto no limite em que $T \rightarrow 0$. Isto ocorre devido à existência, no espaço de fases, de um mínimo no ponto correspondente ao centro de massa das memórias. Quando $T = 0$, o sistema pode se encontrar no fundo da bacia de alguma memória e o estado possui superposição 1 com este padrão e 0 (na verdade $1/\sqrt{N}$) com os restantes. Quando $T \neq 0$, ocorrem flutuações as quais são mais freqüentes na direção do mínimo central do que no sentido oposto. Isto se deve à maior inclinação das paredes da bacia de atração da memória na direção dos estados altamente energéticos (anti-memórias) do que na direção do centro de massa, fazendo com que os estados nesta direção sejam mais visitados, por possuírem menor energia. Ao efetuarmos uma média temporal destas flutuações surgirá uma pequena superposição, não nula, com as outras memórias (ϵ). No caso PAS, como o espaço é isotrópico, as flutuações são equiprováveis, ou seja, ocorrem tantas flutuações na direção de uma memória quanto na direção oposta. Portanto, a média temporal da superposição com as memórias anula-se, originando a solução com $\epsilon = 0$ mesmo quando $T \neq 0$.

Substituindo a eq.(3.45) na eq.(3.44) obtemos

$$m = \left\langle\left\langle \xi^1 \tanh \beta \left[\xi^1 (1 - \epsilon)^{P-1} + (1 - m)(1 - \epsilon)^{P-2} \sum_{\mu=2}^P \xi^\mu \right] \right\rangle\right\rangle \quad (3.46)$$

e

$$\epsilon = \left\langle\left\langle \xi^\nu \tanh \beta \left[\xi^1 (1 - \epsilon)^{P-1} + (1 - m)(1 - \epsilon)^{P-2} \sum_{\mu=2}^P \xi^\mu \right] \right\rangle\right\rangle, \quad \nu \neq 1. \quad (3.47)$$

Substituindo a média $\langle\langle \rangle\rangle$ pela média sobre $z_1 = \xi^1$ e $z_{P-1} = \sum_{\mu=2}^P \xi^\mu$, com os devidos pesos temos, para padrões não correlacionados:

$$m = \frac{1}{2^P} \sum_{\ell=0}^1 \sum_{\ell'=0}^{P-1} \frac{(P-1)!}{\ell'!(P-1-\ell')!} (1-2\ell) \tanh \beta \left[(1-2\ell)(1-\epsilon)^{P-1} + (P-1-2\ell')(1-m)(1-\epsilon)^{P-2} \right] \quad (3.48)$$

e

$$\epsilon = \frac{1}{2^P(P-1)} \sum_{\ell=0}^1 \sum_{\ell'=0}^{P-1} \frac{(P-1)!}{\ell!(P-1-\ell)!} (P-1-2\ell').$$

$$\tanh \beta \left[(1-2\ell)(1-\epsilon)^{P-1} + (P-1-2\ell')(1-m)(1-\epsilon)^{P-2} \right] , \quad (3.49)$$

onde $z_1 = 1 - 2\ell$ e $z_{P-1} = P - 1 - 2\ell'$.

As soluções do sistema de equações transcendentais acopladas acima são obtidas numericamente pelo método de Newton-Raphson [82] e são mostradas na figura 3.5 para diferentes valores de P . Para cada valor de P há dois ramos na figura, um correspondendo à solução simétrica $m = \epsilon$, que existe em qualquer temperatura e outro para as memórias próprias ($m \approx 1$ e $\epsilon \approx 0$) que existe somente abaixo de uma certa temperatura $T(P)$. Esta temperatura $T(P)$, a partir da qual as memórias deixam de ser solução de (3.48)-(3.49) é mostrada, em função de P , na figura 3.6. A descontinuidade que surge quando $P > 2$ no parâmetro de ordem é típica de uma transição de primeira ordem. Para $P = 1$ esta temperatura diverge, pois na verdade só temos uma solução, que corresponde ao único mínimo na energia. Para $P = 2$ podemos calcular analiticamente $T(2)$. Subtraindo (3.48) de (3.49) e, para $P = 2$, tem-se:

$$m - \epsilon = \tanh [\beta(m - \epsilon)] .$$

A solução trivial $m - \epsilon = 0$, ou seja, $m = \epsilon$, sempre existe. Para que exista uma solução não trivial com $m - \epsilon \neq 0$ é preciso que

$$\left[\frac{\partial}{\partial(m - \epsilon)} \tanh [\beta(m - \epsilon)] \right]_{m-\epsilon=0} \geq 1 ,$$

ou seja, $\beta \geq 1$. Portanto, $T(2) = 1$, e a transição entre as duas soluções é contínua (o parâmetro de ordem m que caracteriza as diferentes fases muda suavemente em $T(2)$). Para $P > 2$, $T(P)$ é obtida numericamente, e decai monotonicamente com P , o que significa o seguinte: suponhamos que o sistema esteja, em $T = 0$, em uma das memórias armazenadas, isto é, no fundo da bacia de atração associada com ela. À medida que o ruído (temperatura) na rede aumenta, o estado do sistema flutua em torno daquela memória, sendo a amplitude da flutuação proporcional ao ruído. Em uma determinada temperatura,

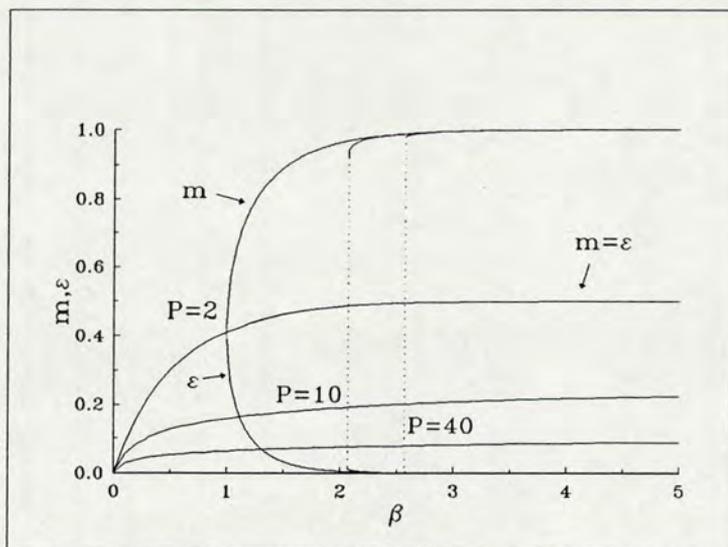


Figura 3.5: As soluções m e ϵ das eqs.(3.48)-(3.49) contra β para vários valores de P .

$T(P)$, as flutuações tornam-se grandes o suficiente para tirar o sistema de dentro da bacia de atração da memória, fazendo com que caia na bacia comum e estabilize no estado simétrico. Como $T(P)$ decai com P , quanto maior é o número de padrões armazenados, menor é a temperatura necessária para desestabilizar as memórias já que as barreiras de energia que separam as memórias do estado simétrico central são menores e também menores são suas bacias de atração. Assim, o tamanho das bacias é determinado por P . Este resultado é corroborado pelas simulações numéricas [17,18] apresentadas no capítulo seguinte. Para $T > T(P)$ a única solução é a simétrica, $\mathbf{m} = (m, \dots, m)$, que corresponde às flutuações ao redor do centro da grande bacia. Em $T = 0$, a solução simétrica pode ser obtida exatamente do mesmo modo que as eqs.(3.33) e (3.34):

$$m = \frac{(P-1)!}{2^{P-1} \left[\left(\frac{P-1}{2} \right)! \right]^2} \quad , \quad (3.50)$$

para P ímpar, e

$$m = \frac{(P-1)!}{2^{P-1} \left(\frac{P}{2} \right)! \left(\frac{P}{2} - 1 \right)!} \quad , \quad (3.51)$$

quando P é par. Como antes, ambas se comportam como $1/\sqrt{P}$ à medida que P cresce

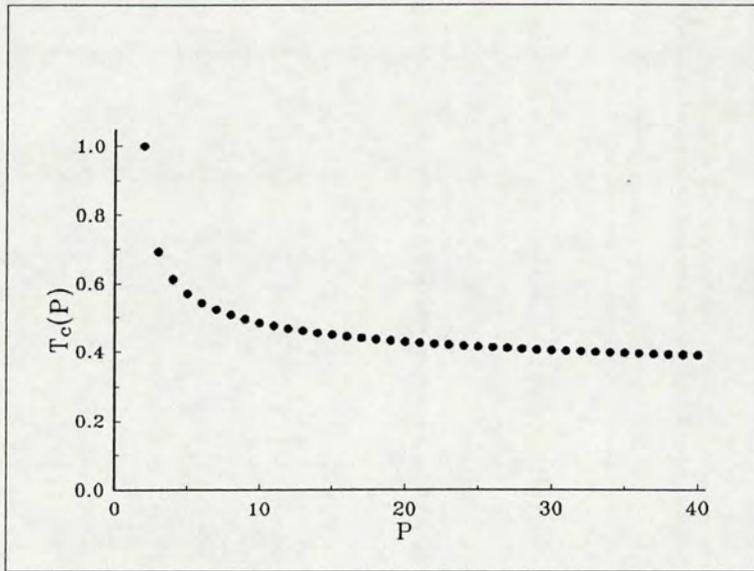


Figura 3.6: A temperatura $T(P)$ contra P na qual as soluções $m \neq \epsilon$ deixam de existir. Esta temperatura é uma função decrescente de P .

muito.

Para estudar a estabilidade de tais soluções precisamos calcular os elementos da matriz estabilidade $A_{\mu\nu} = \partial^2 f / \partial m_\mu \partial m_\nu$ e obter seus autovalores. Após algumas manipulações algébricas, estes elementos são:

$$A_{\mu\nu} = -(1 - \delta_{\mu\nu}) \prod_{\lambda \neq \mu, \nu}^P (1 - m_\lambda) - \beta \sum_{\lambda \neq \mu}^P \sum_{\theta \neq \nu}^P T_{\lambda\theta} \left[\prod_{\kappa \neq \lambda, \mu}^P (1 - m_\kappa) \right] \left[\prod_{\eta \neq \theta, \nu}^P (1 - m_\eta) \right] , \quad (3.52)$$

onde

$$T_{\lambda\theta} = \left\langle \left\langle \xi^\lambda \xi^\theta \operatorname{sech}^2 \left[\beta \sum_{\kappa=1}^P \xi^\kappa \prod_{\substack{\eta=1 \\ \eta \neq \kappa}}^P (1 - m_\eta) \right] \right\rangle \right\rangle . \quad (3.53)$$

Esta matriz tem a seguinte forma

$$A = \begin{pmatrix} A_{1D} & A_{1FD} & A_{1FD} & \dots & A_{1FD} \\ A_{1FD} & A_D & A_{FD} & \dots & A_{FD} \\ A_{1FD} & A_{FD} & A_D & \dots & A_{FD} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{1FD} & A_{FD} & A_{FD} & \dots & A_D \end{pmatrix}$$

onde os elementos são:

$$A_{1D} = \frac{-\beta(1-\epsilon)^{P-1}}{2^P} \sum_{\ell=0}^1 \sum_{\ell'=0}^1 \sum_{\ell''=0}^{P-2} \frac{(P-2)!}{\ell''!(P-2-\ell'')!} [P-1-2(\ell'+\ell'')]^2 \text{sech}^2\Theta_1 \quad , \quad (3.54)$$

$$A_{1FD} = \epsilon - 1 - \frac{\beta(1-\epsilon)^{P-2}}{2^P} \sum_{\ell=0}^1 \sum_{\ell'=0}^1 \sum_{\ell''=0}^{P-2} \frac{(P-2)!}{\ell''!(P-2-\ell'')!} [P-1-2(\ell'+\ell'')] [(1-\epsilon)(1-2\ell) + (1-m)(P-2-2\ell'')] \text{sech}^2\Theta_1 \quad , \quad (3.55)$$

$$A_D = -\frac{\beta(1-\epsilon)^{P-3}}{2^P} \sum_{\ell=0}^1 \sum_{\ell'=0}^1 \sum_{\ell''=0}^{P-2} \frac{(P-2)!}{\ell''!(P-2-\ell'')!} [(1-\epsilon)(1-2\ell) + (1-m)(P-2-2\ell'')]^2 \text{sech}^2\Theta_1 \quad , \quad (3.56)$$

e

$$A_{FD} = -(1-m) - \frac{\beta(1-\epsilon)^{P-3}}{2^P} \sum_{\ell=0}^1 \sum_{\ell'=0}^1 \sum_{\ell''=0}^1 \sum_{\ell'''=0}^{P-3} \frac{(P-3)!}{\ell'''!(P-3-\ell''')!} \{ (1-\epsilon)^2 + (1-m) [P-2-2(\ell'+\ell''')] \times [2(1-\epsilon)(1-2\ell) + (1-m)(P-2-2(\ell''+\ell'''))] \} \text{sech}^2\Theta_2 \quad , \quad (3.57)$$

onde Θ_1 e Θ_2 são definidos como

$$\Theta_1 = \beta(1-\epsilon)^{P-2} \{ (1-2\ell)(1-\epsilon) + (1-m)[P-1-2(\ell'-\ell'')] \} \quad , \quad (3.58)$$

e

$$\Theta_2 = \beta(1-\epsilon)^{P-2} \{ (1-2\ell)(1-\epsilon) + (1-m)[P-1-2(\ell'+\ell''+\ell''')] \} \quad . \quad (3.59)$$

Os autovalores de uma matriz com tal estrutura são obtidos no apêndice B e valem

$$\lambda_1 = A_D - A_{FD} \quad , \quad (3.60)$$

$$\lambda_2 = \frac{1}{2} [A_{1D} + A_D + (P-2)A_{FD}] + \frac{1}{2} \sqrt{ [A_{1D} - A_D - (P-2)A_{FD}]^2 + 4(P-1)A_{1FD}^2 } \quad (3.61)$$

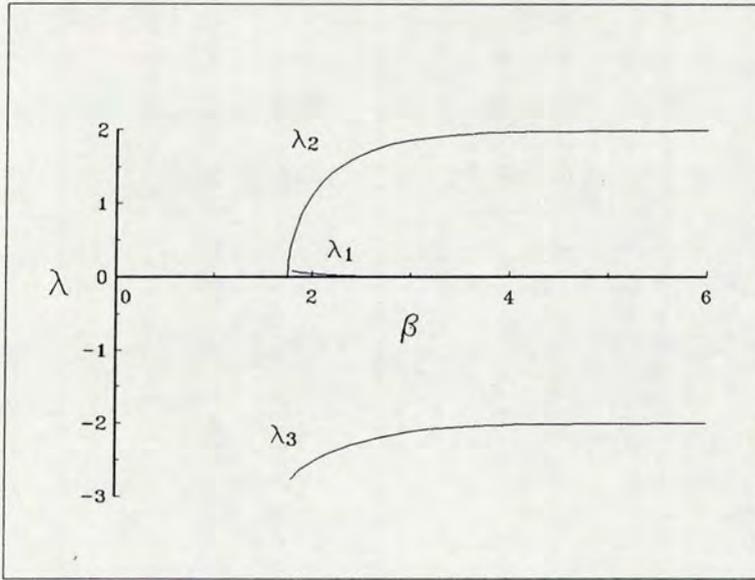


Figura 3.7: Autovalores da matriz estabilidade versus β para a solução $m \neq \epsilon$ e $P = 5$. O autovalor λ_1 é sempre positivo e λ_3 sempre negativo. Já λ_2 vai a zero à medida que T vai a $T(P)$. Este comportamento é análogo para todo valor de P .

e

$$\lambda_3 = \frac{1}{2} [A_{1D} + A_D + (P - 2)A_{FD}] - \frac{1}{2} \sqrt{ [A_{1D} - A_D - (P - 2)A_{FD}]^2 + 4(P - 1)A_{1FD}^2 } \quad (3.62)$$

O autovalor λ_1 é $P - 2$ vezes degenerado enquanto que λ_2 e λ_3 não o são. O gráfico destes autovalores para $P = 5$ é mostrado na figura 3.7 em função de β . O comportamento é análogo para todo $P > 1$. O autovalor λ_2 vai a zero à medida que T se aproxima de $T(P)$, a temperatura na qual a solução $m \neq \epsilon$ deixa de existir, como pode ser visto na figura 3.8 para diferentes valores de P . Portanto, onde a solução $m \neq \epsilon$ existe, nenhum dos três autovalores muda de sinal. Para $T = 0$ ($\beta \rightarrow \infty$) as memórias são claramente estáveis, o que implica que, para todo o intervalo de temperatura em que ela existe, são estáveis. O mesmo método de calibração dos autovalores foi usado por Gardner na referência [83].

Para a solução simétrica somente precisamos fazer $m = \epsilon$ nas eqs.(3.54)-(3.59). Neste

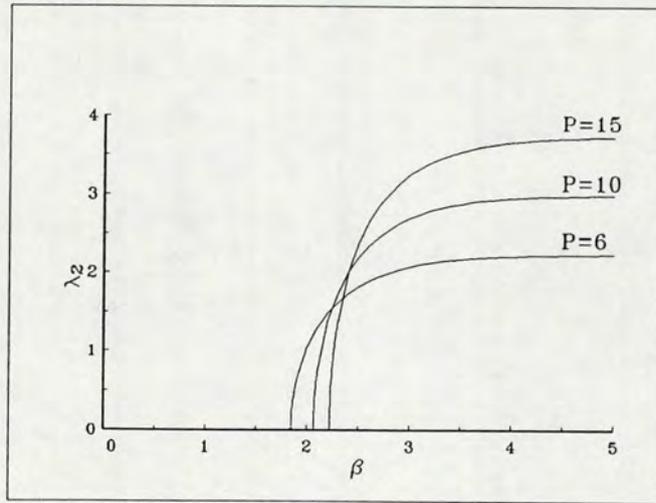


Figura 3.8: O autovalor λ_2 versus β para vários valores de P .

caso, $A_{1D} = A_D$ e $A_{1FD} = A_{FD}$ e os autovalores ficam

$$\lambda'_1 = 1 - \frac{4\beta(1-m)^{P-2}}{2^P P(P-1)} \sum_{\ell=0}^P \frac{P!}{\ell!(P-\ell)!} \ell(P-\ell) \operatorname{sech}^2[\beta(1-m)^{P-1}(P-2\ell)] \quad , \quad (3.63)$$

que é $P-1$ vezes degenerado e

$$\lambda'_2 = 1 - P - \frac{\beta(P-1)^2(1-m)^{P-2}}{P2^P} \sum_{\ell=0}^P \frac{P!}{\ell!(P-\ell)!} (P-2\ell)^2 \operatorname{sech}^2[\beta(1-m)^{P-1}(P-2\ell)] \quad , \quad (3.64)$$

que é não degenerado e sempre negativo para qualquer temperatura. Como mostra a figura 3.9, λ'_1 pode trocar de sinal dependendo de se P é par ou ímpar. Para $T \rightarrow \infty$, quando $m \rightarrow 0$, esta solução é equivalente à paramagnética (desordem) e é claramente estável. Portanto, se P é ímpar, como λ'_1 não muda de sinal, ela é sempre estável. Se P for par, pode se tornar instável a uma temperatura muito baixa T_c^S , a qual pode ser obtida da eq.(3.63) como segue: a principal contribuição dos termos envolvidos no somatório vem de $\ell = P/2$, pois a temperatura onde λ'_1 troca de sinal é muito baixa, e este termo é o único no qual o argumento da secante hiperbólica é nulo. Ficamos então com:

$$\lambda'_1 \simeq 1 - \frac{\beta(1-m)^{P-2} P^2 (P-2)!}{2^P \left[\left(\frac{P}{2}\right)! \right]^2} \quad , \quad (3.65)$$

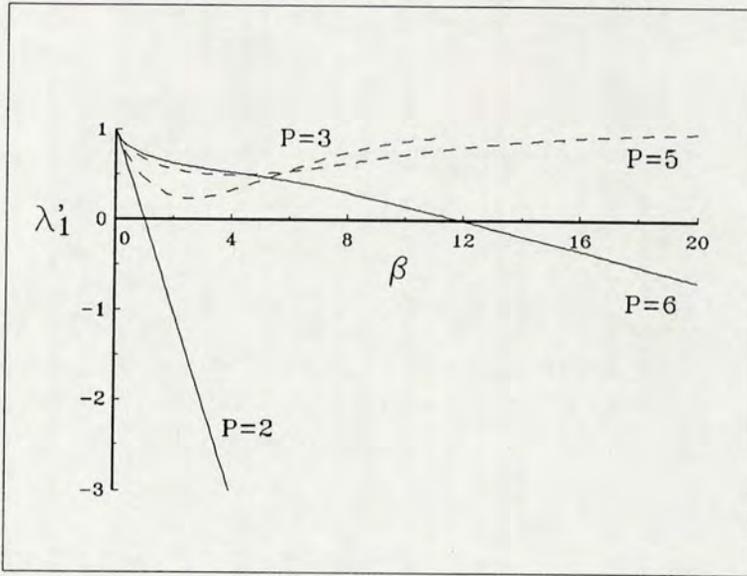


Figura 3.9: O autovalor λ'_1 contra β para vários valores de P . Este autovalor é sempre positivo se P é ímpar e troca de sinal em T_c^S quando P é par.

que é nulo quando T vale

$$T_c^S \simeq \frac{P^2(P-2)!(1-m)^{P-2}}{2^P \left[\left(\frac{P}{2} \right)! \right]^2}, \quad (3.66)$$

onde m é dado, com excelente aproximação, pelo limite $T = 0$ da eq.(3.51). O comportamento linear da eq.(3.65) é claro na fig. 3.9. Para temperaturas altas, o sistema está na bacia principal e a solução estável é a simétrica, isto é, flutua ao redor do estado que se superpõe simetricamente com todas as memórias. À medida que a temperatura baixa, o sistema tem duas alternativas: ou ele entra na bacia de atração de alguma memória ou continua na bacia principal, flutuando ao redor de seu centro. O estado que define o centro da bacia é dado pela regra da maioria

$$S_i^S = \operatorname{sgn} \left(\sum_{\mu=1}^P \xi_i^\mu \right) .$$

Finalmente, quando P é par, em vários sítios o argumento da função sinal é nulo e ela não define o estado do neurônio, sendo razoável que o centro da bacia não seja um estado estável. A diferença entre P par e ímpar só é notável quando P é muito pequeno, pois

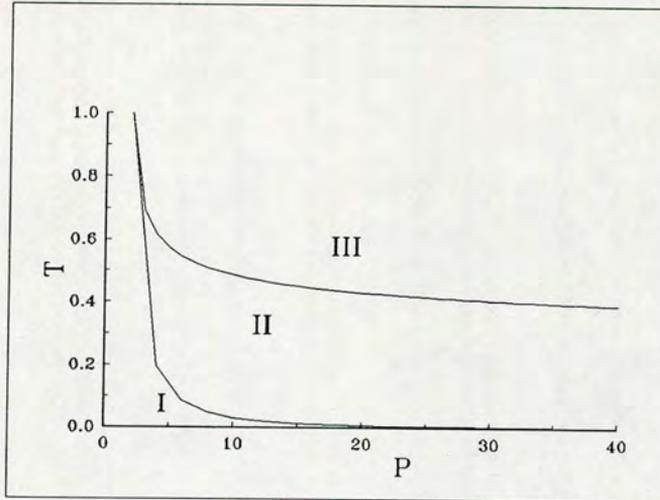


Figura 3.10: Diagrama de fases para a rede OPS. Na região III apenas a solução simétrica ($m = \epsilon$) é estável; na região II tanto esta solução quanto as memórias próprias são estáveis e na região I, as memórias próprias são estáveis e a solução simétrica é instável somente quando P é par.

quando $P \rightarrow \infty$, todas as bacias individuais de atração desaparecem e a única solução estável é a simétrica. Quando P é pequeno, o número de conexões de primeira ordem (J_{1i}) nulas é grande, isto é, o campo que tende a alinhar o estado da rede ao longo da direção que define o centro da bacia simétrica não está bem definido, o que justifica a instabilidade desta solução em $T = 0$. Quanto maior o número de memórias armazenadas P , menor é o número de elementos nulos no campo e portanto a solução se torna estável.

Resumindo, a situação para sistemas OPS pode ser visualizada no seguinte diagrama T versus P (fig. 3.10): na região III, onde T é muito alta, somente a solução simétrica é estável, na região II tanto a solução simétrica quanto as memórias próprias são estáveis e na região I, onde a temperatura é muito baixa, os estados simétricos (para P par) se tornam instáveis.

3.4 Conclusão

Neste capítulo introduzimos o modelo RS e apresentamos os principais resultados obtidos analiticamente. Foram consideradas duas situações de interesse: quando as memórias e as anti-memórias são aprendidas (caso PAS) e quando somente são consideradas as memórias (caso OPS).

Na situação PAS o sistema funciona muito bem como memória associativa e, além disso, o número de estados espúrios é drasticamente reduzido, pelos termos de multi-interação, o que faz com que o espaço ao redor de cada memória, isto é, sua bacia de atração, seja suave e isotrópico. Por outro lado, os sistemas OPS funcionam satisfatoriamente quando o número de padrões armazenados P é finito, resultado que é encontrado analiticamente e comprovado, como veremos a seguir, nas simulações em computador. Este comportamento indica a importância do armazenamento das anti-memórias e sua origem está na forma característica do espaço de fase no caso OPS: uma grande bacia com pequenos furos, cada um deles correspondendo a uma memória. Quando as anti-memórias são também guardadas, todas as componentes $\mathbf{J}\mathbf{1}_i$ do campo que tende a alinhar o estado \vec{S} com o estado central são nulas. Além disso, a região que antes compreendia altas energias agora possui mínimos de energia zero (as anti-memórias), o que destrói a estrutura da grande bacia.

Capítulo 4

Simulação Numérica do Modelo RS

No capítulo precedente estudamos as propriedades de equilíbrio do modelo RS, nas possíveis configurações PAS e OPS, por meio das técnicas da Mecânica Estatística. Agora vamos analisar algumas propriedades dinâmicas do modelo através de simulações numéricas. O desempenho de um modelo pode ser analisado em termos de várias quantidades, como por exemplo, o tamanho das bacias de atração (seção 4.1) de cada memória, que está associado com a informação mínima necessária para recuperar uma memória e o tempo de convergência (seção 4.2) necessário para atingir uma memória após o sistema ter sido colocado na sua bacia de atração, que, por sua vez, se relaciona com a isotropia das bacias. Na seção 4.3 resumimos e concluímos.

4.1 Algoritmo

O algoritmo utilizado nas simulações que seguem é o seguinte:

i) Construimos P padrões com atividade a (probabilidade de um dado neurônio ser +1) e um estado inicial que possui uma superposição m_0 com um dos padrões escolhidos ao acaso. Calculamos as superposições m_μ com todas as memórias e então a energia inicial do sistema;

ii) Consideramos uma inversão virtual no i -ésimo neurônio e calculamos as superposições

virtuais por

$$m_{\mu}^* = m_{\mu} - \frac{2}{N} S_i \xi_i^{\mu} \quad ;$$

- iii) Obtemos a energia virtual E^* usando m_{μ}^* ;
- iv) Sempre que $E^* \leq E$ efetuamos a inversão, atualizando a energia e as superposições;
- v) Procedemos seqüencialmente desde o primeiro até o último neurônio, várias vezes, até que um estado estável seja alcançado.

No passo v, a atualização é feita seqüencialmente porque os resultados são os mesmos se escolhermos aleatoriamente o neurônio a ser invertido. A atividade a dos padrões nos diz qual fração da rede está ativa, ou seja, possui neurônios no estado +1. Atividades diferentes de 0.5 são usadas para introduzir correlações entre os padrões e, em particular, padrões com baixa atividade são importantes por várias razões. Entre elas a de que existem evidências experimentais de que o cérebro possui atividade menor do que 0.5. Por outro lado, em quase todas situações práticas, os padrões a serem armazenados possuem um pano de fundo comum (e geralmente irrelevante) sendo, portanto, correlacionados.

4.2 Bacias de Atração

Para estimar os tamanhos das bacias de atração seguimos o procedimento de Forrest [43], sendo que todas as simulações foram realizadas com ruído nulo ($T = 0$). Medimos a fração $f(m_0)$ de vezes em que uma memória é recuperada a partir de um estado inicial que possui uma superposição m_0 com ela. Quanto maior é m_0 , maior é a quantidade de informação que fornecemos a respeito da memória escolhida, ou seja, mais próximos estamos dela. Se o estado inicial converge para a memória escolhida é porque ele foi colocado dentro de sua bacia de atração. Para obter resultados confiáveis, repetimos o procedimento várias vezes de tal modo que os erros estatísticos sejam cada vez menos importantes. Como toda simulação é efetuada com redes finitas, é preciso simular vários tamanhos de rede e tentar alguma extrapolação para $N \rightarrow \infty$.

A figura 4.1 mostra $f(m_0)$ para redes PAS com atividade 0.5 (padrões não correlaciona-

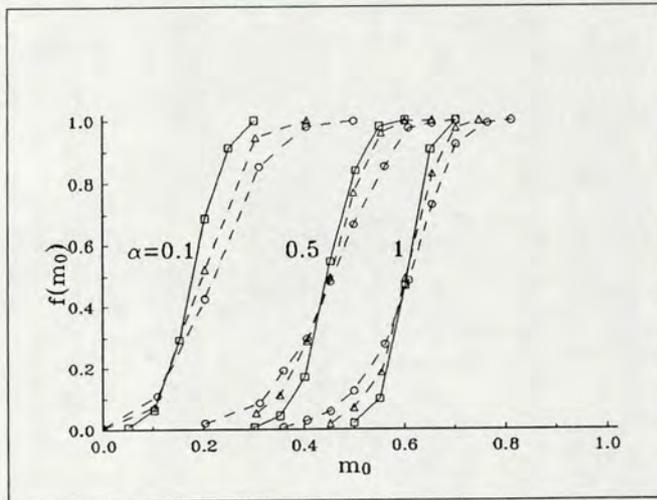


Figura 4.1: Fração $f(m_0)$ de padrões recuperados com uma superposição inicial m_0 para $N = 128(\circ)$, $256(\triangle)$ e $512(\square)$ e $\alpha = 0.1, 0.5$ e 1 no caso PAS, e padrões descorrelacionados.

dos). As redes possuem 128, 256 e 512 neurônios e as simulações foram feitas com $\alpha = 0.1, 0.5$ e 1 . Para cada ponto do gráfico foram feitas 1000 relaxações usando 5 diferentes conjuntos de memórias, exceto para a rede de 512 elementos onde foram feitas 250 relaxações. Como as memórias são sempre mínimos da energia, a superposição final, quando a memória é recuperada, é sempre 1, ou seja, não há erros no estado recuperado. Como no modelo de Hopfield, as curvas correspondendo ao mesmo α se superpõem e, à medida que N aumenta, a inclinação da curva também aumenta. A diferença é que no modelo de Hopfield, se $\alpha > \alpha_c \simeq 0.14$, a inclinação diminui com N e aqui, até $\alpha \simeq 1$, este efeito não existe. Quando N aumenta, as curvas de $f(m_0)$ se aproximam de uma função degrau, isto é

$$f(m_0) = \Theta(m_0 - m_c) \quad , \quad N \rightarrow \infty \quad ,$$

onde a função Θ vale 1 se o argumento é positivo e 0 se é negativo. O valor de $m_c(\alpha)$, que é obtido através da extrapolação para $N \rightarrow \infty$ da figura 4.1 (é o ponto de giro das curvas), nos dá uma estimativa do tamanho das bacias de atração. Da figura, obtemos os dados da tabela 3.1:

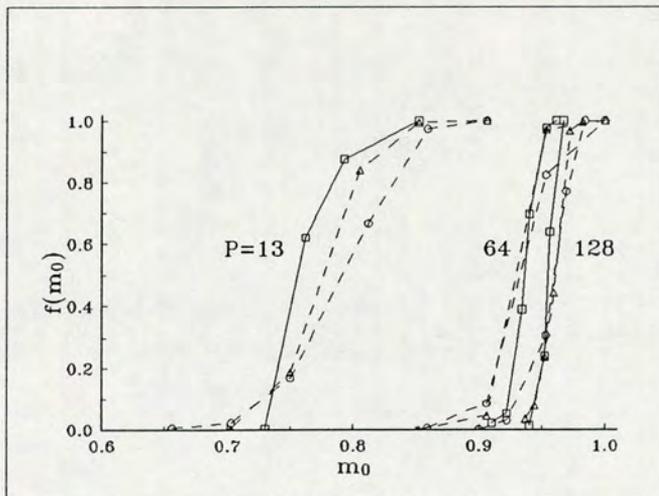


Figura 4.2: $f(m_0)$ contra m_0 para $N = 128(o)$, $256(\Delta)$ e $512(\square)$ e $P = 13, 64$ e 128 para o caso OPS com padrões descorrelacionados. Note a diferente escala no eixo horizontal.

α	m_c
0.1	0.12
0.5	0.43
1	0.61

Tabela 3.1

Quanto maior o valor de $m_c(\alpha)$, menor é a bacia de atração, pois é necessário uma quantidade maior de informação no estado inicial para que a memória seja recuperada. Comparando estes resultados com aqueles de Forrest [43] para o modelo de Hopfield, mostrados na seção 2.3.4, (p.ex., $m_c^H(\alpha = 0.1) \simeq 0.37$), podemos ver que os tamanhos das bacias de atração são significativamente maiores. Outro efeito visível é a dependência de m_c com α : no modelo RS as bacias diminuem muito menos quando α cresce.

Continuando nosso estudo, na figura 4.2 apresentamos $f(m_0)$ para redes OPS com atividade 0.5 e cujos tamanhos são $N = 128, 256$ e 512 neurônios. Como no caso anterior, são feitas 1000 relaxações, exceto na rede de 512 onde são feitas 250. Aqui, ao contrário do

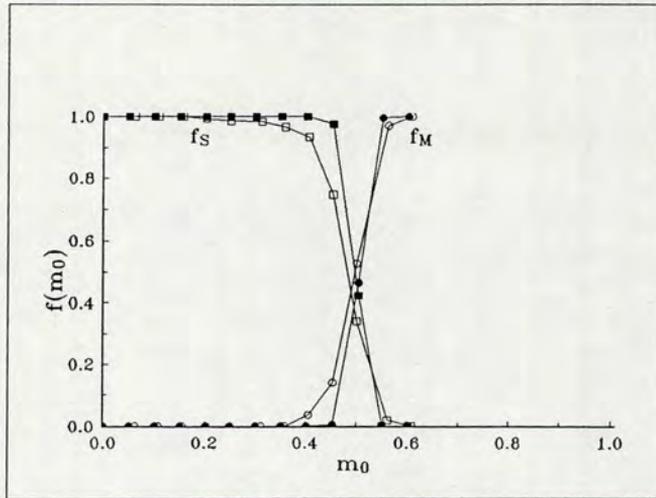
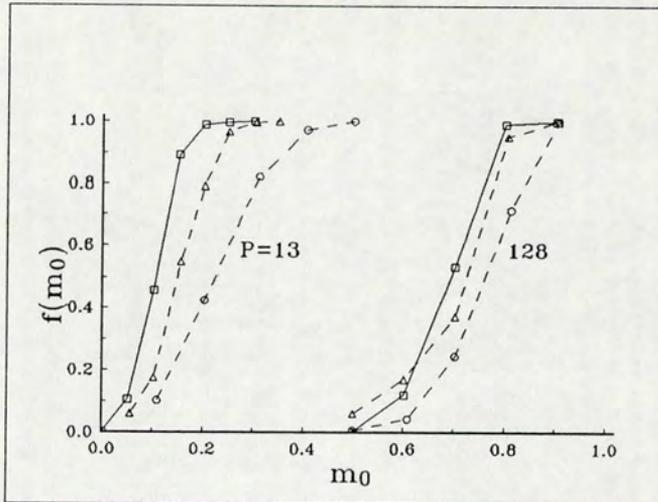


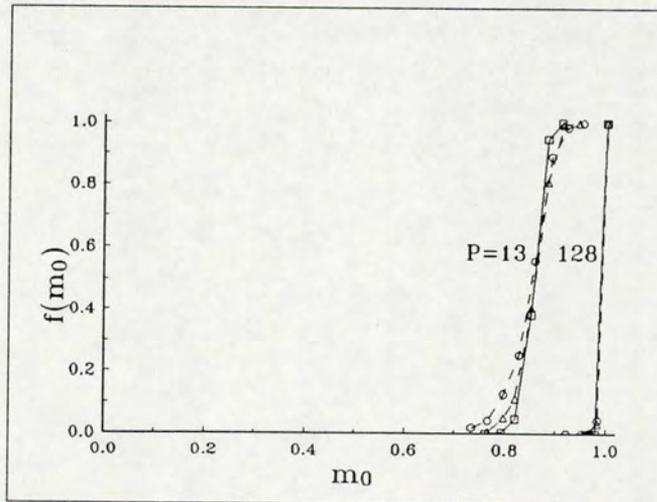
Figura 4.3: A fração de vezes com que o estado central é recuperado, f_S (\square e \blacksquare), contra m_0 para $N = 128$ e 512 (símbolos hachurados). Os símbolos \circ e \bullet representam a fração de vezes em que uma das memórias é recuperada. O número de memórias é $P = 5$.

caso PAS, as curvas com diferentes N se superpõem para o mesmo valor de P (e não α). Isto comprova os resultados da seção anterior, onde mostramos que as bacias associadas às memórias tornam-se cada vez menores à medida que P aumenta, independentemente do tamanho da rede. Esta dependência no número de padrões armazenados é um reflexo direto do fato das anti-memórias não estarem armazenadas originando, assim, a forma de uma grande bacia do espaço de fase. Novamente, a taxa com que as curvas aumentam desde 0 até 1 aumenta à medida que o tamanho da rede cresce. A existência de um mínimo local localizado no centro de massa das memórias pode ser vista na figura 4.3 onde, além da fração de memórias recuperadas, apresentamos a fração f_S de vezes com que este estado central, construído pela regra da maioria como explicado no capítulo anterior, é recuperado.

As figuras 4.4a 4.4b mostram os mesmos tipos de gráficos mas agora com atividade diferentes de 0.5, para o caso PAS. As curvas com mesmo P (e não com o mesmo α) se superpõem, como no caso OPS. Aqui também o efeito de N crescente é aumentar a inclinação das curvas. Acontece que agora o espaço de fase possui uma região de energias



a)



b)

Figura 4.4: $f(m_0)$ contra m_0 para o caso PAS com atividade a) $a = 0.4$ e b) $a = 0.2$ para dois valores diferentes de P e para $N = 128$ (o), 256 (Δ) e 512 (\square).

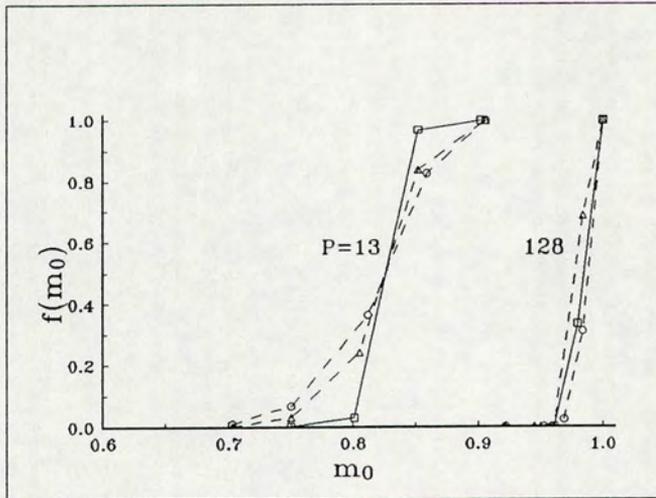
altas, ocupada pelos estados com atividade 0.5, adquirindo assim uma forma análoga ao caso OPS. Esta região separa duas grandes bacias, cada uma contendo ou as memórias ou as anti-memórias. Estas bacias têm uma estrutura similar às daquelas do caso OPS. Outro efeito de introduzir correlação entre as memórias aparece quando diminuimos a atividade, ou seja, aumentamos a semelhança entre elas. As bacias tornam-se mais próximas umas das outras e mais informação é necessária para distinguir uma memória da outra, diminuindo, deste modo, os tamanhos das bacias.

As figuras 4.5a e 4.5b mostram os resultados para redes OPS e baixas atividades. O único efeito é a redução das bacias de atração, como esperado.

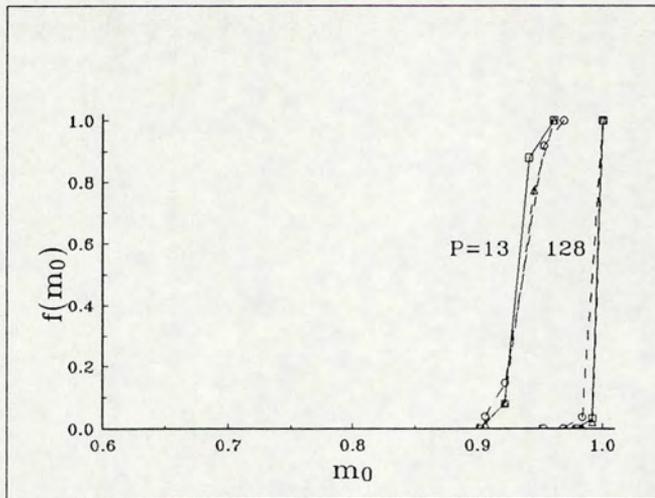
Resumindo a situação para o caso PAS, a capacidade de memória do sistema é incrementada, principalmente devido à existência de termos de interação entre mais de dois neurônios. Também a recuperação é sempre perfeita, ou seja, a superposição final é sempre 1, o que reflete o fato das memórias serem sempre mínimos da energia.

Para o caso OPS, a novidade é o fato do tamanho das bacias de atração dependerem, devido à forma do espaço de fase, do número de padrões armazenados. Estas bacias de atração são também pequenas, efeito que também é visto quando temos memórias correlacionadas, para ambas as configurações, PAS e OPS.

Quanto a uma estimativa do valor crítico de α , isto é, o número máximo de padrões por neurônio que se pode armazenar em redes descritas pelo modelo RS, podemos dizer que, pelos menos no intervalo varrido, nenhum efeito foi encontrado. Provavelmente, novamente pela existência de uma enorme quantidade de sinapses, este valor seja muito alto, o que tornaria a simulação muito dispendiosa em termos de tempo computacional. Além disso, como o modelo de Hopfield é reobtido quando expandimos a função energia do modelo RS até a segunda ordem, o valor de $\alpha_c \simeq 0.14$ pode ser tomado como um limite inferior para a capacidade de memória, ou seja, se cortarmos a expansão para o caso PAS na quarta, sexta, ..., $2k$ -ésima ordem, a capacidade de armazenamento da rede deve aumentar continuamente.



a)



b)

Figura 4.5: $f(m_0)$ contra m_0 para o caso OPS com atividade a) $a = 0.4$ e b) $a = 0.2$ para dois valores diferentes de P e para $N = 128(\circ)$, $256(\triangle)$ e $512(\square)$. Note a diferente escala no eixo horizontal.

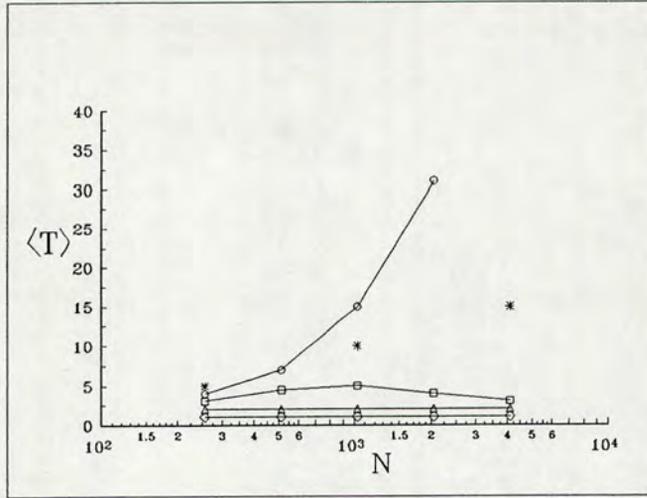


Figura 4.6: Tempo de convergência $\langle T \rangle$ como função de N com $\alpha = 0.1$ para diferentes superposições iniciais: $0.1(\circ)$, $0.2(\square)$, $0.3(\triangle)$ e $0.8(\diamond)$. Os dados representados por asteriscos (*), $m_0 = 0.35$, são para o modelo de Hopfield [45].

4.3 Tempo de Convergência

Outro aspecto dinâmico analisado foi o tempo de convergência médio $\langle T \rangle$, isto é, o número de passos dispendidos para o sistema atingir um estado estacionário. Vamos considerar o tamanho da rede como a unidade de tempo, ou seja, cada vez que percorremos seqüencialmente toda a rede, o tempo é acrescido de uma unidade. Também estimamos a dispersão no tempo de convergência, $\sigma = \langle T^2 \rangle - \langle T \rangle^2$, a qual está relacionada com a isotropia das bacias de atração ao redor de cada memória. Se o espaço é isotrópico e suave, $\langle T \rangle$ e σ devem ser ambos pequenos, ao passo que se tivermos estados metaestáveis (espúrios) ao redor das memórias, vão existir vários caminhos diferentes que levam desde o estado inicial até a memória, no fundo da bacia. Além do aumento no tempo de convergência, esta diversidade de caminhos provoca uma maior dispersão [18,44,45].

Foram simuladas redes com até 4096 neurônios, nas quais se mediu o tempo de convergência $\langle T \rangle$ e a dispersão σ para diferentes superposições iniciais m_0 , capacidades α e

atividades a dos padrões armazenados. $\langle T \rangle$ é obtido tanto para a fase de recuperação ($m_0 > m_c$) quanto fora dela ($m_0 < m_c$). O número de conjuntos de memórias e relaxações por conjunto depende do tamanho da rede e do valor de α . Por exemplo, para redes com $N = 4096$ foram utilizados 50 conjuntos e 10 relaxações para cada um, enquanto que para as menores redes foram usados 8 conjuntos e 50 relaxações.

Na figura 4.6 podemos ver que na fase de recuperação, poucos passos são necessários para atingir o estado estável, não dependendo do tamanho da rede. Esta característica é importante para aplicações práticas, uma vez que permite o reconhecimento após alguns poucos passos, rapidamente, como faz o cérebro humano e como deve fazer qualquer máquina artificial.

A figura 4.7 mostra a dispersão σ contra o tamanho da rede N para estados dentro da fase de recuperação, ou seja, dentro da bacia de atração. O fato das dispersões se anularem mostra a isotropia e suavidade do espaço de fase ao redor das memórias, uma vez que diferentes caminhos levando ao centro da bacia não implicam em tempos de convergência diferentes. Na figura 4.8 são mostrados os resultados para diferentes N e atividades a . Quando consideramos atividades diferentes de $a = 0.5$, isto é, quando correlações são introduzidas, $\langle T \rangle$ é menor (para a fase de não-recuperação) do que no caso descorrelacionado. Isto é devido ao fato das memórias estarem mais próximas umas das outras, o que faz com que a dinâmica leve rapidamente a uma delas ao invés de cruzar o espaço de fases. Finalmente na figura 4.9 são mostrados os resultados para diferentes valores de α . Na fase de recuperação o sistema necessita de poucos passos para atingir o estado de equilíbrio, enquanto que fora da fase de recuperação o tempo de convergência cresce com α .

Resumindo, se a recuperação é bem sucedida, o tempo de convergência é muito pequeno e depende somente das condições iniciais e do parâmetro α , independentemente do tamanho da rede. A dispersão destes tempos é também pequena, indicando que o espaço de fase é homogêneo e suave. Estes resultados candidatam o modelo para servir de base para futuras máquinas de reconhecimento, uma vez que a recuperação sendo rápida e não existindo estados meta-estáveis capazes de retardar ou mesmo enganar a dinâmica, a eficiência do

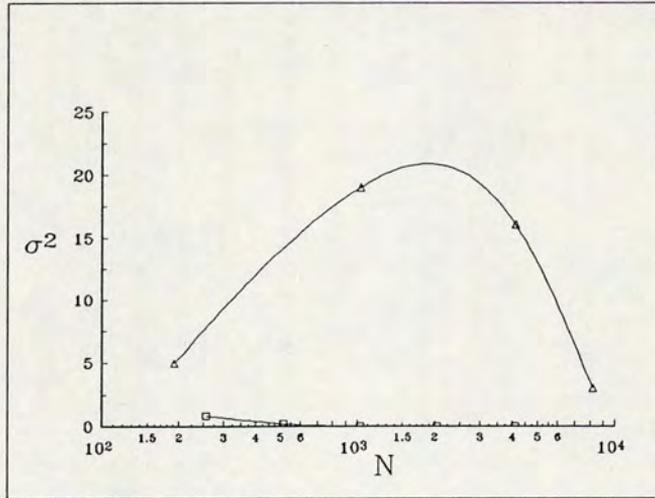


Figura 4.7: Variância da distribuição de $\langle T \rangle$ como uma função de N para os modelos de Hopfield (Δ) e RS (\square). A superposição inicial é $m_0 = 0.4$ e $\alpha = 0.1$.

algoritmo é muito alta.

4.4 Conclusão

Neste capítulo apresentamos os principais resultados obtidos por meio de simulações numéricas para o modelo RS. Foram consideradas duas situações de interesse: quando as memórias e as anti-memórias são aprendidas (caso PAS) e quando somente são consideradas as memórias (caso OPS).

Na situação PAS o sistema funciona muito bem como memória associativa, podendo armazenar uma quantidade muito grande de padrões, com tempos de recuperação muito pequenos. Além disso, o número de estados espúrios é drasticamente reduzido, o que faz com que o espaço ao redor de cada memória, isto é, sua bacia de atração, seja suave e isotrópico. Por outro lado, os sistemas OPS somente funcionam satisfatoriamente quando o número de padrões armazenados P é finito, resultado que é encontrado tanto analítica quanto numericamente. Este comportamento indica a importância do armazenamento das

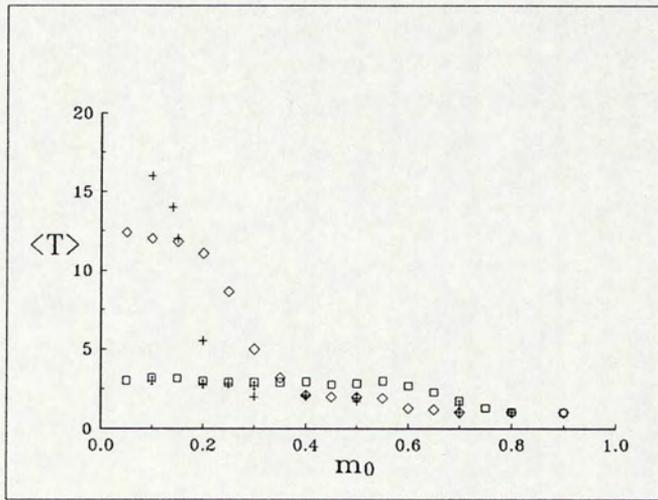


Figura 4.8: Tempo de convergência em função da superposição inicial para dois tamanhos de rede: $N = 512(a = 0.4(\square), 0.5(\diamond))$ e $N = 1024(a = 0.4(*), 0.5(+))$. Note que os tempos de convergência são menores na fase de não recuperação se introduzimos correlação. A capacidade é $\alpha = 0.2$.

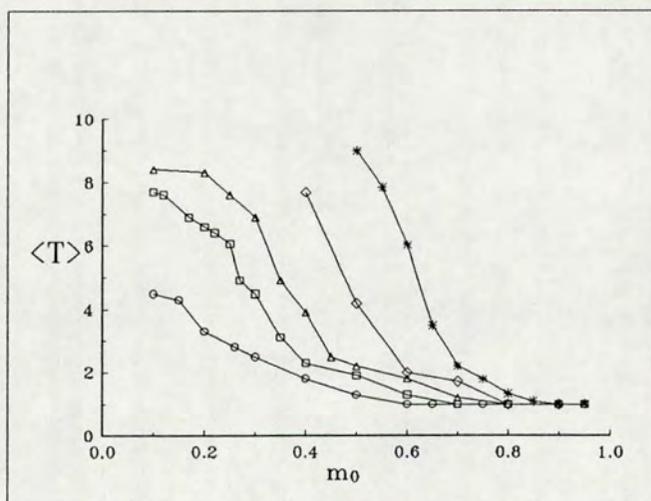


Figura 4.9: $\langle T \rangle$ para diferentes valores de α : 0.1 (o), 0.2 (□), 0.3 (△), 0.5 (◇) e 1 (*). A fase de recuperação pode ser identificada como sendo a região onde um ou dois passos são necessários para atingir o estado estável. As redes são de $N = 256$.

anti-memórias e sua origem está na forma característica do espaço de fase no caso OPS: uma grande bacia centrada no centro de massa das memórias, com pequenos furos, cada um deles correspondendo a um dos padrões armazenados.

No capítulo seguinte comparamos os resultados aqui obtidos com aqueles encontrados para outros modelos, especialmente o modelo de Hopfield e concluímos.æ

Capítulo 5

Conclusões

Os métodos da mecânica estatística são muito úteis no entendimento do comportamento cooperativo das redes de neurônios. Tal comportamento tem sido investigado por estes métodos fornecendo informações sobre as propriedades de estados estacionários para padrões randômicos típicos.

No capítulo 2 foram apresentados alguns resultados básicos para o modelo de Hopfield nos limites quando $\alpha = 0$ (seção 2.3.1) e $\alpha \neq 0$ (seção 2.3.2), bem como algumas informações que são extraídas através de simulações numéricas (2.3.4). Estes resultados mostram que este modelo, ao menos na sua forma original, não é muito eficiente no armazenamento de memórias (a capacidade é pequena), na recuperação destas (o tempo de convergência é grande e podem ocorrer erros no estado final) e na manipulação de padrões correlacionados (é incapaz de distinguir, por exemplo, o E e o F). Alguns destes problemas podem ser efetivamente melhorados substituindo-se a regra de Hebb por outras especialmente construídas para tratar destes problemas específicos.

5.1 Resumo dos resultados para o Modelo RS

O modelo RS foi criado tendo-se em mente um sistema onde as memórias nunca deixassem de ser mínimos absolutos de energia quando a carga da rede ultrapassasse um determinado valor. Isto definitivamente é o que ocorre com os outros modelos. Além desta

diferença fundamental, o modelo também possui algumas características que o aproximam dos análogos biológicos, como por exemplo, a assimetria das sinapses e a existência de interações n -árias entre os neurônios. Outra propriedade, a ordem das interações aumentar com o número de memórias armazenadas P , ainda carece de comprovação experimental, mas tal hipótese não deve ser descartada. A maneira compacta como estas multi-sinapses são tratadas pelo modelo facilita em muito o tratamento analítico, além de propiciar uma interpretação simples do espaço de fases em termos das distâncias entre as memórias e o estado da rede.

O modelo RS foi estudado em dois limites específicos: quando temos um conjunto de P padrões armazenados (caso OPS) e quando, além destes, as correspondentes anti-memórias são também guardadas (caso PAS). Ambos os casos foram estudados analiticamente no limite em que $\alpha = 0$ e numericamente quando $\alpha \neq 0$. Os resultados são brevemente resumidos a seguir. A configuração PAS possui as seguintes características:

- 1) O sistema sofre uma transição de fase contínua em $T = 2$, abaixo da qual o sistema se condensa num estado que possui superposição macroscópica com apenas uma das memórias (estado de Mattis). Acima desta temperatura o sistema se encontra na fase paramagnética.
- 2) Os estados que se superpõem com mais do que uma memória (estados simétricos ou espúrios) são sempre instáveis.
- 3) As bacias de atração são grandes quando as memórias são descorrelacionadas e a taxa com que são reduzidas à medida em que α aumenta é pequena. Quando introduzimos correlação entre as memórias há uma mudança na estrutura do espaço de fase, pois as memórias encontram-se muito mais próximas umas das outras o que faz com que o tamanho das bacias dependa do número de memórias armazenadas, como foi explicado no capítulo anterior. Quando as memórias são correlacionadas, o ato de aumentar a correlação provoca uma previsível diminuição das bacias, ou seja, como a semelhança entre elas aumenta, é preciso dar mais detalhes a fim de se reconhecer um determinado padrão.
- 4) O tempo médio que o sistema leva para, a partir de um determinado estado inicial, chegar até a memória, ou seja, para que ocorra o reconhecimento, é muito pequeno. Quando o

ponto inicial já se encontra dentro da região compreendida pela bacia, são necessários apenas uns poucos passos para atingir uma configuração estável, independentemente do tamanho da rede. A dispersão no tempo médio se anula, o que mostra que diferentes caminhos levam ao mesmo ponto com o mesmo intervalo de tempo, não ficando o sistema retido por nenhum estado meta-estável ou tendo que fazer desvios por causa deles. Esta isotropia e suavidade do espaço de fases ao redor das memórias é devida à ausência de estados espúrios.

O comportamento do modelo RS, na configuração PAS, como memória associativa é altamente eficiente: possui alta capacidade de memória e o tempo necessário para recuperá-las é muito pequeno, tornando-se assim um excelente candidato para implementações práticas.

Já a configuração OPS, cujo relevo de energia é peculiar, apresenta os seguintes resultados:

1) O sistema apresenta duas fases distintas, uma que é chamada de fase simétrica e outra que possui correlação com uma das memórias armazenadas. A transição entre estas fases é de primeira ordem, exceto quando $P = 2$, onde temos uma transição contínua. Na fase simétrica o sistema se superpõe igualmente com todas as memórias, estando, portanto, numa grande bacia cujo centro é o estado dado pela regra da maioria (ver capítulo 3). Esta fase é a única que ocorre quando a temperatura possui um valor acima de $T(P)$ ou quando o número de memórias armazenadas é muito grande. Abaixo desta temperatura, o sistema entra na bacia de atração de alguma das memórias. A temperatura de transição $T(P)$ vai a zero quando P cresce muito, o que significa que as bacias de atração individuais deixam de existir e o sistema para de funcionar como memória associativa.

2) A forma do espaço de fases é também aparente nas simulações numéricas onde a dependência em P no tamanho das bacias é explícita. O desaparecimento das bacias de atração individuais quando $P \rightarrow \infty$ também é claro nas simulações.

Portanto o caso OPS apresenta propriedades de associatividade no limite em que a rede não está muito carregada, situação na qual ainda temos bacias de atração individuais e o

sistema ainda é capaz de reconhecer padrões.

5.2 Futuras extensões

Entre as muitas possíveis extensões do trabalho, existem aquelas que podem ser tomadas como o passo imediato a seguir, como por exemplo:

a) Os cálculos apresentados no capítulo 3 são válidos no limite em que $P/N \rightarrow 0$, à medida que $N \rightarrow \infty$, ou seja, $\alpha \rightarrow 0$. É importante estudar a situação em que $\alpha \neq 0$ por meio das técnicas introduzidas por Amit *et al* [12,13]. A existência ou não de uma fase de vidro de spin deve ser aparente neste tipo de cálculo e o valor de α_c , se existir, também surge neste limite.

b) A energia do modelo RS pode, no caso PAS, ser escrita da seguinte forma:

$$E = N \sum_{l=1}^k (-1)^l \sum_{\mu_1 < \dots < \mu_l} m_{\mu_1}^2 m_{\mu_2}^2 \dots m_{\mu_l}^2 \quad (5.1)$$

Quando $k = P$ temos a energia completa do modelo RS ao passo que, se $k < P$, temos o que denominamos de Modelo RS Truncado (TRS). Quando $k = 1$ temos, a menos de uma constante aditiva, o modelo de Hopfield. Simulações numéricas e alguns resultados analíticos preliminares [84] mostram que mesmo quando k é pequeno, por exemplo, 2 ou 3, a performance do sistema é sensivelmente melhorada em relação ao caso $k = 1$.

c) Podemos também definir um Modelo RS Generalizado (GRS), o qual é dado por

$$E = N \prod_{\mu=1}^P (1 - m_{\mu}^K) \quad (5.2)$$

Quando $K = 1$ temos o caso OPS e quando $K = 2$ o PAS é obtido. Ao expandirmos (5.2), o primeiro termo envolve interações de ordem K entre os neurônios e é exatamente o modelo estudado por vários autores [48]-[55]. Este sistema aparentemente apresenta três regimes qualitativamente distintos [85]: quando $K = 1, 2$ e > 2 . Quando $K > 2$ a ordem da transição entre as fases para e ferromagnética é de primeira ordem, ao contrário da transição contínua que ocorre quando $K = 2$. No limite em que $K \rightarrow \infty$, o espaço de fases

adquire uma forma peculiar, análoga ao modelo citado acima [48]-[55]: todo o espaço torna-se plano exceto nas configurações que representam as memórias, onde teremos pequenos furos, à semelhança de um campo de golfe plano. Neste limite provavelmente a solução exata do modelo pode ser obtida de forma similar ao modelo de energias randômicas [56]-[61].

Apêndice A

O Método do Ponto de Sela

Vamos neste apêndice expor brevemente os fundamentos do método do ponto de sela ('saddle-point method'), que é eficaz quando a função examinada pode ser escrita como uma integral do tipo [86]-[87]

$$J(z) = \int_C e^{zf(t)} dt \quad , \quad (\text{A.1})$$

onde C é o caminho de integração e t e z são variáveis complexas. Frequentemente é possível transformar uma integral, que pode não estar nesta forma, numa do tipo (A.1). Para exemplificar podemos citar a função partição como é calculada nos capítulos 2 e 3.

Vamos examinar o comportamento de $J(z)$ quando $|z| \rightarrow \infty$. No nosso caso, onde z é o tamanho da rede N , este é o limite termodinâmico, no qual estamos interessados. Grandes valores de $|z|$ usualmente resultam em flutuações muito rápidas no valor do integrando: se z ou $f(t)$ são complexas ao longo de porções do caminho C , quando $|z|$ cresce a parte imaginária de $[zf(t)]$ geralmente aumenta, fazendo com que o fator $e^{i\text{Im}[zf(t)]}$ oscile rapidamente, com uma frequência proporcional a $|z|$. A existência de tais oscilações inviabiliza a avaliação da integral, que requer uma estimativa acurada em cada um dos pontos do percurso, sem alguma modificação no caminho de integração devido aos constantes cancelamentos do integrando. Portanto é desejável deformar o contorno para minimizar tais efeitos. Vamos assumir que tais deformações são possíveis e se houver alguma singularidade intervindo, que seja possível avaliar seu efeito.

ESCOLHENDO O CONTORNO:

Em geral, o contorno tomado pela integral corre por regiões onde a parte real de $zf(t)$ é positiva e por outras onde é negativa. As primeiras são mais importantes uma vez que o integrando é maior e, onde $\text{Re}[zf(t)]$ é máximo, é mais importante reduzir as oscilações. O que devemos fazer é procurar um contorno ao longo do qual a parte imaginária de $zf(t)$ é constante na região (ou regiões) onde a parte real é máxima. Então, onde temos a maior contribuição para o valor final, a integral pode ser escrita como

$$J(z) = \int_C e^{zf(t)} dt = e^{i\text{Im}[zf(t)]} \int_C e^{\text{Re}[zf(t)]} dt \quad . \quad (\text{A.2})$$

O contorno, onde $\text{Re}[zf(t)]$ é mínimo, pode ser escolhido de tal modo que $\text{Im}[zf(t)]$ varie, se isto for necessário para completar o contorno. Deste modo garantimos que as oscilações do integrando afetam o mínimo possível.

O ponto t_0 onde temos o máximo de $\text{Re}[zf(t)]$ é também o máximo de $f(t)$, no qual a condição

$$f'(t_0) = \left. \frac{df}{dt} \right|_{t=t_0} = 0$$

é satisfeita. Isto é consequência das equações de Cauchy- Riemann (C-R) para uma função analítica $W(z)$, onde $z = x + iy$:

$$\frac{\partial(\text{Re}W)}{\partial x} = \frac{\partial(\text{Im}W)}{\partial y} \quad \text{e} \quad \frac{\partial(\text{Im}W)}{\partial x} = -\frac{\partial(\text{Re}W)}{\partial y} \quad .$$

Perto deste ponto é particularmente importante que as oscilações sejam suprimidas, portanto requeremos que próximo de t_0 o contorno seja escolhido de tal modo que

$$\text{Im}[zf(t)] = \text{Im}[zf(t_0)] \quad . \quad (\text{A.3})$$

isto é, a parte imaginária é constante.

O comportamento do integrando em (A.2) é visto na figura A.1, onde plotamos a superfície $\text{Re}f(t)$. A superfície deve ser plana na vizinhança de t_0 . Entretanto, ela não pode ser nem máxima nem mínima ali, pois a parte real de uma função de variável complexa não possui máximos nem mínimos. O ponto t_0 deve ser então um 'minimax', ou ponto

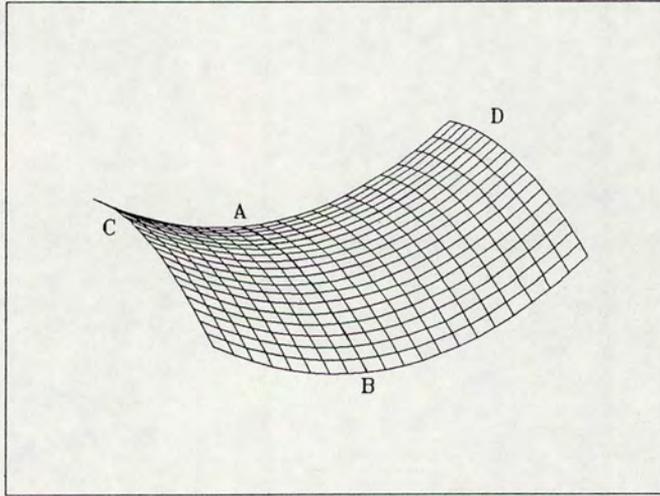


Figura A.1: Superfície $\text{Ref}(t)$. Para a curva AB temos um ponto de máximo, enquanto que um mínimo para CD .

de sela. Isto se conclui das equações de C-R acima, as quais fornecem $\nabla^2(\text{Re}W) = 0$ e $\nabla^2(\text{Im}W) = 0$.

Para entendermos melhor o que significa o caminho onde (A.3) é satisfeita, expandimos $f(t)$ ao redor do ponto t_0 :

$$f(t) = f(t_0) + \frac{1}{2}f''(t_0)(t - t_0)^2 \quad . \quad (\text{A.4})$$

Seja

$$f''(t_0) = \rho e^{i\theta}$$

e

$$t - t_0 = \lambda e^{i\phi}$$

Usando estas equações e lembrando que $f(t) = \text{Ref}(t) + i\text{Imf}(t)$, obtemos:

$$\text{Ref}(t) = \text{Ref}(t_0) + \frac{1}{2}\rho\lambda^2 \cos(\theta + 2\phi)$$

e

$$\text{Imf}(t) = \text{Imf}(t_0) + \frac{1}{2}\rho\lambda^2 \text{sen}(\theta + 2\phi)$$

Vemos então que na superfície $\operatorname{Re} f(t)$ os caminhos com descida mais acentuada ('steepest descent') a partir do ponto de sela em direção aos vales saem nas direções onde $\cos(\theta + 2\phi) = -1$. Nestas direções $\operatorname{Im} f(t)$ é constante, pois $\sin(\theta + 2\phi) = 0$. Portanto, ao longo deste caminho, $\operatorname{Re} f(t)$ varia tão rapidamente quanto possível. Note-se que à medida que $|z|$ aumenta, o máximo torna-se cada vez mais acentuado e a descida para os pontos finais cada vez mais rápida, ou seja a região de $\operatorname{Re} f(t)$ que encerra o máximo diminui cada vez mais, o que significa que quando $|z| \rightarrow \infty$, a integral é totalmente dominada pelo ponto de máximo e o método do ponto de sela torna-se exato.

Apêndice B

Cálculo dos Autovalores

Primeiramente vamos nos reportar a um resultado bem conhecido da teoria dos determinantes. Se temos uma matriz cujos elementos diagonais valem A e os que estão fora da diagonal são B , o seu determinante é [91]:

$$\begin{vmatrix} A & B & \dots & B \\ B & A & \dots & B \\ \vdots & \vdots & \ddots & \vdots \\ B & B & \dots & A \end{vmatrix} = (A - B)^{n-1} [A + (n - 1)B]$$

onde n é a dimensão da matriz.

Agora queremos calcular os autovalores da seguinte matriz M :

$$M = \begin{pmatrix} A & B & \dots & B & E & E & \dots & E \\ B & A & \dots & B & E & E & \dots & E \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ B & B & \dots & A & E & E & \dots & E \\ E & E & \dots & E & C & D & \dots & D \\ E & E & \dots & E & D & C & \dots & D \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ E & E & \dots & E & D & D & \dots & C \end{pmatrix}$$

onde a dimensão da matriz é P e temos quatro agrupamentos principais, um deles, de ordem $n \times n$ contém somente os elementos A e B , outro, de ordem $(P - n) \times (P - n)$,

contém os elementos C e D e os outros dois, de ordens $n \times (P - n)$ e $(P - n) \times n$, com o elemento E . Queremos os autovalores λ , os quais satisfazem

$$\begin{vmatrix} A - \lambda & B & \dots & B & E & E & \dots & E \\ B & A - \lambda & \dots & B & E & E & \dots & E \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ B & B & \dots & A - \lambda & E & E & \dots & E \\ E & E & \dots & E & C - \lambda & D & \dots & D \\ E & E & \dots & E & D & C - \lambda & \dots & D \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ E & E & \dots & E & D & D & \dots & C - \lambda \end{vmatrix} = 0$$

Usando a seguinte propriedade dos determinantes [91], $|AB| = |A||B|$, vamos procurar uma matriz (com determinante conhecido) que, ao ser multiplicada pela matriz acima, forneça outra com determinante também conhecido. Uma possível escolha é a seguinte:

$$N = \begin{pmatrix} C - \lambda & \frac{D(P-n-1)}{n-1} & \dots & \frac{D(P-n-1)}{n-1} & 0 & 0 & \dots & 0 \\ \frac{D(P-n-1)}{n-1} & C - \lambda & \dots & \frac{D(P-n-1)}{n-1} & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{D(P-n-1)}{n-1} & \frac{D(P-n-1)}{n-1} & \dots & C - \lambda & 0 & 0 & \dots & 0 \\ -E & -E & \dots & -E & 1 & 0 & \dots & 0 \\ -E & -E & \dots & -E & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ -E & -E & \dots & -E & 0 & 0 & \dots & 1 \end{pmatrix}$$

cujo determinante vale

$$\det N = \left[C - \lambda - \frac{D(P - n - 1)}{n - 1} \right]^{n-1} [C - \lambda + (P - n - 1)D]$$

uma vez que, como temos um bloco cujos termos são todos nulos, o determinante passa a ser o produto dos determinantes dos blocos diagonais. O produto das duas matrizes acima fornece a seguinte matriz:

$$M \times N = \begin{pmatrix} T_D & T_{FD} & \dots & T_{FD} & E & E & \dots & E \\ T_{FD} & T_D & \dots & T_{FD} & E & E & \dots & E \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ T_{FD} & T_{FD} & \dots & T_D & E & E & \dots & E \\ 0 & 0 & \dots & 0 & C - \lambda & D & \dots & D \\ 0 & 0 & \dots & 0 & D & C - \lambda & \dots & D \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & D & D & \dots & C - \lambda \end{pmatrix}$$

onde

$$T_D = (A - \lambda)(C - \lambda) + (P - n - 1)BD - E^2(P - n)$$

e

$$T_{FD} = (A - \lambda) \frac{D(P - n - 1)}{n - 1} + (C - \lambda)B + \frac{(P - n - 1)}{n - 1} BD(n - 2) - E^2(P - n) \quad ,$$

cujo determinante vale

$$\det(M \times N) = (C - \lambda - D)^{P-n-1} [C - \lambda + (P - n - 1)D] (T_D - T_{FD})^{n-1} [T_D + (n - 1)T_{FD}]$$

e, portanto, o determinante que nos interessa, já igualado a zero vale:

$$(C - \lambda - D)^{P-n-1}(A - \lambda - B)^{n-1}\{[C - \lambda + (P - n - 1)D][A - \lambda + B(n - 1)] - nE^2(P - n)\} = 0 ,$$

cujas soluções são:

$$\lambda_1 = C - D \quad , \quad P - n - 1 \text{ vezes} \quad (B.1)$$

$$\lambda_2 = A - B \quad , \quad n - 1 \text{ vezes} \quad (B.2)$$

$$\lambda_3 = \frac{1}{2}[A + C + B(n - 1) + D(P - n - 1)] + \frac{1}{2}\sqrt{[A - C + B(n - 1) - D(P - n - 1)]^2 + 4E^2n(P - n)} \quad , \quad 1 \text{ vez} \quad (B.3)$$

$$\lambda_4 = \frac{1}{2}[A + C + B(n - 1) + D(P - n - 1)] - \frac{1}{2}\sqrt{[A - C + B(n - 1) - D(P - n - 1)]^2 + 4E^2n(P - n)} \quad , \quad 1 \text{ vez} \quad (B.4)$$

Note-se que esta demonstração não vale quando $n = 1$. Neste caso, a matriz multiplicativa é a seguinte:

$$N = \begin{pmatrix} A - \lambda & -E & -E & \dots & -E \\ 0 & A - \lambda & 0 & \dots & 0 \\ 0 & 0 & A - \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & A - \lambda \end{pmatrix}$$

Os autovalores produzidos são os mesmos listados acima, exceto λ_2 , que não existe.

Apêndice C

Cálculo das Médias

Vamos mostrar como são calculadas as médias configuracionais $\langle\langle \ \rangle\rangle$ para padrões não correlacionados:

$$\langle\langle \ \rangle\rangle = \frac{1}{2^P} \sum_{\xi^1} \cdots \sum_{\xi^n} \sum_{\xi^{n+1}} \cdots \sum_{\xi^P}$$

Definimos a seguinte variável:

$$z_n = \sum_{\mu=1}^n \xi^\mu$$

e a média $\langle\langle \ \rangle\rangle$ é escrita como uma média sobre tal variável.

Seja $f(z_n)$ uma função qualquer de z_n . Temos:

$$\begin{aligned} \langle\langle f(z_n) \rangle\rangle &= \frac{1}{2^P} \sum_{\xi^1} \cdots \sum_{\xi^n} f(z_n) \underbrace{\sum_{\xi^{n+1}} \cdots \sum_{\xi^P}}_{2^{P-n}} \\ &= \frac{1}{2^n} \sum_{\xi^1} \cdots \sum_{\xi^n} f(z_n) \\ &= \sum_{k=-n}^{+n} p(k) f(k) \end{aligned}$$

onde $p(k)$ é a probabilidade com que z_n assume o valor k . Como temos C_n^ℓ configurações distintas que dão o mesmo valor de $z_n = n - 2\ell$ (ℓ é o número de spins $+1$), $p(k)$ vale:

$$p(k = n - 2\ell) = \frac{C_n^\ell}{2^n}$$

Então:

$$\langle\langle f(z_n) \rangle\rangle = \frac{1}{2^n} \sum_{\ell=0}^n \frac{n!}{\ell!(n-\ell)!} f(n-2\ell) \ .$$

Referências Bibliográficas

- [1] FERREIRA, A. B. H., *Novo Dicionário da Língua Portuguesa*, Editora Nova Fronteira.
- [2] HOPFIELD, J. J., *Proc. Natl. Acad. Sci. USA* **79**, p. 2254-2558, 1982.
- [3] AMIT, D., *Modelling Brain Function*. Cambridge University Press, 1989.
- [4] BRADLEY, A. e MYERS, C., *Physics World*, **2**, p. 17-18, 1989.
- [5] CAMPBELL, P., *Physics World* **2**, p. 8-10, 1989.
- [6] HUMPERT, B., *Computer Physics Communications*, **56**, p. 299-311, 1990.
- [7] ALTHERR, T. e SEIXAS, J. C., *preprint*, CERN-TH.5758/89 , 1990.
- [8] ANGEL, J. R. P., WIZINOWICH, P., LLOYD-HART, M. e SANDLER, D. G., *Nature*, **348**, p. 221-224, 1990.
- [9] SANDLER, D. G., BARRET, T. K., PALMER, D. A., FUGATE, R. Q. e WILD, W. J., *Nature*, **351**, p. 300-351, 1991.
- [10] LEE, H. e KANG, I. S., *Journal of Computational Physics*, **91**, p. 110-131, 1990.
- [11] AMIT, D., GUTFREUND, H. e SOMPOLINSKY, H., *Phys. Rev.*, **A32**, p. 1007, 1985.
- [12] AMIT, D., GUTFREUND, H. e SOMPOLINSKY, H. , *Phys. Rev. Lett.*, **55**, p. 1530, 1985.

- [13] AMIT, D., GUTFREUND, H. e SOMPOLINSKY, H. , *Ann. Phys.*, **173**, p. 30, 1987.
- [14] DE ALMEIDA, R. M. C. e IGLESIAS, J. R., *Phys. Lett.*, **A146**, p. 239, 1990.
- [15] ARENZON, J. J., DE ALMEIDA, R. M. C. e IGLESIAS, J. R., submetido para publicação, 1991.
- [16] ARENZON, J. J., DE ALMEIDA, R. M. C. e IGLESIAS, J. R.. In: *NONLINEAR PHENOMENA IN FLUIDS, SOLIDS AND OTHER COMPLEX SYSTEMS*, p. 121-129, Elsevier Science Publishers, 1991.
- [17] ARENZON, J. J., DE ALMEIDA, R. M. C., IGLESIAS, J. R., PENNA, T. J. P. e DE OLIVEIRA, P. M. C., *J. Physique I*, a ser publicado, 1991.
- [18] PENNA, T. J. P., DE OLIVEIRA, P. M. C., ARENZON, J. J., DE ALMEIDA, R. M. C. e IGLESIAS, J. R., *Int. Journ. Mod. Phys. C*, a ser publicado, 1991.
- [19] COTTERILL, R. M. J., *Physica Scripta*, **T13**, p. 161-168, 1986.
- [20] KANDEL, E. R., *Cellular Basis of Behavior*. Freeman, San Francisco, 1976.
- [21] CAREW, T. J., CASTELLUCCI, V. F. e KANDEL, E. R., *Int. J. Neurosci.*, **2**, p. 79-98, 1971.
- [22] KANDEL, E. R. e TAUC, L., *J. Physiol.*, **181**, p. 1-27, 1965.
- [23] HEBB, D. O., *The Organization of Behavior*, Wiley, New York, 1949.
- [24] McCULLOCH, W. S. e PITTS, W. A., *Bull. Math. Biophys.*, **5**, p. 115, 1943.
- [25] DOMANY, E., *J. Stat. Phys.*, **51**, No. 5/6, p. 743-775, 1988.
- [26] SOMPOLINSKY, H., *Physics Today*, p. 70-80, dezembro 1988.
- [27] PARGA, N., *Comp. Phys. Comm.*, **55**, p. 77-84, 1989.
- [28] GUTFREUND, H., *Physica* , **A163**, p. 373-385, 1990.

- [29] DE ALMEIDA, R. M. C.. In: III ENCONTRO REGIONAL DE ATUALIZAÇÃO EM FÍSICA, Santa Maria, RS. Proceedings, a ser publicado na *Natura*.
- [30] VIANA, L. *Memoria Natural y Artificial*, SEP, Mexico, 1990.
- [31] TOULOUSE, G., *J. Phys.*, **A22**, p. 1959-1968, 1989.
- [32] LAUTRUP, B., *preprint*, The Niels Bohr Institute, NBI-HE-88-06, 1988.
- [33] CLARK, J. W., *Phys. Reports*, **158** (2), p. 91-157, 1988.
- [34] PERETTO, P. e NIEZ, J. J., *Biol. Cybern.*, **54**, p. 53-63, 1986.
- [35] MEZARD, M., PARISI, G. e VIRASORO, M. A., *Spin Glass Theory and Beyond*, World Scientific Publishing Co., 1987.
- [36] BINDER, K. e YOUNG, A. P., *Rev. Mod. Phys.*, **58**, p. 801-976, 1986.
- [37] CHOWDHURY, D., *Spin Glasses and other Frustrated Systems*, World Scientific Publishing Co., 1986.
- [38] MATTIS, D. C., *Phys. Lett.*, **A56** (5), p. 421-422, 1976.
- [39] WEISBUCH, G. e FOGELMAN-SOULIÉ, F., *J. Physique Lett.*, **46**, p. L623-L630, 1985.
- [40] CRISANTI, A., AMIT, D. e GUTFREUND, H., *Europhys. Lett.*, **2**, p. 337-341, 1986.
- [41] AMIT, D., GUTFREUND, H. e SOMPOLINSKY, H., *Phys. Rev.* **A35**, p. 2293, 1987.
- [42] KOHRING, G. A., *J. Stat. Phys.*, **59**, 1990.
- [43] FORREST, B. M., *J. Phys.*, **A21**, p. 245-255, 1988.
- [44] KANTER, I., *Phys. Rev.*, **A40**(5), p. 2611-2614, 1989.
- [45] KOHRING, G. A., *J. Phys.*, **A23**, p. 2237-2241, 1990.

- [46] GARDNER, E., *Nuclear Physics*, **B257** [FS14], p. 747-765, 1985.
- [47] STARIOLO, D., *Physica*, **A166**, p. 622-632, 1990.
- [48] GARDNER, E., *J. Phys.*, **A20**, p. 3453-3464, 1987.
- [49] ABBOTT, L. F. e ARIAN, Y., *Phys. Rev.*, **A36**, p. 5091-5094, 1987.
- [50] BALDI, P. e VENKATESH, S. S., *Phys. Rev. Lett.*, **58** (9), p. 913-916, 1987.
- [51] PERSONNAZ, L., GUYON, I. e DREYFUS, G., *Europhys. Lett.*, **4**, p. 863-867, 1987.
- [52] HORN, D. e USHER, M., *J. Phys. France*, **51**, p. 389-395, 1988.
- [53] PEREZ, P. e MATUS, I. J., *preprint*, 1990.
- [54] KOHRING, G. A., *J. Phys. France*, **51**, p. 145-155, 1990.
- [55] TAMARIT, F. A., STARIOLO, D. A. e CURADO, E. M. F., *Phys. Rev.*, **A43** (2), 1991.
- [56] DERRIDA, B., *Phys. Rev. Lett.*, **45**(2), p. 79-82, 1980.
- [57] DERRIDA, B., *Phys. Rev.*, **B24**(5), p. 2613-2626, 1981.
- [58] DERRIDA, B., *Phys. Rep.*, **67**(1), p. 29-35, 1980.
- [59] GROSS, D. J. e MEZARD, M., *Nuclear Physics* **B240** [FS12], p. 431-452, 1984.
- [60] DERRIDA, B. e GARDNER, E., *J. Phys.*, **C19**, p. 2253-2274, 1986.
- [61] MOTTISHAW, P., *Europhys. Lett.*, **1**(8), p. 409-414, 1986.
- [62] PARGA, N. e VIRASORO, M. A., *J. Physique*, **47**, p. 1857, 1986.
- [63] FEIGELMAN, M. V. e IOFFE, L. B., *Int. J. Mod. Phys. B*, **1**, p. 51, 1987.
- [64] GUTFREUND, H., *Phys. Rev.*, **A37**, p. 570, 1988.

- [65] BACCI, S., MATO, G. e PARGA, N., *J. Phys.*, **A23**, p. 1801-1810, 1990.
- [66] ENGEL, A., *J. Phys.*, **A23**, p. L285-L289, 1990.
- [67] IDIART, M. A. P. e THEUMANN, A., *J. Phys.*, **A24**, p. L649-L658, 1991.
- [68] PARISI, G., *J. Phys.*, **A19**, p. L617-L620, 1986.
- [69] VIANA, L. *J. Physique*, **49**, p.167-174, 1988.
- [70] NADAL, J. P., TOULOUSE, G., CHANGEUX, J. P. e DEHAENE, S., *Europhys. Lett.*, **1** (10), p. 535-542, 1986.
- [71] NADAL, J. P., TOULOUSE, G., CHANGEUX, J. P. e DEHAENE, S., *Europhys. Lett.*, **2** (4), p. 346, 1986.
- [72] MEZARD, M., NADAL, J. P. e TOULOUSE, G., *J. Physique*, **47**, p. 1457-1462, 1986.
- [73] KINZEL, W., *Z. Phys. B*, **60**, p. 205-213, 1985.
- [74] PARISI, G., *J. Phys.*, **A19**, p. L675-L680, 1986.
- [75] DERRIDA, B., GARDNER, E. e ZIPPELIUS, A., *Europhys. Lett.*, **4**, p. 167, 1987.
- [76] BARKAI, E., KANTER, I. e SOMPOLINSKY, H., *Phys. Rev.*, **A41** (2), p. 590-597, 1990.
- [77] NOEST, A. J., *Phys. Rev. Lett.*, **63** (13), p. 1739-1742, 1989.
- [78] FORREST, B. M., *J. Phys. France*, **50**, p. 2003-2017, 1989.
- [79] KÜRTEEN, K. E., *J. Phys. France*, **51**, p. 1585-1594, 1990.
- [80] GUYON, I., PERSONNAZ, L., NADAL, J. P. e DREYFUS, G., *Phys. Rev.*, **A38** (12), p. 6365-6372, 1988.
- [81] VIRASORO, M. A., *Phys. Reports*, **184**, p. 301-306, 1989.

- [82] CARNAHAN, B., LUTHER, H. A. e WILKES, J. O., *Applied Numerical Methods*, John Wiley & Sons, Inc., 1969.
- [83] GARDNER, E., *J. Phys.*, **A21**, p. 257, 1988.
- [84] ARENZON, J. J., DE ALMEIDA, R. M. C., IGLESIAS, J. R., PENNA, T. J. P. e DE OLIVEIRA, P. M. C., em preparação.
- [85] ARENZON, J. J., DE ALMEIDA, R. M. C. e IGLESIAS, J. R., em preparação.
- [86] MORSE, P. M. e FESHBACH, H., *Methods of Theoretical Physics*, McGRAW-HILL BOOK COMPANY, INC., 1953.
- [87] MATHEWS, J. e WALKER, R. L., *Mathematical Methods of Physics*, W. A. BENJAMIN, INC., 1965.
- [88] AYRES, F., *Matrizes*, Editora McGRAW HILL DO BRASIL LTDA, 1975.