

**Caracterização de Perfis de Reatividade  
Imunológica**

Leonardo Costa Ribeiro

Março de 2004

# Caracterização de Perfis de Reatividade Imunológica

Leonardo Costa Ribeiro

Orientador: Prof. Américo Tristão Bernardes

Co-orientador: Prof. Ronald Dickman

Dissertação apresentada à UNIVERSIDADE FEDERAL DE MINAS GERAIS, como requisito parcial para a obtenção do grau de MESTRE EM CIÊNCIAS (Física).

Março de 2004

# Agradecimentos

Ao Prof. Américo, pelos 4 anos de amizade e constante bom humor, e pela orientação instigante que me motiva a cada dia.

Ao Prof. Ronald Dickman, pela co-orientação e discussões de grande valor.

Ao Prof. José Guilherme, pela revisão deste texto, por deletar minha área na Tabajara e pela companhia durante o copos de cerveja.

Ao grupo de Imunologia do Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais, pelo material que permitiu desenvolvimento desse trabalho, em particular ao Luiz pela sua iniciativa.

À minha mãe, Beatriz, pelo incentivo à minha carreira e pelo apoio nos momentos difíceis e amizade.

À Aurea, pelo seu amor, carinho, companherismo, amizade, dedicação, apoio, força...

Aos companheiros de física, incluindo, Wanderson, Breno, Indhira, Júlia, Tiago, Schneider, Ana Júlia, Pablo, Catão, Planeta, Gordinho, Bira, Cabelinho, Madá, Rodrigo, etc, pelos inenarráveis momentos de diversão e “lama”.

Ao Gordinho, por me apresentar o carnaval em Diamantina.

Aos amigos de infância Willian, Luiz Phillipe, Mauricio, Márcia Aline, por terem crescido junto comigo, e ao demais grandes amigos de Elói Mendes.

Ao Franscisco Carlos, pela seu visão de mundo.

Ao CNPQ pela apoio finaceira a esse projeto.

*“Entre o céu e a terra, existem mais coisas que sonha a nossa vã filosofia.”*

*William Shakespeare*

*À minha mãe, à Minha Linda, que  
reluz como ouro iluminando os  
meus dias e ao grande amigo  
Francisco Carlos.*

# Conteúdo

<b>RESUMO</b>	<b>ix</b>
<b>ABSTRACT</b>	<b>x</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Sistema Imunológico . . . . .	1
1.2 Sistemas Complexos . . . . .	8
1.3 Sistema Imune como um Sistema Complexo . . . . .	15
<b>2 Revisão Bibliográfica</b>	<b>19</b>
2.1 Fractais . . . . .	19
2.2 Análise de Componentes Principais (PCA) . . . . .	28
2.3 Agrupamento Super-paramagnético . . . . .	30
<b>3 Materiais e Métodos</b>	<b>39</b>
3.1 Experimento . . . . .	39
3.2 Tratamentos Digitais . . . . .	41
3.2.1 Digitalização . . . . .	41
3.2.2 Suavização . . . . .	43
3.2.3 Normalização . . . . .	46

<b>4</b>	<b>Resultados e Discussões</b>	<b>50</b>
4.1	Expoente de Hurst . . . . .	51
4.2	Análise de Variância . . . . .	53
4.3	Análise por PCA . . . . .	56
4.4	Análise por Agrupamento Super-paramagnético . . . . .	58
<b>5</b>	<b>Conclusões</b>	<b>61</b>
	<b>BIBLIOGRAFIA</b>	<b>63</b>

# Lista de Figuras

1.1	A figura mostra esquematicamente a estrutura de uma imunoglobulina, à esquerda. E, à direita, a reação de diferentes imunoglobulinas com um mesmo antígeno. . . . .	2
1.2	A figura mostra esquematicamente a estrutura de uma molécula de TCR. . . . .	3
1.3	Esquematização dos três mecanismos que atuam sobre os linfócitos de transição quando esses reagem com componentes que constituem o organismo. . . . .	5
1.4	Ilustração da ligação entre imunoglobulinas proposta por Jerne. . . . .	6
1.5	O antígeno é parcialmente digerido, ligado ao MHC e então expresso pela célula apresentadora. Quando o linfócito T reage com essa molécula ele se torna ativo, eliminando linfocinas, na presença das quais os linfócitos B que reagem diretamente com os antígenos se tornam ativos. . . . .	7
1.6	Solução exata para o modelo de Ising bi-dimensional numa rede quadrada infinita. Na parte (a) está o gráfico da magnetização e na parte (b) a sua derivada. Figura baseada em [7] . . . . .	10
1.7	Solução exata para o modelo de Ising bi-dimensional sem interação numa rede quadrada infinita na presença de um campo externo. Na parte (a) está o gráfico da magnetização e na parte (b) sua derivada. . . . .	11
1.8	Ilustração de como seria um gráfico da energia em função dos estados para um sistema sem estados meta-estáveis (a) e para um sistema com estados meta-estáveis (b). . . . .	11
1.9	Seqüência de estados do autômato celular, onde um ponto preto indica um sítio cheio e um ponto branco um sítio vazio. . . . .	13



1.10	Ilustração pictórica de uma pilha de areia. . . . .	14
1.11	Lei de potência encontrada na frequência com que ocorrem terremotos que liberam energia $E$ , à esquerda. À direita, a região em que esses terremotos ocorreram. . . . .	15
1.12	Lei de potência encontrada na frequência com que as palavras aparecem num texto de língua inglesa. . . . .	16
1.13	Perfil de reatividade obtido em um experimento. Os tons de cinza da imagem indicam a intensidade da reação naquele ponto. . . . .	17
2.1	Objetos com dimensões 1, 2 e 3. . . . .	19
2.2	Linha de comprimento $L$ unitário dividida em $N$ elementos de comprimento $\epsilon$ e quadrado de área $L$ unitária também dividida em $N$ elementos de área $\epsilon$ . . . . .	21
2.3	Ilustração do procedimento para gerar uma curva fractal de Koch. . . . .	22
2.4	Na parte (a), uma curva fractal com expoente de Hurst 0.5 gerada pela função real de Weierstrass. Na parte (b), a lei de potência da rugosidade como função da largura da escala. . . . .	24
2.5	Transformada de Fourier da mesma curva da figura anterior. . . . .	26
2.6	Fotos de estruturas fractais encontradas na natureza. Em (a) um feto de galinha e em (b) uma concha. . . . .	27
2.7	Distribuição de pontos nos eixos $y_1$ e $y_2$ e também os eixos $z_1$ e $z_2$ definidos pelos autovetores da matriz de covariância. . . . .	29
2.8	Gráficos obtidos por simulação do modelo de Ising mudando os spins um a um. A parte (a) mostra o processo de termalização. Na parte (b) o processo de decorrelação entre configurações subseqüentes. . . . .	33
2.9	Gráficos obtidos por simulação do modelo de Ising mudando blocos de spins. A parte (a) mostra o processo de termalização. Na parte (b) o processo de decorrelação entre configurações sub-seqüente. . . . .	34
2.10	Gráficos obtidos por simulação do modelo de Ising mudando blocos de spins. Na parte (a) a magnetização e na (b) a suscetibilidade. . . . .	35

2.11	Distribuição de pontos utilizada para exemplificar a técnica de agrupamento super-paramagnético. . . . .	37
2.12	Gráficos da magnetização e suscetibilidade que identificam as três fases do sistema. . . . .	38
3.1	Imagem das proteínas coradas após o processo de separação e fixação na membrana de nitro-celulose. . . . .	40
3.2	Membrana resultante do experimento. Nela estão coradas as proteínas do extrato (linhas que cruzam todo o perfil) e a imunoglobulinas do tipo M que reagiram (linhas restritas às canaletas). . . . .	41
3.3	Parte do arquivo texto no qual a imagem é codificada. A parte superior é o cabeçalho e a inferior o corpo do arquivo. . . . .	42
3.4	Curva obtida da decodificação dos tons de cinza de uma canaleta. No eixo $y$ o tom de cinza médio da linha correspondente à proteínas de mesmo peso molecular e no eixo $x$ a posição desta linha na canaleta. . . . .	43
3.5	A parte (a) mostra um intervalo do perfil antes da suavização e o perfil após o processo de suavização. E na parte (b) o ruído retirado durante a suavização. . . . .	44
3.6	Gráfico da rugosidade com função da largura da escala. Observa-se que a lei de potência se mantém somente para escalas pequenas. . . . .	45
3.7	Três diferentes perfis artificiais, já suavizados, gerados com diferentes variâncias médias das gaussianas. . . . .	46
3.8	Histograma do ruído adicionado ao perfil artificial e histograma do ruído retirado pelo processo de suavização desse mesmo perfil. . . . .	47
3.9	Na parte (a), as rugosidade para os perfis com diferentes variâncias e em (b) as rugosidades para os perfis com diferentes larguras. . . . .	48
3.10	Na parte (a), o perfis antes da normalização e em b após a normalização. . . . .	49
4.1	Expoente de Hurst dos perfis obtidos utilizando extrato de fígado, separados pelo quadro clínico da malária dos indivíduos. . . . .	52

4.2	Expoente de Hurst dos perfis obtidos utilizando extrato de cérebro, separados pelo quadro clínico da malária dos indivíduos. . . . .	53
4.3	Na parte (a), a projeção dos pontos correspondentes aos perfis obtidos através do extrato de fígado nos 1 <sup>o</sup> e 3 <sup>o</sup> eixos principais. Na parte (b), a projeção dos pontos correspondentes aos perfis obtidos através do extrato de cérebro nos 1 <sup>o</sup> e 2 <sup>o</sup> eixos principais. . . . .	57
4.4	A parte (a) mostra a suscetibilidade obtida aplicando o agrupamento super-paramagnético. Na parte (b), o tamanhos do dois maiores agrupamentos identificados. . . . .	58
4.5	A parte (a) mostra a suscetibilidade obtida aplicando o agrupamento super-paramagnético. Na parte (b), os tamanhos dos três maiores agrupamentos identificados. . . . .	59
4.6	Projeção dos pontos correspondentes aos perfis obtidos através do extrato de cérebro nos 1 <sup>o</sup> e 2 <sup>o</sup> eixos principais, separados no grupos identificados pela técnica de agrupamento super-paramagnético. . . . .	59

# Resumo

Recentemente, comprovou-se a existência de interações inter-linfocitárias. Essas interações suprimem a ativação de alguns linfócitos, o que seria responsável pela inexistência de respostas imunológicas contra o próprio organismo.

A introdução desse tipo de interação nos permite ver o sistema imunológico como um sistema complexo. Assim, para se medir as propriedades dos sistema imune, como um todo, deve-se fazer experimentos que possibilitem a interação inter-linfocitária. Um desses experimentos mede a reatividade de uma população de proteínas quando apresentada às imunoglobulinas presentes no soro sanguíneo de um indivíduo.

O perfil de reatividade obtido nesse experimento é digitalizado e transformado num gráfico entre o peso molecular da proteína e a reatividade que esta apresentou. Calculamos o expoente de Hurst dessa curva, aplicamos a análise de componentes principais e o agrupamento super-paramagnético. Com essas análises, obtemos a formação de grupos distintos segundo o quadro clínico da malária que os indivíduos apresentavam.

# Abstract

Recently, the existence of interactions between lymphocytes that suppress the activation of certain other lymphocytes was demonstrated. These interactions prevent an immune response against the body's own proteins.

The existence of this kind of interaction allows us to see the immune system as a complex system. Thus, to characterize the immune system requires experiments that allow interlymphocyte interactions. One such experiment measures the reactivity of a protein population when presented to the immunoglobulins present in the serum of an individual.

The profile of reactivity obtained in this experiment is scanned and transformed into a plot between the molecular weight of the protein and the observed reactivity. We apply three statistical techniques to the resulting function: determination of the Hurst exponent, principal component analysis and super-paramagnetic clustering. With these analyses, we find that population falls into groups according to the clinical picture of malaria exposure.

# Capítulo 1

## Introdução

### 1.1 Sistema Imunológico

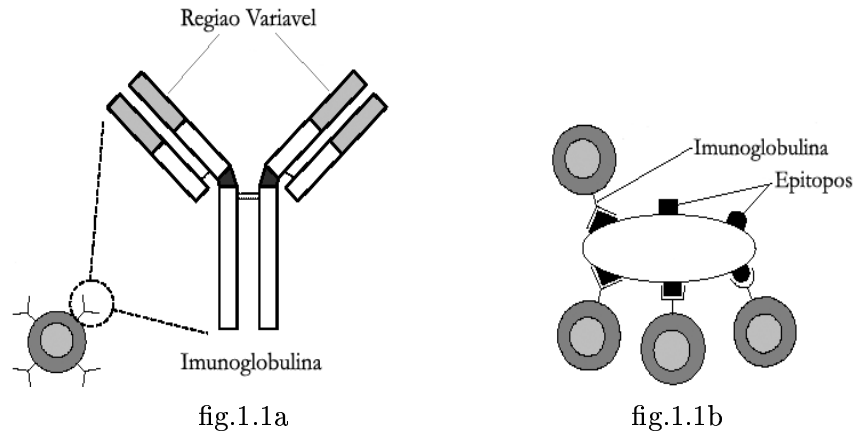
Numa visão pictórica que tradicionalmente nos é passada, o sistema imunológico é formado por todos os entes (órgãos, células, proteínas, etc) que desempenham um papel na tarefa de proteger o organismo contra o que lhe é estranho<sup>1</sup>. Assim o sistema imunológico teria a propriedade de identificar materiais estranhos (antígenos), substâncias que não constituem o organismo, reagindo com elas e acionando mecanismos que induzem, de alguma forma, a destruição dessas substâncias.

Em geral, essa identificação ocorre através da reação do corpo estranho com um tipo de proteína denominada imunoglobulina ou anticorpo. A reação é não-covalente, envolvendo, em geral, forças de Van der Waals, interações entre grupos carregados e pontes de hidrogênio. Devido a sua natureza, essas fracas interações são de intensidade suficiente para manter as macromoléculas ligadas, em geral, quando existe uma relação de complementariedade entre o formato dos ligantes, como um modelo chave-fechadura. A região do corpo estranho onde ocorre a ligação com a imunoglobulina é denominada epítipo. E a região da imunoglobulina onde ocorre a ligação é a parte variável da imunoglobulina. Diferentes imunoglobulinas podem reagir com um mesmo antígeno com variados graus de afinidade dependendo da relação entre o formato dessas moléculas (Figura 1.1). E, reciprocamente, uma imunoglobulina pode reagir com uma multiplicidade de antígenos com diferentes

---

<sup>1</sup>O texto dessa seção está baseado nas referências [1, 2, 3]

afinidades, da mesma forma, dependendo da relação entre o formato dos ligantes. A esta afinidade de ligação iremos nos referir, futuramente, como a reatividade da imunoglobulina ou do antígeno. Assim uma alta reatividade corresponde a uma alta afinidade de ligação.

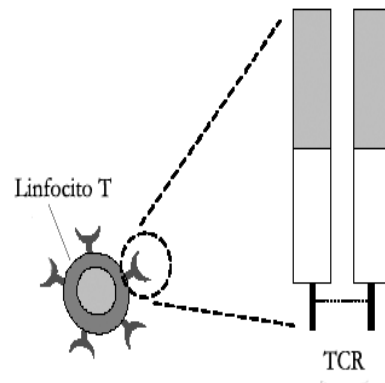


**Figura 1.1:** A figura mostra esquematicamente a estrutura de uma imunoglobulina, à esquerda, ressaltando a região variável da molécula. E, à direita, a reação de diferentes imunoglobulinas com um mesmo antígeno através da ligação da parte variável da imunoglobulina com os epitopos do corpo estranho. Figura baseada em [2]

Para executar de forma eficiente a identificação do enorme número de diferentes antígenos que podem invadir o organismo, o sistema imune produz, através de processos de recombinação gênica, uma enorme gama de imunoglobulinas que diferem pelas suas regiões variáveis (Figura 1.1a). Pela análise da quantidade de DNA nuclear, pode-se calcular que um organismo como o corpo humano utiliza cerca de  $10^5$  proteínas diferentes em sua constituição; em comparação, calcula-se que cerca de  $10^8$  diferentes moléculas de imunoglobulinas sejam produzidas por um indivíduo. Em torno de mil vezes mais que todos os outros componentes do organismo juntos. Este fato mostra a extraordinária diversidade de imunoglobulinas produzidas.

Um fato espantoso é que cada imunoglobulina é gerada por uma célula (linfócito) encarregada da produção daquele específico tipo de proteína. Cada linfócito produz somente um tipo de imunoglobulina. Assim, para sintetizar cem milhões de variedades de anticorpos, o organismo utiliza cem milhões de variedades de linfócitos.

Nestas células produtoras, as imunoglobulinas ficam expressas, ou seja, grudadas na superfície externa da membrana celular juntamente com outras proteínas. Apresentadas desta forma, dizemos que estas moléculas são receptores. Dependendo de quais são as outras proteínas expressas, os linfócitos podem desempenhar funções bem diferentes no funcionamento do sistema imune. Dividindo estas células em dois grandes grupos funcionalmente distintos: os linfócitos T que expressam a molécula de TCR (*T Cell Receptor*) e os B, que não a expressam (Figura 1.2).



**Figura 1.2:** A figura mostra esquematicamente a estrutura de uma molécula de TCR que é expressa somente nos linfócitos T. Figura baseada em [2]

Para compreendermos como a diversidade linfocitária é formada e também a questão, primordial na imunologia, de elucidar os mecanismos que levam o sistema imune a não reagir com os componentes próprios do organismo, temos que analisar passo a passo a formação e maturação dos linfócitos.

Os linfócitos B são derivados de células localizadas primariamente nos tecidos hematopoiéticos encontrados no fígado fetal e na medula óssea, por um complexo conjunto de eventos de diferenciação.

Na medula óssea encontram-se os linfócitos B imaturos que expressam baixos níveis de receptores, incluindo as imunoglobulinas. A primeira função conhecida destas imunoglobulinas está relacionada com a indução do desenvolvimento da célula que a expressa, pois ela inibe a ação da enzima de ativação da recombinação gênica, conseqüentemente induzindo a finalização dos rearranjos das cadeias das imunoglo-



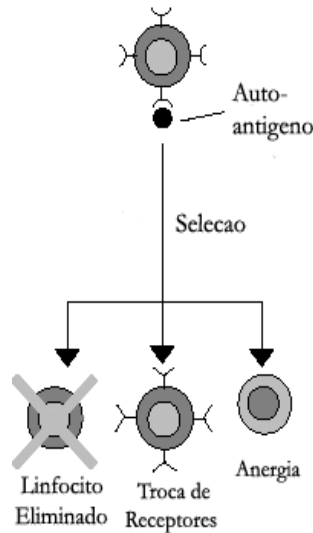
bulinas.

Depois de cerca de 3 a 5 dias, as células se encontram mais avançadas no processo de diferenciação expressando maiores níveis de receptores. Neste estágio, em que elas são denominadas células de transição, os linfócitos B, ainda em formação, iniciam uma caminhada até os órgãos linfóides secundários (baço, linfonodos e placas de Payer) durante a qual os receptores atingem altos níveis de concentração. Nestes órgãos, os linfócitos de transição se instalam nos folículos e formam verdadeiras piscinas cheias de linfócitos, onde o processo de maturação se concluirá e cerca de 97 a 99% das células encontrarão seu fim.

Nessas “piscinas de maturação”, os linfócitos naturalmente entram em contato com componentes próprios do organismo que se localizam próximos aos folículos ou que circulam pelo indivíduo. Ao entrarem em contato com esses componentes, pode haver a ligação destes com as imunoglobulinas que são expressas, agora em altas concentrações, pelos linfócitos. Quando ocorre a reação, o próprio linfócito que ligou ao componente aciona mecanismos que levarão ou a sua auto-destruição, ou a não apresentar mais resposta se tornando inutilizado, ou a recombinar suas imunoglobulinas criando um novo linfócito que não reagirá mais com os auto-componentes do organismo (Figura 1.2).

Caso esses linfócitos não fossem eliminados, ao atingirem seu estado final de maturação iriam provocar uma resposta imunológica contra aqueles componentes do organismo que reagiriam com suas imunoglobulinas expressas, levando o organismo a eliminar componentes que o constituem, ou seja, a uma auto-degradação, o que chamamos patologias auto-imunes.

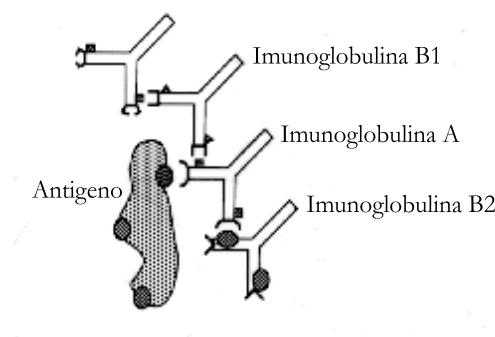
Assim, os mecanismos de seleção dos linfócitos asseguram a eliminação das células que levariam o organismo a sua auto-destruição. No entanto, isso ocorre em parte, pois somente são eliminados as células reativas a auto-componentes encontrados no seu micro-ambiente de maturação (piscinas de linfócitos). Linfócitos que sobreviveram ao processo de maturação podem ainda reagir com auto-componentes que não estavam presentes no seu micro-ambiente de maturação. Portanto, esta forma de seleção dos linfócitos não é suficiente para assegurar a tolerância do sistema imune contra antígenos próprios. É necessário a existência de algum outro mecanismo para assegurar em sua totalidade tal tolerância.



**Figura 1.3:** Esquemática dos três mecanismos que atuam sobre os linfócitos de transição quando esses reagem com componentes que constituem o organismo. Os mecanismos asseguram, ou através da destruição do linfócito, ou pela troca dos seus receptores ou por suprimir a sua ativação, que não haja uma resposta imune contra o próprio organismo. Figura baseada em [2]

Em resposta a essa questão, surgem modelos que propõem a presença de populações de linfócitos T capazes de suprimir uma resposta de auto-agressão. Estes linfócitos, ditos reguladores, exercem uma supressão específica ao antígeno considerado, deixando intacta a resposta do sistema frente a outros antígenos não relacionados. Ou seja, um linfócito regulador faz com que o sistema imune não produza nenhuma resposta contra um certo antígeno e deixa inalterada a resposta do sistema contra os demais antígenos. Para explicar o fenômeno da tolerância é necessário, então, introduzir conceitualmente a existência de interações entre linfócitos, pois para que o linfócito regulador iniba a resposta contra um certo antígeno ele deve, de alguma forma, agir sobre o linfócito que responderia contra o corpo estranho em questão. Considerando esta imuno-regulação, um alto grau de complexidade é adicionado à teoria conceitual do sistema imune, visto que as interações entre linfócitos podem se dar continuamente, de maneira recursiva, estabelecendo intrincados ciclos dinâmicos. Historicamente, e de maneira independente do fenômeno de tolerância, o primeiro modelo para o sistema imune envolvendo interações linfocitárias foi proposto por Niels Jerne em 1974 [5] (a teoria da Rede Idiotípica). O modelo considera que imunoglobulinas

podem reagir entre si através dos mesmos tipos de ligações que podem se formar entre as imunoglobulinas e os epítomos, ou seja, pode ocorrer que a região variável de duas imunoglobulinas tenham formas complementares (chave-fechadura) levando essas moléculas a reagirem formando ligações agora denominadas idiotípicas (Figura 1.4).

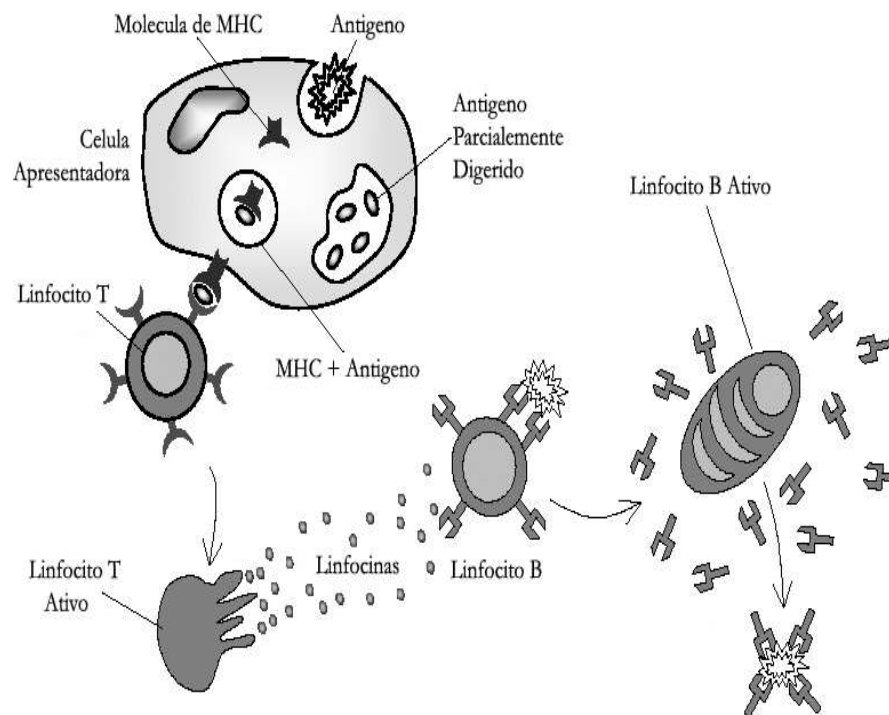


**Figura 1.4:** Ilustração da ligação entre imunoglobulinas proposta por Jerne. A imunoglobulina A pode se ligar ao antígeno e, também, às imunoglobulinas B1 e B2 que, por sua vez, podem se ligar à outras moléculas. Formando, assim, uma cadeia de reações. Figura baseada em [5]

Finalizando seu processo de maturação, os linfócitos de transição sobreviventes ao mecanismo de deleção se diferenciam rapidamente em linfócitos recirculantes, que são a principal população de linfócitos de longa duração que circulam pelo organismo. Estes emergem dos órgãos, onde são originalmente gerados, em um estado de repouso no qual o tamanho da célula é pequeno, a atividade metabólica é mínima e existem poucos receptores para moléculas que induzem o crescimento. No momento em que os linfócitos de repouso reagem com algum corpo estranho, eles se tornam ativos, aumentando de tamanho, dividindo-se por mitose, gerando clones de si mesmos, exibindo altos níveis de receptores para fatores de crescimento e, eles próprios, secretando moléculas que induzem o crescimento de outras células (linfocinas) e primordialmente secretando suas imunoglobulinas.

A ativação dessas células guarda uma interessante peculiaridade, nem sempre o contato simples com o antígeno é capaz de promover a sua ativação. Agora surge uma diferença crucial entre os linfócitos B e T. Os linfócitos B podem se ligar aos antígenos

por união direta às suas imunoglobulinas expressas, tornando-se ativos quando isso ocorre na presença de certas linfocinas. Em contraste, os linfócitos T são incapazes de ligar materiais como moléculas de antígenos em sua forma nativa diretamente a seus receptores. Para que a ligação ocorra, o material precisa ser apresentado aos linfócitos T incorporado na membrana de outras células (as apresentadoras de antígenos). Antes de incorporá-los a sua membrana, as células apresentadoras digerem parcialmente, em compartimentos intracelulares, os antígenos e os conjugam às moléculas que se ligam exclusivamente ao MHC (*Major Histocompatibility Complex*, que é um dos receptores das células T). Para que a ativação do linfócito T efetivamente ocorra é necessário que haja tanto a reação do antígeno parcialmente digerido com sua respectiva imunoglobulina quanto a reação da molécula conjugada com seu respectivo MHC (Figura 1.5).



**Figura 1.5:** O antígeno é parcialmente digerido, ligado ao MHC e então expresso pela célula apresentadora. Quando o linfócito T reage com essa molécula ele se torna ativo, eliminando linfocinas, na presença das quais os linfócitos B que reagem diretamente com os antígenos se tornam ativos. Figura baseada em [2]

Experimentos mostram que em indivíduos normais cerca de 20% dos linfócitos recirculantes se encontram no estado ativado. E surpreendentemente, quase a totalidade desses linfócitos ativos correspondem aos que reagiram com componentes próprios do organismo. Somente uma pequena parte corresponde aos que foram ativados devido a estímulos do meio-ambiente, tais como infecções subclínicas de microorganismos, antígenos provenientes da microbiota intestinal ou proteínas da alimentação.

Como estão ativos, esta parte dos linfócitos recirculantes secretam imunoglobulinas idênticas às que são expressas por eles. Esses anticorpos formados pelo organismo em condições sadias são denominados anticorpos naturais (NAb). A função primária e mais importante, proposta por Boydem em 1965 e por Jerne, em sua Teoria da Rede Idiópica, dos NAb seria mediar a interação inter-linfocitária responsável pelo controle da auto-reatividade imune impedindo a expansão descontrolada dos linfócitos auto-reativos.

Dada a natureza múltipla da interação interlinfocitária, decorre que cada linfócito pode, em teoria, vir a ser afetado por uma alteração funcional ou quantitativa de outro linfócito, alteração esta que se propagaria no sistema imune através de interações em cadeia. Embora não se conheça a extensão de tais interações, estas contribuem de forma substancial para a seleção da quantidade de células B e T e com quais outras moléculas elas reagem. Assim a população de linfócitos (repertório) que um indivíduo apresenta em um dado momento, é determinada pela “topologia” (qual linfócito interage com qual, de que forma, qual é sua funcionalidade) desta rede de interações mútuas.

## 1.2 Sistemas Complexos

Nesta visão do sistema imune, suas sofisticadas propriedades, como auto-regulação, memória, tolerância a componentes próprios do organismo, surgem devido à existência de interações inter-linfocitárias. Assim, esse rico comportamento que possibilita que o sistema imunológico execute a tarefa de proteger o organismo contra corpos estranhos é fruto da interação entre seus componentes. O comportamento do sistema não é derivado do comportamento de seus componentes isolados uns dos outros, mas

sim do comportamento de seus componentes quando interagem entre si.

Embora este trabalho se dedique ao tratamento de dados experimentais, uma revisão sobre os sistemas complexos se torna interessante, aqui, pois, além de motivar o uso de várias ferramentas estatísticas, que serão utilizadas nos capítulos seguintes, dá uma noção de como o sistemas imune pode ser abordado no contexto da física estatística. A tendência na física é de procurar modelos mais simples possíveis mas que, mesmo assim, sejam capazes de apresentar um comportamento, em algum sentido, “complexo”.

Um exemplo simples de um sistema cujas propriedades dependem fortemente das interações entre seus componentes é dado pelo modelo de Ising [7]. Embora o sistema imune seja infinitamente mais complexo que o modelo do Ising, esse exemplo serve para ilustrar as repercussões das interações entre os elementos no comportamento do sistema. Este modelo, que é utilizado para explicar a magnetização espontânea de materiais ferromagnéticos, é constituído por spins, dispostos nos sítios de uma rede, que são caracterizados pelo seu momento angular que pode ter dois valores, um positivo ou outro negativo com ambos de mesmo módulo ( $\sigma = \pm 1$ ). O modelo considera, ainda, a interação de troca entre os primeiros vizinhos de cada sítio. Esta interação, de natureza quântica, faz com que haja uma menor energia armazenada quando dois sítios vizinhos, digamos  $i$  e  $j$ , têm o mesmo valor do momento angular de spin ( $\sigma_i = \sigma_j$ ). Utilizando a mecânica estatística, pode-se deduzir analiticamente que este sistema simples apresenta um comportamento peculiar. Para baixas temperaturas o momento angular médio de spin (magnetização) tem seu valor máximo, indicando que o spin da grande maioria dos sítios tem o mesmo sinal, que continua o mesmo quando se aumenta gradativamente a temperatura, até que a uma certa temperatura a magnetização cai abruptamente a zero, indicando que a fração dos spins com valor positivo é meio. A queda na magnetização é de tal forma abrupta que faz sua derivada primeira divergir (Figura 1.6). Nessa temperatura crítica, os grupos de spins vizinhos com mesmo sinal de spin formam uma estrutura fractal (descrição detalhada na secção 2.2), fazendo com que o comprimento de correlação<sup>2</sup> entre spins tenda a infinito.

---

<sup>2</sup>Definida a correlação  $C(r)$  como a probabilidade de se encontrar os sítios, afastados de uma distância  $r$ , com mesmo valor de spin, definimos o alcance  $\xi$  de correlação pela relação  $C(r) \sim e^{-r/\xi}$ . No ponto crítico,  $C(r)$  decai algebricamente e assim  $\xi$  é infinito.

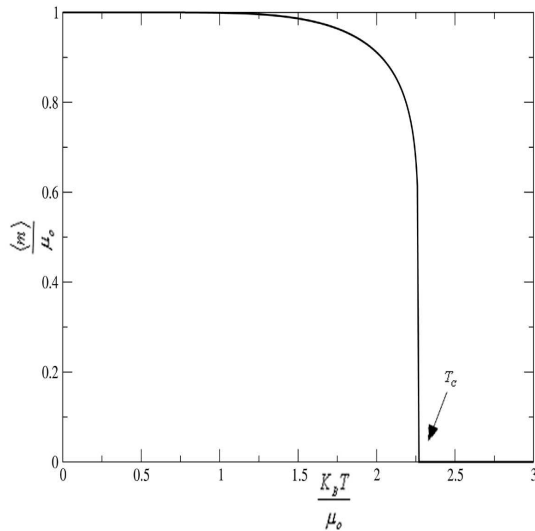


fig.1.6a

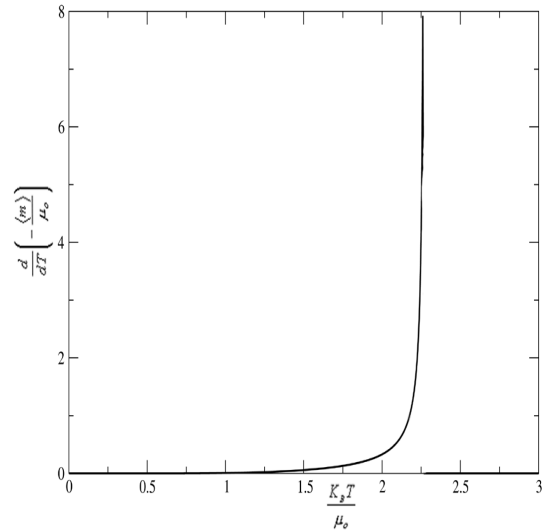


fig.1.6b

**Figura 1.6:** Solução exata para o modelo de Ising bi-dimensional numa rede quadrada infinita. Na parte (a) está o gráfico da magnetização e na parte (b) a sua derivada. Figura baseada em [7]

Todo este rico comportamento é apagado quando se resolve o problema considerando que a interação entre os sítios é desprezível, ou seja, quando estudamos um sistema paramagnético cujos sítios são isolados uns dos outros [7]. Neste caso, aplicando-se um campo externo para alinhar inicialmente os spins, a magnetização tem seu valor máximo para temperaturas baixas e cai gradativamente a zero com o aumento da temperatura. Sua derivada primeira é suave e o comprimento de correlação é sempre finito (Figura 1.7).

A interação entre os sítios influencia de forma crucial o comportamento do sistema como um todo. Quando há a interação, o sistema apresenta uma transição de fase, ou seja, a mudança abrupta na magnetização numa temperatura bem definida, e uma estrutura fractal nesta temperatura. Já quando não há a interação entre os sítios, o sistema não apresenta essas inusitadas propriedades. O sistema tem propriedades bem diferentes do sistema interagente.

Podemos aumentar um pouco a complexidade do sistema, modificando a interação entre os sítios, fazendo com que o sistema apresente ainda mais propriedades novas. Por exemplo, trocando, aleatoriamente, uma porcentagem considerável das intera-

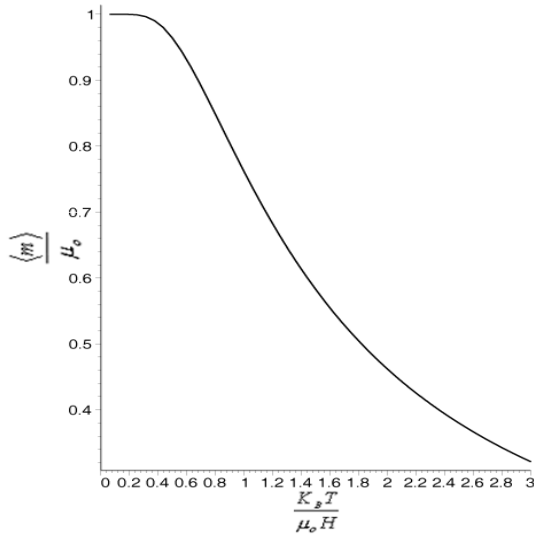


fig.1.7a

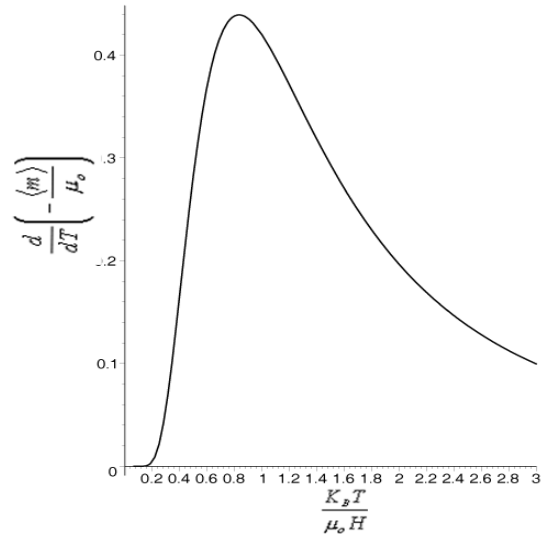


fig.1.7b

**Figura 1.7:** Solução exata para o modelo de Ising sem interação numa rede quadrada infinita na presença de um campo externo. Na parte (a) está o gráfico da magnetização e na parte (b) sua derivada. Figura baseada em [7]

ções do modelo de Ising por interações que fazem com que a energia armazenada seja menor quando dois vizinhos tenham seus spins com sinais opostos, que ficou conhecido como vidro de spins [8]. Agora alguns pares de vizinhos tendem a manter o mesmo valor de spin para minimizar a energia e outros pares tendem a mantê-los com valores opostos. Isso faz com que nem todas as interações possam ser satisfeitas, isto é, ficarem com energia mínima, simultaneamente. Este fenômeno é chamando frustração. O valor de spin de cada sítio será aquele que deixe o menor número de interações insatisfeitas e, geralmente, existem várias configurações com essa propriedade. Podemos reduzir lentamente a energia desse sistema, a partir de uma distribuição de valores de spin qualquer, até que se alcance seu valor mínimo. O interessante é que repetindo diversas vezes tal procedimento, não se encontra a mesma energia mínima. Diferentes estados finais com energias “minimizadas” serão atingidos após cada procedimento (Figura 1.8). Essas energias são mínimos locais, correspondentes a estados metaestáveis nos quais o sistema irá permanecer por longos períodos se forem feitas somente pequenas mudanças na configuração de seus spins. Tal propriedade não estava presente no modelo de Ising no qual o mínimo de energia é global e alcançado quando todos os spins tiverem o mesmo valor.





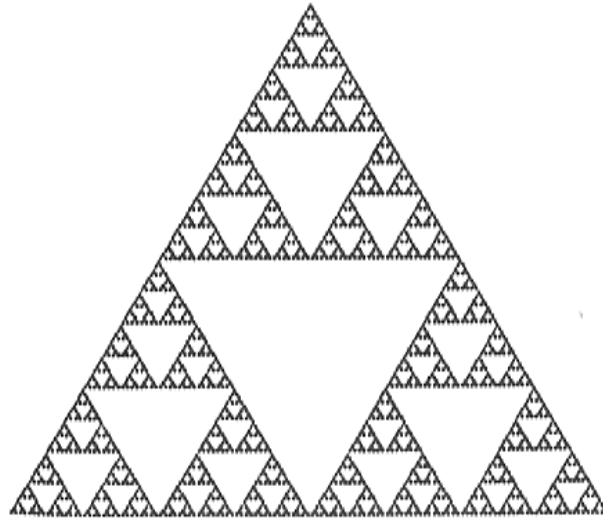
**Figura 1.8:** Ilustração de como seria um gráfico da energia em função dos estados para um sistema sem estados meta-estáveis (a) e para um sistema com estados meta-estáveis (b). Figura baseada em [8]

Ilustrando mais claramente o quanto a interação entre os componentes de um sistema pode influenciar o seu comportamento, introduzimos aqui alguns sistemas que ficaram conhecidos como autômatos celulares [9, 10]. Genericamente, estes sistemas são constituídos por sítios que são caracterizados pelo seu estado, que evolui no tempo discreto ( $t = 1, 2, 3, \dots$ ). O estado de cada sítio pode mudar de acordo com alguma regra. Existe uma infinidade de diferentes regras que podemos definir para governar a dinâmica dos estados destes sistemas. Um grande número de regras foram tabeladas, algumas dessas levando os sistemas a apresentarem uma grande gama de diferentes comportamentos. Comportamentos estes que são gerados pela regra escolhida para governar o sistema, reforçando a idéia de que as interações entre os componentes são cruciais no comportamento do sistema.

Como exemplo desse tipo de sistema considere uma linha unidimensional na qual sítios igualmente espaçados estão dispostos [8]. Cada um desses sítios pode estar cheio ou vazio. E a regra de evolução destes estados é a seguinte: caso somente um dos vizinhos de um sítio estiver cheio, no próximo passo de tempo este sítio também estará cheio; caso os dois vizinhos estiverem ambos cheios ou vazios, o sítio estará vazio no próximo passo de tempo. A seqüência de configurações que esse sistema apresenta ao longo do tempo forma uma estrutura fractal como mostrado na figura 1.9.

Como consequência disso, o número  $N$  de sítios vizinhos (consideramos vizinhos, agora, sítios que estão lado a lado num dado passo de tempo, e também, um mesmo sítio em dois passos de tempo consecutivos) no mesmo estado, esse conjunto de sítios chamaremos agrupamento, como função da freqüência  $f$  com que o agrupamento

deste tamanho aparece é uma lei de potência ( $N = f^\tau$ ), o que faz com que não haja uma escala bem definida no sistema, relação essa válida quando de inicia com apenas um sítio cheio. Tanto *clusters* pequenos como *clusters* muito grandes são encontrados, diferindo apenas pela frequência com que aparecem. Isto se deve à existência de uma interação (regra de evolução) entre os sítios, de tal forma que a influência do estado de um sítio sobre os demais se estende por quase todo o sistema. Influência esta que se dá através de uma cadeia de interações entre os vizinhos, um sítio interfere no estado do seu vizinho que interfere no estado do seu vizinho, e assim por diante, até que todo o sistema seja influenciado. Assim, dependendo da configuração do sistema, um sítio pode alterar o estado de apenas alguns poucos sítios próximos a ele ou também alterar quase toda a configuração do sistema. A mesma coisa ocorre no modelo de Ising na temperatura crítica, o que pode ser confirmado pela divergência no comprimento de correlação entre os spins.



**Figura 1.9:** Seqüência de estados do autômato celular, onde um ponto preto indica um sítio cheio e um ponto branco um sítio vazio. O estado inicial é apenas um sítio cheio e a evolução temporal é de cima para baixo. O conjunto fractal gerado é conhecido como tapete de Sierpinski. Figura baseada em [8]

Outro exemplo deste tipo de sistema são os modelos que ficaram conhecidos como pilhas de areia (Figura 1.10). Numa versão mais simples, cada sítio de uma rede

quadrada  $d$ -dimensional é caracterizado por um número inteiro que corresponde ao número de grãos da pilha de areia naquele ponto da rede [8]. Na dinâmica do sistema, a inclinação da pilha em algum ponto pode se tornar grande, maior que uma certa inclinação crítica pré-definida. Caso isto ocorra, uma quantidade de grãos de areia do sítio irá “deslizar” indo para seus vizinhos de acordo com a regra:

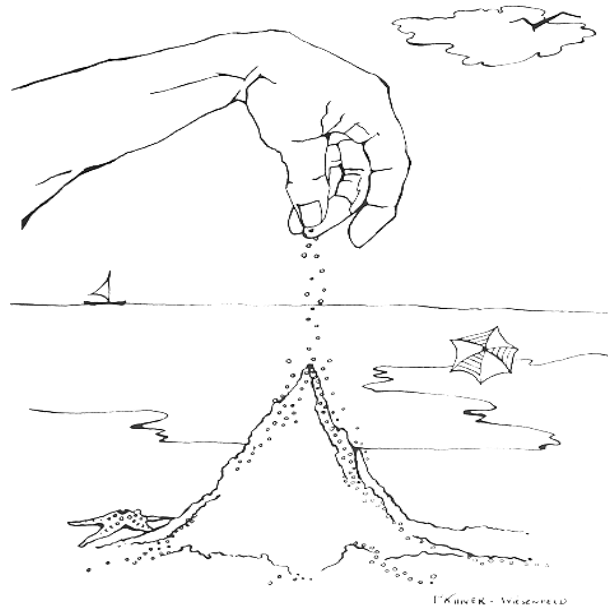
$$z \rightarrow z - 2d \quad (1.1)$$

$$z_{nn} \rightarrow z_{nn} + 1 \quad (1.2)$$

sendo  $z$  o número de grão da pilha naquele sítio e  $z_{nn}$  o de seus vizinhos. Se, após esta relaxação, a inclinação crítica de algum dos vizinhos for excedida, aplica-se novamente a regra até que a avalanche de difusão da inclinação pare. Inicia-se o sistema com uma distribuição aleatória de inclinações e, a partir deste estado, interage-se o sistema até que não haja mais avalanches. Depois, adiciona-se grãos de areia em algum sítio, escolhido aleatoriamente, aumentando de uma unidade o valor de sua inclinação, sempre mantendo  $z = 0$  na fronteira da rede. Este sistema também apresenta uma lei de potência, agora tanto na frequência das avalanches que atingem um número  $s$  de sítios, quanto na frequência do tempo de duração das avalanches. No entanto, a interação entre os sítios é de tal forma que este comportamento não aparece somente para um determinado valor de algum parâmetro do sistema, como ocorre no modelo de Ising que só apresenta uma lei de potência na temperatura crítica, e sim sem o ajuste do valor de algum determinado parâmetro. Este fenômeno ficou conhecido como criticalidade auto-organizada.

Um enorme número de sistemas com estas propriedades podem ser listados aqui: terremotos que apresentam lei de potência na frequência com que ocorrem terremotos que liberam energia  $E$  (Figura 1.11); Obras literárias também apresentam uma lei de potência na frequência com que as palavras aparecem no texto (Figura 1.12).

O comportamento de sistemas formados por componentes interagentes, cuja interação é de tal forma que a influência do estado de um componente se estende por quase todo o sistema, é derivado basicamente da forma de interação entre seus componentes. Estes sistemas ficaram conhecidos como sistemas complexos [11, 12].

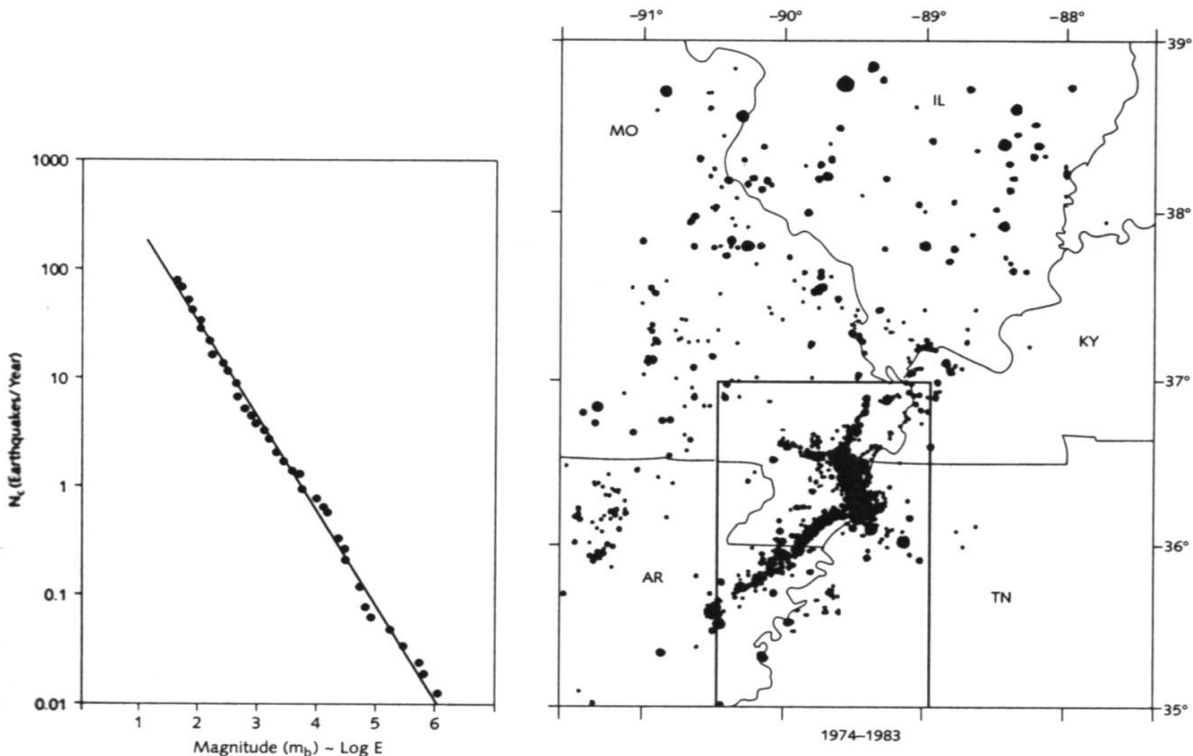


**Figura 1.10:** Ilustração pictórica de uma pilha de areia. Figura baseada em [8]

### 1.3 Sistema Imune como um Sistema Complexo

O sistema imune é um desses sistemas em que a interação é responsável pelas propriedades que ele apresenta [6, 13, 14, 15]. No sistema imune, a interação inter-linfocitária desempenha um papel crucial no seu funcionamento, pois conecta os comportamentos dos vários linfócitos através da supressão de um certo linfócito devido à ativação de um outro linfócito. Esta ação de supressão e ativação pode influenciar o comportamento de todo o sistema através de uma cadeia de reações na qual a supressão de um linfócito faz com que um outro linfócito, que era suprimido por ele, possa voltar a apresentar resposta, podendo tornar-se ativo novamente. Este linfócito, por sua vez, pode suprimir a resposta de um terceiro linfócito que irá influenciar o comportamento de outros linfócitos e assim por diante até que quase todo o sistema seja influenciado.

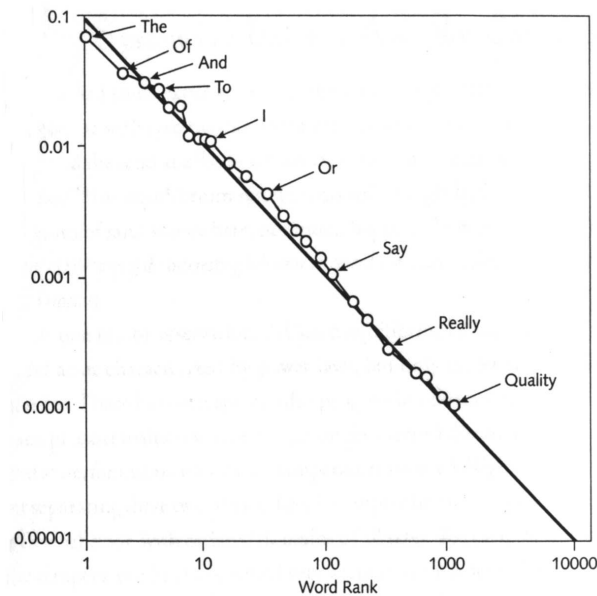
Este tipo de reação em cadeia pode fazer com que o sistema apresente propriedades bem distintas das que o sistema apresenta quando não há este tipo de interação. Portanto, para obtermos as propriedades globais deste tipo de sistema, temos que fazer medidas sobre este de tal forma que a interação entre seus componentes seja preservada. Se o processo de medida impossibilita que haja a interação, estaremos



**Figura 1.11:** Lei de potência encontrada na frequência com que ocorrem terremotos que liberam energia  $E$ , à esquerda. À direita, a região em que esses terremotos ocorreram. Figura baseada em [8]

medindo as propriedades dos componentes do sistema isolados uns dos outros, que podem diferir bastante das propriedades globais do sistema. Este último tipo de medida é que tradicionalmente aplicou-se na imunologia. A propriedade de interesse a ser medida era a reatividade de uma certa imunoglobulina. Esta imunoglobulina era isolada das demais que estavam presentes no soro de algum indivíduo e media-se com quais substâncias e com qual intensidade ela reagia. O fato da imunoglobulina ser isolada das demais torna claro que a medida é feita sobre as propriedades dos componentes do sistema e não do sistema como um todo.

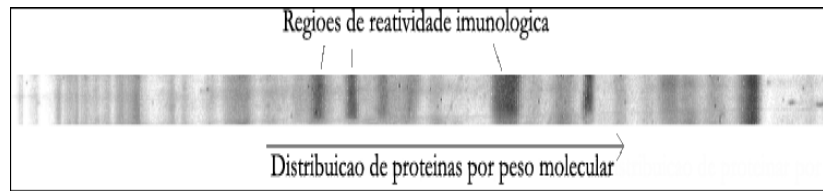
Neste trabalho estamos interessados nas propriedades globais do sistema imunológico, portanto foram feitas medidas que refletem a interação inter-linfocitária presente no organismo. Durante a dinâmica do sistema imune, a interação inter-linfocitária, com sua ação ativadora/supressora, define quais linfócitos irão apresentar resposta imunológica e com qual intensidade o fazem. Isto se reflete diretamente em quais,



**Figura 1.12:** Lei de potência encontrada na frequência com que as palavras aparecem num texto de língua inglesa. Figura baseada em [8]

e com qual quantidade, as imunoglobulinas são encontradas no organismo, pois dependendo da intensidade da resposta, os linfócitos irão produzir uma maior ou menor quantidade de imunoglobulinas. Uma medida de quantas imunoglobulinas de cada tipo estão presentes num organismo irá refletir o comportamento do sistema como um todo, pois a propriedade medida, a intensidade de resposta dos linfócitos, é uma consequência da interação entre os componentes do sistema. Ou seja, este tipo de medida preserva a interação entre os componentes, refletindo as propriedades do sistema e não as de seus componentes isolados.

A forma utilizada aqui para medir as quantidades de imunoglobulinas foi extrair o soro sanguíneo de um indivíduo e colocá-lo para reagir com uma grande quantidade de diferentes proteínas, obtidas de um extrato de fígado ou cérebro, previamente separadas por peso molecular através do processo de eletroforese e fixadas numa membrana de nitro-celulose. Após a reação, corando-se as imunoglobulinas que se ligaram ao painel de proteínas, podemos ver com quais proteínas houve reação e com qual intensidade. O resultado deste experimento é o que chamamos de perfil de reatividade imunológica (Figura 1.13).



**Figura 1.13:** Perfil de reatividade obtido em um experimento. Os tons de cinza da imagem indicam a intensidade da reação naquele ponto. As proteínas estão distribuídas com as de maior peso molecular à esquerda, e as de menor peso molecular à direita.

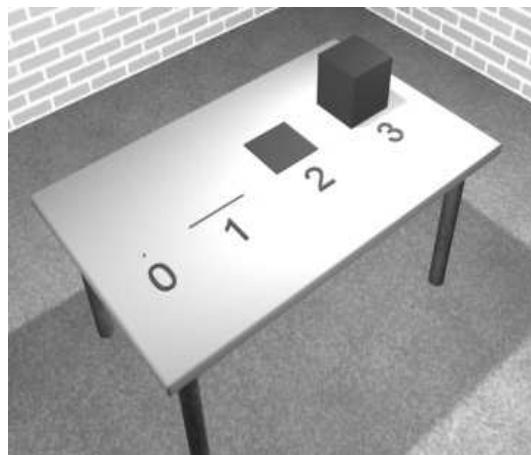
Iremos aqui fazer uma análise estatística desse perfil, visando a estudar as propriedades que não são vistas nos experimentos de medida da reatividade individual das imunoglobulinas. Uma das análises baseia-se nas propriedades fractais do perfil de reatividade, mais especificamente no expoente de Hurst, por isso é encontrado no Capítulo 2 uma descrição detalhada sobre os fractais, assim como suas propriedades e como estudá-las numericamente. Duas outras análises foram também aplicadas aos perfis. Uma delas é uma análise multivariada baseada em álgebra linear denominada Análise de Componentes Principais (PCA). A outra é baseada nas propriedades do modelo de Potts com interação não homogênea entre os vizinhos. Ambas técnicas estão descritas no Capítulo 2. Toda uma descrição dos tratamentos feitos nos perfis até que fossem feitas as análises e do experimento de obtenção destes perfis está no Capítulo 3. E os resultados dessas análises estão apresentados e discutidos no Capítulo 4.

# Capítulo 2

## Revisão Bibliográfica

### 2.1 Fractais

Observando a figura 2.1, é esperado e completamente natural que se diga que o objeto de número zero trata-se de um ponto e que tem dimensão zero<sup>1</sup>. Que o segundo objeto define um segmento de reta cuja dimensão é 1. Que o terceiro objeto define uma seção de um plano e que tem dimensão 2. E que o quarto objeto é tri-dimensional.



**Figura 2.1:** Objetos com dimensões 1, 2 e 3.

---

<sup>1</sup>O texto dessa seção está baseado nas referências [16, 17, 18]



No entanto, a pergunta: “Qual a dimensão de dois pontos?” provoca grande estranheza, porque este conceito não está presente no nosso cotidiano. Será que faz sentido perguntas deste tipo? Se fizer, como definir a dimensão de um conjunto de pontos, de segmentos de reta, de esferas, etc?

O campo elétrico de duas cargas puntuais afastadas de uma distância  $d$ , quando olhado em comprimentos muito maiores que este, é igual ao de uma carga puntual. Mas para comprimentos próximos de  $d$ , o campo não se assemelha nem um pouco com o de uma carga puntual. Quando estamos extremamente afastados dos dois pontos, estes se comportam como se fossem um só ponto, como um único objeto, mas na sua proximidade eles se comportam como se não o fossem, como um objeto com dimensão diferente de 1. Devido a esta distinção, o fato de que para certas escalas o conjunto é unidimensional e para outras não o é, não parece fazer sentido o conceito da dimensão neste caso. Pois a dimensão é uma propriedade intrínseca do conjunto, e não deve depender da escala na qual o observamos.

Mas se o conjunto em questão, ao contrário do caso anterior, apresentar o mesmo comportamento para qualquer escala, propriedade conhecida como invariância por escala, faz sentido o conceito de dimensão. Mas como definir a dimensão destes conjuntos invariantes por escala? A forma mais natural é generalizando o conceito de dimensão utilizado para classificar os objetos da figura 2.1.

Um segmento de reta de comprimento  $L$  pode ser dividido em  $N$  sub-segmentos de comprimento  $\epsilon$ . Considerando a linha inicial com comprimento unitário, obtém-se

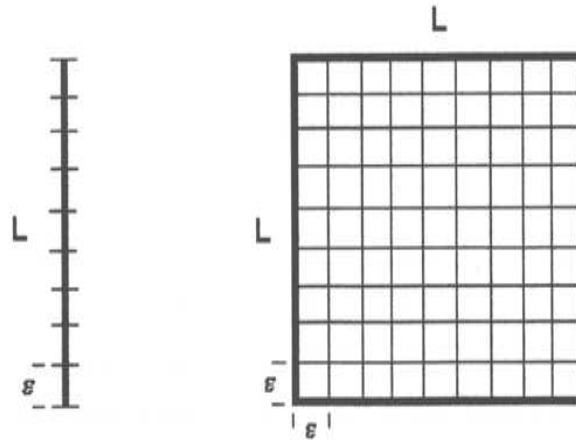
$$L = N\epsilon^1 = 1, \quad (2.1)$$

logo a linha unitária é composta de  $N$  partes de comprimento  $\epsilon = 1/N$ . Um quadrado, inicialmente com área unitária, pode ser dividido em  $N$  quadrados com área  $\epsilon^2$  cada um. Tem-se assim

$$A = N\epsilon^2 = 1 \quad (2.2)$$

e a área unitária é composta de  $N$  quadrados de lado  $\epsilon$  (Figura 2.2). Aplicando a mesma lógica para o volume unitário, obtém-se

$$V = N\epsilon^3 = 1 \quad (2.3)$$



**Figura 2.2:** Linha de comprimento  $L$  unitário dividida em  $N$  elementos de comprimento  $\epsilon$  e quadrado de área  $A$  unitária também dividida em  $N$  elementos de área  $\epsilon^2$ . Figura baseada em [16]

Examinado as expressões acima, vê-se que o expoente de  $\epsilon$  em cada caso é a dimensão do objeto. Generalizando tem-se

$$N\epsilon^D = 1, \tag{2.4}$$

aplicando o logaritmo

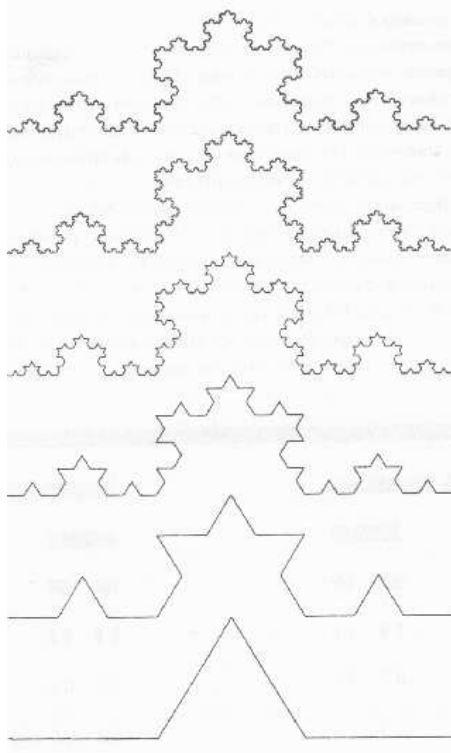
$$D = \frac{\log(N)}{\log(1/\epsilon)}. \tag{2.5}$$

Para estudarmos esta generalização da dimensão, agora chamada dimensão fractal, vamos aplicá-la a um conjunto simples de segmentos de reta, conhecido como curva de Koch. Para gerar este objeto iniciamos com um segmento de reta e retiramos seu segundo terço. No seu lugar adicionamos duas novas linhas, com mesmo comprimento da retirada, de modo que formaríamos um triângulo equilátero com o segmento que existia. Aplicando infinitas vezes este procedimento para cada novo segmento formado, obtemos uma curva com a propriedade de invariância por escala (Figura 2.3). Após  $n$  iterações são necessários  $4^n$  segmentos de comprimento  $(1/3)^n$  para cobrir toda a curva. Portanto sua dimensão é

$$D = \frac{\log(N)}{\log(1/\epsilon)} = \frac{n \log(4)}{n \log(3)} = 1.2618... \tag{2.6}$$

A curva de Koch apresenta uma dimensão não-inteira, assim como os demais conjuntos invariantes por escala. Esta é uma consequência da generalização da dimensão:

antes somente definia-se dimensão como um número inteiro, agora é possível encontrá-la como um número não-inteiro.



**Figura 2.3:** Ilustração do procedimento para gerar uma curva fractal de Koch. Figura baseada em [19]

Para que uma curva (aqui considerada como o gráfico de uma função real de uma variável real) seja invariante por escala, é necessário que

$$g(x) = b^{-H} f(bx) \quad (2.7)$$

onde o perfil  $g(x)$  resultante é o perfil  $f(x)$  reescalado, sendo estas funções estatisticamente equivalentes entre si, ou seja, suas densidades de probabilidade são as mesmas. O expoente  $H$  é chamado de expoente de Hurst e a curva que apresenta todas essas propriedades é fractal.

Pela propriedade de invariância por escala (equação 2.7), a largura da função dentro

de um intervalo com largura  $\epsilon$  se comporta como

$$W(\epsilon) \sim \epsilon^H. \quad (2.8)$$

Por simplicidade, irermos, aqui, discretizar a função. Assim, podemos medir a sua largura de maneira mais simples e intuitiva como

$$W(\epsilon) \equiv \frac{1}{N} \sum_x [\max_{\epsilon} f(x) - \min_{\epsilon} f(x)], \quad (2.9)$$

onde  $\max_{\epsilon} f(x)$  é o máximo de  $f(x)$  em um intervalo de largura  $\epsilon$  centrado em  $x$ ,  $\min_{\epsilon} f(x)$  é o mínimo da função no mesmo intervalo e  $N$  o número de pontos, do perfil discretizado, para os quais o intervalo não extrapola a região onde a função está definida. Outra forma mais fina de medir a largura da função é através da sua rugosidade média  $W(\epsilon)$ . Primeiro, definimos a rugosidade local  $w(x, \epsilon)$  como

$$w^2(x, \epsilon) \equiv \frac{1}{\epsilon} \sum_{x-\frac{\epsilon}{2} < y < x+\frac{\epsilon}{2}} [(f(y) - \bar{f}(y))]^2, \quad (2.10)$$

onde  $\bar{f}(x)$  pode ser a altura média ou a regressão linear da função no intervalo definido por  $\epsilon$  [20]. A seguir definimos a rugosidade média

$$W(\epsilon) \equiv \frac{1}{N+1} \sum_x w(x, \epsilon). \quad (2.11)$$

Dentro de uma escala  $\epsilon$ , a área ocupada pela função é a soma da sua largura em cada intervalo multiplicada pela distância entre dois pontos no perfil discretizado. No entento, essa distância entre dois pontos é uma constante e pode ser considerada como unitária, obtendo-se [21, 22, 23, 24]

$$A(\epsilon) \propto \epsilon^H. \quad (2.12)$$

Por outro lado, pela definição de dimensão, o número de intervalos necessários para recobrir a função é

$$N(\epsilon) \propto \epsilon^{-D}, \quad (2.13)$$

e a sua área

$$A(\epsilon) \propto N(\epsilon)\epsilon^d, \quad (2.14)$$

onde  $d$  é a dimensão do espaço no qual a curva está imersa. Resultando em

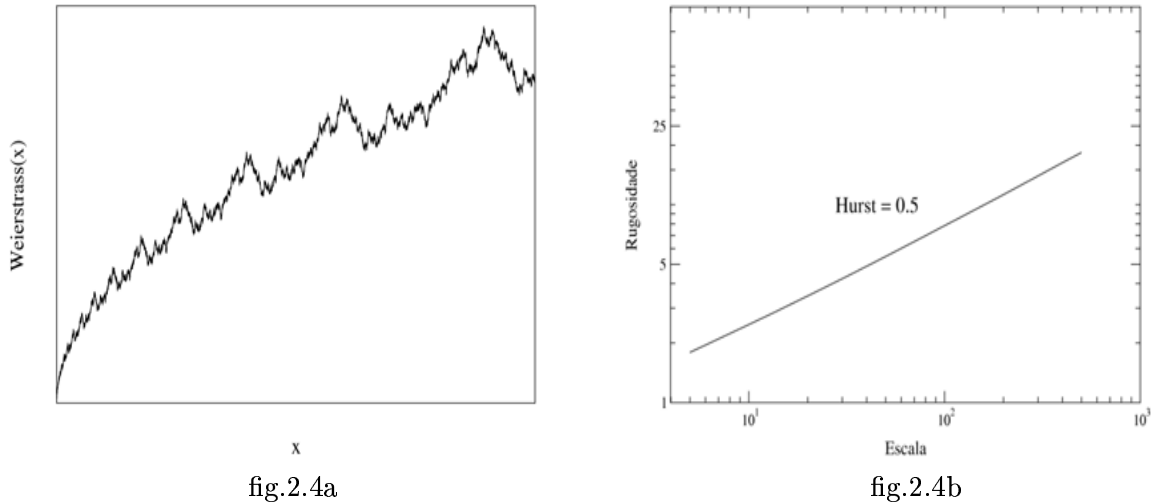
$$A(\epsilon) \propto \epsilon^{d-D}. \quad (2.15)$$

Comparando com a equação 2.12 temos

$$H = d - D, \quad (2.16)$$

que relaciona o expoente de Hurst com a dimensão fractal da curva considerada.

Esta relação é de extrema importância pois, com o cálculo do expoente de Hurst, obtemos facilmente a dimensão do conjunto. E a relação 2.8 nos permite calcular o expoente. Basta, dada a função, calcular sua rugosidade para várias escalas  $\epsilon$ . A rugosidade como função da escala é uma lei de potência com expoente definido como o expoente de Hurst (Figura 2.4). Portanto, com uma série de medidas da rugosidade podemos calcular a dimensão de um conjunto invariante por escala.



**Figura 2.4:** Na parte (a), uma curva fractal com expoente de Hurst 0.5 gerada pela função real de Weierstrass, definida como  $W(x) = \sum_{n=-\infty}^{\infty} \frac{1-\cos(\gamma^n x)}{\gamma^{Hn}}$ , onde  $n$  é um número inteiro,  $\gamma$  um real maior que 1 e  $H$  o expoente de Hurst. Na parte (b), a lei de potência da rugosidade como função da largura da escala, cujo expoente é o expoente de Hurst [17].

Existe uma outra maneira de se calcular a dimensão, que é através do espectro de potência da função. Para este cálculo iremos adotar a seguinte notação:

- $F_f(k, L)$  e  $F_g(k, L)$ , são as transformadas de Fourier de  $f(x)$  e  $g(x)$  no intervalo  $0 < x < L$ ;

- $s_f(k, L)$  e  $s_g(k, L)$ , são as densidades espectrais de  $f(x)$  e  $g(x)$  no intervalo  $0 < x < L$ ;

- $S_f(k)$  e  $S_g(k)$ , são os espectros de potências de  $f(x)$  e  $g(x)$  quando  $L \rightarrow \infty$ .

A partir da relação entre as funções  $f(x)$  e  $g(x)$ , dada pela equação 2.7, e usando a transformada de Fourier, temos

$$F_g(k, L) = \int_0^L g(x)e^{-ikx} dx = \frac{1}{b^H} \int_0^{bL} f(s)e^{-iks/b} \frac{ds}{b}, \quad (2.17)$$

onde  $s = bx$ . Podemos isolar na segunda integral os termos que não dependem de  $s$  e chegamos a

$$F_g(k, L) = \frac{1}{b^{H+1}} F_f\left(\frac{k}{b}, bL\right). \quad (2.18)$$

A densidade espectral de  $g(x, L)$  é

$$s_g(k, L) = \frac{1}{L} |F_g(k, L)|^2 = \frac{1}{b^{2H+1}} \frac{1}{bL} \left| F_f\left(\frac{k}{b}, bL\right) \right|^2 = \frac{1}{b^{2H+1}} s_f\left(\frac{k}{b}, bL\right). \quad (2.19)$$

No limite  $L \rightarrow \infty$ , de modo que  $bL \rightarrow \infty$ , temos

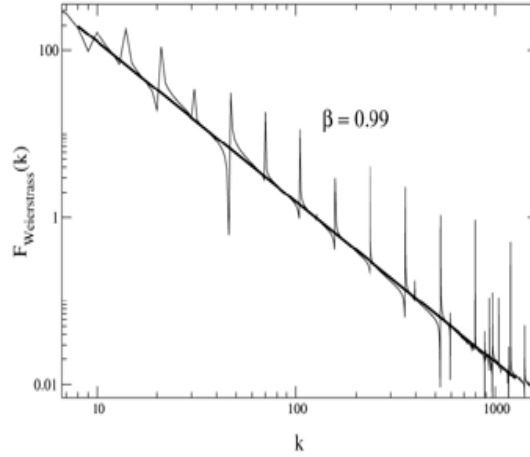
$$S_g(k) = \frac{1}{b^{2H+1}} S_f\left(\frac{k}{b}\right). \quad (2.20)$$

A equação acima é válida para quaisquer valores de  $k$  e  $b$ , e dessa maneira podemos fazer  $k/b = 1$  obtendo

$$S_g(k) \propto \frac{1}{b^{2H+1}} \propto \frac{1}{k^\beta}. \quad (2.21)$$

sendo  $\beta = 2H + 1$ . Assim uma função que seja invariante por escala apresenta seu espectro de potência como um lei de potência cujo expoente se relaciona diretamente com o expoente de Hurst (Figura 2.5).

O impressionante é que, além um devaneio da mente humana, os fractais são encontrados a todo instante em formas criadas pela natureza ou em comportamento de grandezas de vários sistemas [19]. Inúmeros exemplos desses objetos podem ser dados aqui, como a topografia de montanhas, as ramificações do pulmão, as nuvens *cumulus*, redes pluviais, a figura impressa em algumas conchas, etc (Figura 2.6).



**Figura 2.5:** Transformada de Fourier da mesma curva da figura anterior. Novamente observa-se uma lei de potência cujo expoente se relaciona com o expoente de Hurst.

Além de ser um valor relacionado com a dimensão, o expoente de Hurst também é uma importante medida das correlações existentes na curva analisada. Em outras palavras, o expoente representa quanto o próximo passo depende do anterior. Para estabelecer esta relação consideremos o *Movimento Browniano Fracionário (FBM)*<sup>2</sup>. Seja  $P(x(t) - x(t_0))$  a distribuição de probabilidade de uma partícula executando um FBM

$$P(x(t) - x(t_0)) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left[-\frac{(x(t) - x(t_0))^2}{\sigma}\right], \quad (2.22)$$

onde

$$\langle x(t) - x(t_0) \rangle = 0, \quad (2.23)$$

$$\langle [x(t) - x(t_0)]^2 \rangle \equiv \sigma = |t - t_0|^{2H}. \quad (2.24)$$

Podemos definir a função de correlação  $C(t)$  que mede qual a probabilidade da partícula experimentar um deslocamento  $x(t) - x(0)$  se anteriormente houver um deslocamento de  $x(0) - x(-t)$

$$C(t) = \frac{\langle [x(t) - x(0)][x(0) - x(-t)] \rangle}{\langle x(t)^2 \rangle} \quad (2.25)$$

<sup>2</sup>Neste movimento, uma partícula se desloca uma quantidade  $\Delta x \equiv x(t_2) - x(t_1)$  em cada instante de tempo. O comprimento desse deslocamento segue uma distribuição Gaussiana de modo que  $\langle [x(t_2) - x(t_1)]^2 \rangle = |t_2 - t_1|^{2H}$ , onde  $H$  é o expoente de Hurst.

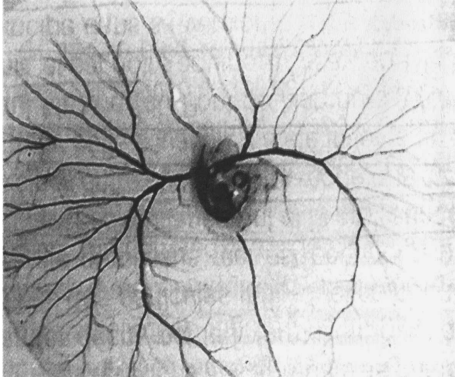


fig.2.6a

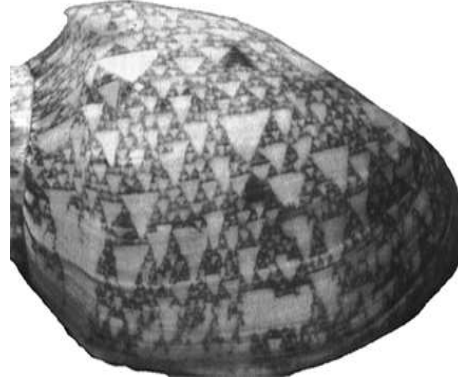


fig.2.6b

**Figura 2.6:** Fotos de estruturas fractais encontradas na natureza. Em (a) um feto de galinha e em (b) uma concha. Figura baseada em [19]

Podemos escolher, por questão de facilidade,  $x(0) = 0$ , logo

$$C(t) = \frac{-\langle [x(t)][x(-t)] \rangle}{\langle x(t)^2 \rangle}. \quad (2.26)$$

Expandindo o termo  $\langle [x(t) - x(t_0)]^2 \rangle$ , temos

$$\langle x(t)^2 \rangle - 2\langle x(t)x(t_0) \rangle + \langle x(t_0)^2 \rangle = |t - t_0|^{2H}. \quad (2.27)$$

Fazendo

$$t_0 = 0 \quad \Rightarrow \quad \langle x(t)^2 \rangle = t^{2H}, \quad (2.28)$$

e

$$t_0 = -t \quad \Rightarrow \quad \langle x(t)x(-t) \rangle = t^{2H} - \frac{1}{2}(2t)^{2H}. \quad (2.29)$$

Substituindo na equação 2.26 temos

$$C(t) = 2^{2H-1} - 1. \quad (2.30)$$

De acordo com essa última relação, podemos claramente estabelecer a relação entre  $H$  e a correlação

$$H > \frac{1}{2} \quad \Rightarrow \quad C(t) > 0, \quad (2.31)$$

$$H < \frac{1}{2} \quad \Rightarrow \quad C(t) < 0. \quad (2.32)$$



$$H = \frac{1}{2} \quad \Rightarrow \quad C(t) = 0, \quad (2.33)$$

Se  $H > \frac{1}{2}$ , há uma correlação positiva ou persistente entre os passos do movimento, o que significa que o próximo passo tende a ser igual ao anterior, resultando em curvas mais lisas. Se  $H < \frac{1}{2}$ , há uma correlação negativa ou anti-persistente. Nesse caso, o próximo passo possui a tendência de se opor ao anterior, ocasionando o aparecimento de curvas mais rugosas. Se  $H = \frac{1}{2}$ , há inexistência de correlações entre os passos, ou seja, um passo não influencia o comportamento do outro, como observado na caminhada aleatória.

## 2.2 Análise de Componentes Principais (PCA)

A análise que primeiro se aplicou, e tradicionalmente é utilizada pelos imunologistas para estudar os perfis de reatividade é a Análise de Componentes Principais (PCA) [25]. Esta é uma análise multi-variada que tem como objetivo reduzir a dimensão de um espaço  $d$ -dimensional para posteriormente se aplicar outras técnicas ou apenas para possibilitar a visualização dos pontos nele distribuídos. Para reduzir a dimensão do espaço basta projetá-lo num conjunto menor de vetores do que a base do espaço. Mas quais vetores escolher para projetar adequadamente o espaço?

Para entendermos a resposta a esta pergunta, lembramos que estamos interessados em caracterizar os pontos pertencentes a grupos distintos. Se houver, a diferença no comportamento entre pontos de grupos distintos se reflete na existência de correlações entre as suas componentes; assim como, a semelhança no comportamento entre pontos de um mesmo grupo. Se a projeção for feita em um conjunto no qual ainda haja correlação entre as componentes dos pontos, iremos eliminar as correlações entre as componentes que foram perdidas com a projeção. Portanto, para não perdermos correlações que possibilitem a diferença no comportamento entre grupos, temos que projetar o espaço num conjunto de vetores nos quais não haja mais correlação entre os pontos nele distribuídos.

Para encontrar este conjunto, calculamos a matriz da covariância entre os pontos

$\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N$ , definida como:

$$\begin{pmatrix} \text{Cov}(\mathbf{x}_1, \mathbf{x}_1) & \text{Cov}(\mathbf{x}_1, \mathbf{x}_2) & \dots & \text{Cov}(\mathbf{x}_1, \mathbf{x}_d) \\ \text{Cov}(\mathbf{x}_2, \mathbf{x}_1) & \text{Cov}(\mathbf{x}_2, \mathbf{x}_2) & \dots & \text{Cov}(\mathbf{x}_2, \mathbf{x}_d) \\ \vdots & \vdots & \dots & \vdots \\ \text{Cov}(\mathbf{x}_d, \mathbf{x}_1) & \text{Cov}(\mathbf{x}_d, \mathbf{x}_2) & \dots & \text{Cov}(\mathbf{x}_d, \mathbf{x}_d) \end{pmatrix} \quad (2.34)$$

onde

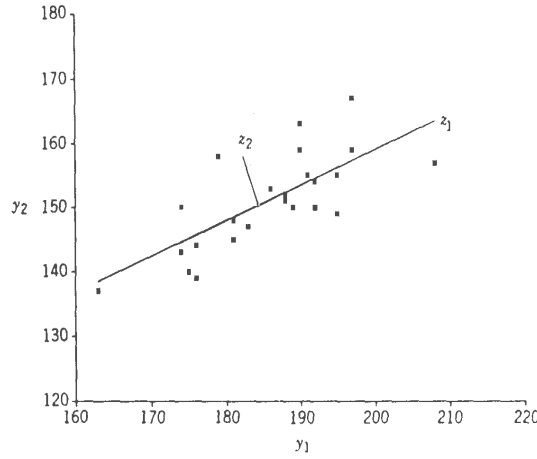
$$\text{Cov}(\mathbf{x}_i, \mathbf{x}_j) = \langle (\mathbf{x}_i - \langle \mathbf{x}_i \rangle) \cdot (\mathbf{x}_j - \langle \mathbf{x}_j \rangle) \rangle \quad (2.35)$$

onde  $\mathbf{x}_i$  é o vetor formado pelas  $i$ -ésimas componentes dos vetores  $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N$ .

A matriz das covariâncias é Hermitiana, pois é real e simétrica, portanto sempre podemos escrevê-la numa base, formada pelos seus auto-vetores, em que ela se torne diagonal. Assim, quando projetamos os pontos nos auto-vetores da matriz das covariâncias, não existem mais correlações entre as componentes dos pontos, pois a matriz é diagonal nesta base. Agora, sabemos qual o conjunto adequado de vetores para projetar o espaço. No entanto, ainda continuamos com um espaço  $d$ -dimensional. Para diminuir essa dimensão temos que escolher alguns dos auto-vetores e projetar o espaço apenas neles. Os auto-vetores escolhidos são aqueles que têm maiores auto-valores, pois os pontos apresentam maior dispersão quando projetados nestes vetores, facilitando a visualização dos diferentes grupos.

Um exemplo pedagógico da eliminação das correlação através deste método é dado considerando a distribuição de pontos, no espaço bi-dimensional, da figura 2.7.

Quando escritos na base formada pelos eixos  $y_1$  e  $y_2$  os pontos apresentam um correlação entre si, pois têm a tendência de seguir uma linha com certa inclinação. Esta correlação se reflete no fato de que, calculando a matriz de covariância, esta apresenta todo os elementos não nulos. Diagonalizando esta matriz, obtemos seus autovetores como sendo os eixos  $z_1$  e  $z_2$  mostrados na figura 2.7. Projetando os pontos nestes novos eixos, eles passam a não ter mais correlações entre si ( $-\text{Cov}(\mathbf{z}_1, \mathbf{z}_2)$ ), pois a tendência linear é eliminada e, os pontos se tornam homoganeamente distribuídos em torno dos eixos.



**Figura 2.7:** Distribuição de pontos nos eixos  $y_1$  e  $y_2$  e também os eixos  $z_1$  e  $z_2$  definidos pelos autovetores da matriz de covariância. Figura baseada em [25]

## 2.3 Agrupamento Super-paramagnético

A última análise aplicada, aqui, aos perfis, foi proposta por Domany e está baseada numa série de generalizações do modelo de Ising<sup>3</sup>. Como discutido brevemente na Introdução, o modelo de Ising consiste de  $N$  sítios dispostos em uma rede. Cada sítio é caracterizado pelo seu estado,  $+1$  ou  $-1$ , que representa o valor esperado da projeção do seu momento magnético de spin no eixo  $z$  (neste momento arbitrário). Os sítios interagem com seus primeiros vizinhos (os quatro sítios mais próximos a ele, no caso de uma rede quadrada) de forma que se dois vizinhos estão no mesmo estado, uma quantidade  $J$  é subtraída da energia do sistema; enquanto, se dois vizinhos estão em estados diferentes a mesma quantidade  $J$  é adicionado à energia. Assim, escrevemos o Hamiltoniano de cada sítio como

$$H_i = -J \sum_{\langle i,j \rangle} s_i s_j \quad (2.36)$$

onde a soma é feita sobre os primeiro vizinhos do sítio  $i$ , e  $J$  é uma constante positiva. Esta forma de interação proporciona ao sistema propriedades peculiares. Para estudá-las, definimos a magnetização

$$m = \frac{\sum_{i=1}^N s_i}{N} \quad (2.37)$$

---

<sup>3</sup>O texto dessa seção está baseado nas referências [26, 27, 28]

que é a fração da quantidade de sítios que um estado tem a mais que o outro. Se a magnetização é nula, o número de sítios em cada estado é o mesmo; se a magnetização é unitária, todos os sítios estão no mesmo estado. Podemos estudar este modelo tanto analiticamente, através da mecânica estatística, quanto simulando-o computacionalmente. Os resultados analíticos foram listados na introdução e a descrição de como estudá-lo computacionalmente será feita a seguir.

Devido à natureza do sistema, temos que estudar as suas propriedades médias, tais como magnetização média, energia média, etc. Por definição, o valor médio de uma variável aleatória  $A$  é

$$\langle A \rangle = \sum_i p_i A_i \quad (2.38)$$

onde  $p_i$  é a probabilidade da variável apresentar o valor  $A_i$ . Por exemplo, podemos calcular a magnetização média, neste caso  $A_i$  corresponde à magnetização para cada configuração de spins dos sítios. Para sistemas como o modelo de Ising, utilizando a mecânica estatística, podemos calcular a probabilidade de cada configuração [29] que é

$$p_i = \frac{e^{-\frac{E_i}{k_b T}}}{Z} \quad (2.39)$$

onde  $E_i$  é a energia do sistema naquela configuração de spins,  $k_b$  a constante de Boltzmann,  $T$  a sua temperatura e  $Z$  a função de partição. Portanto, para calcularmos a magnetização média basta encontrarmos todas as configurações possíveis, assim como a energia de cada uma.

Para sistemas com pequeno número de sítios, isso não representa nenhum grande problema, e pode ser facilmente calculado. Mas o número de configurações no modelo de Ising é  $2^N$ , que cresce vertiginosamente com o número de sítios, tornando inviável, devido ao tempo necessário para se encontrar todas as configurações, o cálculo das médias para sistemas grandes. Temos, então, que encontrar uma outra forma de obter os valores médios.

Esta outra forma está baseada no fato de que se simplesmente somarmos os valores  $A_i$ , de modo que a frequência relativa com que o termo  $A_i$  aparece nessa soma seja a mesma que a probabilidade  $p_i$ , encontraremos também o valor médio dessa variável. Temos que criar uma seqüência de configurações de forma que as que apresentam energia  $E_i$  se repitam, relativamente,  $p_i$  vezes na seqüência. Uma forma de obtermos

essa seqüência é definirmos uma caminhada no espaço de configurações (espaço no qual cada ponto representa uma configuração de spins da rede) cuja distribuição de probabilidades de equilíbrio seja dada por  $p_i$ . Esse método introduzido muitos anos atrás por Metropolis et al [30] consiste em definir um processo estocástico que apresenta esta propriedade iniciando com uma configuração qualquer de spins e, a cada passo de tempo, escolher um sítio e calcular a sua energia com seu estado original e com seu estado invertido de valor. Essa troca de estado é aceita com probabilidade

$$P(\Delta E) = \frac{1}{1 + e^{\frac{E_f - E_i}{k_b T}}} \quad (2.40)$$

onde  $E_i$  e  $E_f$  são, respectivamente, as energias antes e depois da inversão.

No entanto dois problemas são inerentes a esta forma de se calcular as médias. Após gerada a configuração inicial, devemos proceder um número tal de passos que duas diferentes configurações convirjam para o mesmo comportamento. Em certos casos, o número de passos necessários pode ser grande, demandando bastante tempo computacional. Outro fato é a grandeza de interesse relacionada com duas configurações consecutivas podem ser correlacionadas. Por isso, devemos proceder um número de passos, entre as configurações cuja grandeza será somada no cálculo da média, tal que seus valores sejam descorrelacionados. Novamente, este número de passos pode se tornar muito grande, principalmente para temperaturas próximas da crítica, demandando bastante tempo computacional. Como, na análise que iremos aplicar, estaremos sempre trabalhando próximos da temperatura crítica, esta não é uma técnica muito eficiente (Figura 2.8).

Existe um outro processo estocástico, introduzido por Swendsen e Wang [31], com a mesma distribuição de equilíbrio, mas que em vez de invertermos somente um spin a cada passo, invertemos um conjunto deles, fazendo com que o tempo necessário para descorrelacionar as configurações, seja consideravelmente menor (Figura 2.9). Nesta nova caminhada no espaço de configurações, iniciamos com uma configuração qualquer de spins, ligamos os sítios vizinhos que estão no mesmo estado com probabilidade

$$P = e^{\frac{-2J}{k_b T}} \quad (2.41)$$

e nunca ligamos sítios com estados diferentes. Após identificar todos os aglomerados,

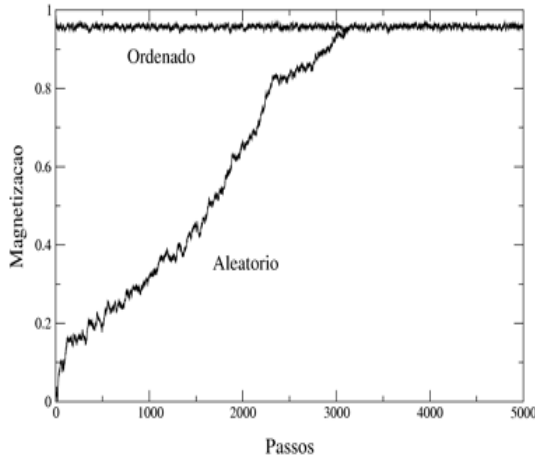


fig.2.8a

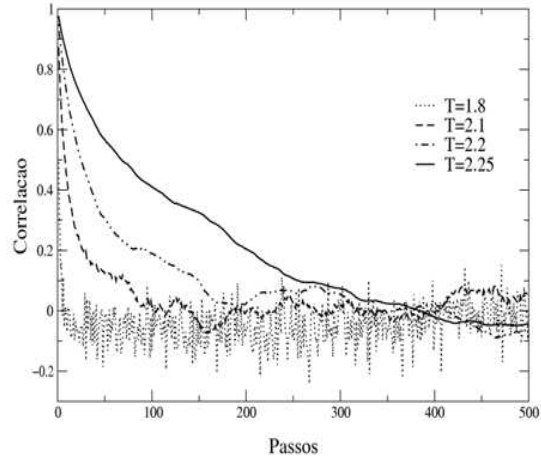


fig.2.8b

**Figura 2.8:** Gráficos obtidos por simulação do modelo de Ising mudando os spins um a um, para uma rede quadrada com 100x100 pontos. A parte (a) mostra o processo de termalização, em que duas configurações iniciais, uma com todos sítios no mesmo estado e outra aleatória, convergem para o mesmo comportamento. Na parte (b) o processo de decorrelação entre configurações subsequentes.

usando, aqui, o algoritmo de Kopelman [32], invertemos o spin de todos os conjuntos de sítios que foram ligados no processo anterior.

Através desta ultima técnica, simulamos o modelo de Ising numa rede quadrada, encontrando as suas propriedades, que são iguais às deduzidas analiticamente (Figura 2.10).

Em uma generalização do modelo de Ising, conhecida como modelo de Potts, o estado de cada sítio é caracterizado por um entre  $q$  valores. Este se reduz ao modelo de Ising quando  $q = 2$ . Temos que adaptar a definição da magnetização para este novo modelo, mas mantendo o mesmo conceito que o do modelo de Ising,

$$m = \frac{qN_{max} - N}{(q - 1)N}, \quad (2.42)$$

com

$$N_{max} = \max\{N_1, N_2, \dots, N_q\}, \quad (2.43)$$

onde  $N_\beta$  é o número de sítios que estão no estado  $\beta$ . Este modelo apresenta, basicamente, propriedades semelhantes às do modelo de Ising. No entanto, podemos utilizar uma generalização dele, proposta recentemente por Domany [26] e obtermos novas propriedades, que serão utilizadas na análise dos perfis de reatividade.

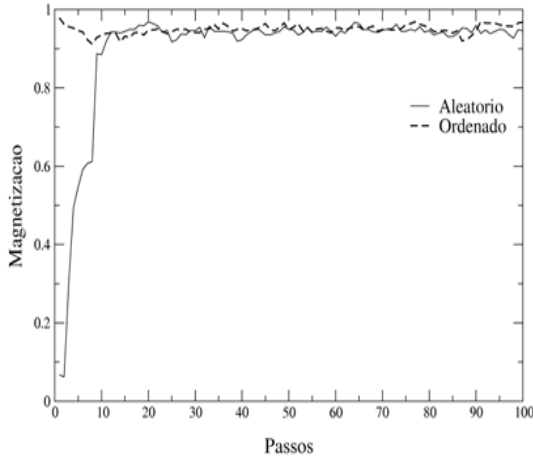


fig.2.9a

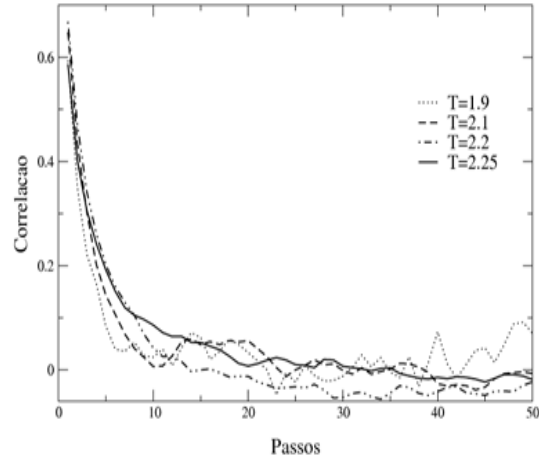


fig.2.9b

**Figura 2.9:** Gráficos obtidos por simulação do modelo de Ising mudando blocos de spins, para uma rede quadrada com 100x100 pontos. A parte (a) mostra o processo de termalização, em que duas configurações iniciais, uma com todos sítios no mesmo estado e outra aleatória, convergem para o mesmo comportamento. Na parte (b) o processo de decorrelação entre configurações sub-seqüente.

A generalização feita é permitir que os sítios sejam distribuídos continuamente no espaço, ao invés de fixá-los numa rede regular. Com isso, temos que definir como encontrar os vizinhos de cada sítio, operação que é simples numa rede regular. O conceito utilizado aqui é o de vizinhança mútua. O sítio  $i$  é vizinho do  $j$  se  $j$  está entre os  $K$  sítios mais próximos de  $i$ ; e também,  $i$  está entre os  $K$  sítios mais próximos de  $j$ . Assim, o número máximo de vizinhos que um sítio tem é  $K$ .

Uma segunda generalização é que a interação  $J$  entre os sítos vizinhos não é mais uma constante, e sim, uma função da distância entre os sítios. O comportamento exigido para essa função é que para distâncias menores que a distância média  $a$  entre todos sítios haja uma forte interação e, para distâncias maiores que esta, a interação seja fraca. A função utilizada aqui, que atende à esses requisitos, é

$$J_{ij} = \frac{1}{k} \exp\left(-\frac{d_{ij}^2}{2a^2}\right) \quad (2.44)$$

onde  $d_{ij}$  é a distância entre os sítios e  $\bar{k}$  o número médio de vizinhos. Esta interação define um escala local de interação, sítios próximos (com distância menor que  $a$ ) interagem fortemente e sítios afastados (com distância maior que  $a$ ) interagem fracamente. Isso faz com que, numa distribuição não-homogênea em que há algumas

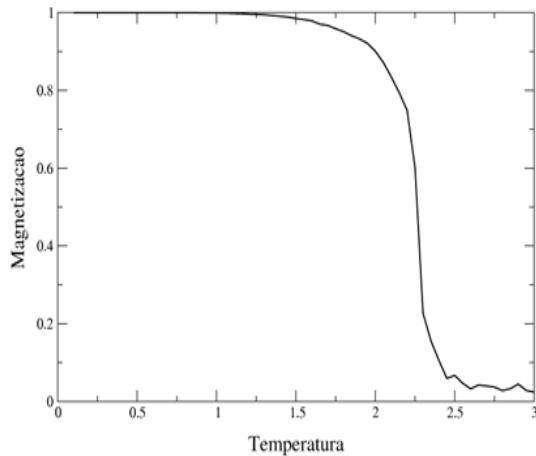


fig.2.10a

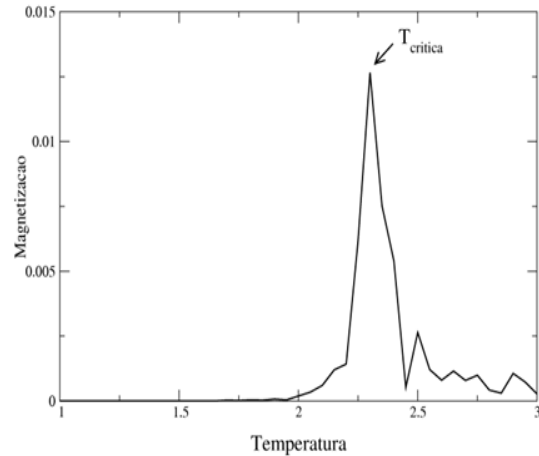


fig.2.10b

**Figura 2.10:** Gráficos obtidos por simulação do modelo de Ising mudando blocos de spins, para uma rede quadrada com 500x500 pontos. Na parte (a) a magnetização e na (b) a suscetibilidade, a partir das quais caracterizamos as duas fase e sua transição do modelo de Ising.

regiões de alta densidade de pontos e outras regiões de baixa densidade, haja uma forte interação dentro das regiões de alta densidade e interação fraca dentro das de baixa densidade.

Dada uma distribuição de sítios, podemos estudar este modelo, através das mesmas técnicas descritas para o modelo de Ising. Para temperaturas baixas o sistema apresenta magnetização unitária, estando todos os sítios no mesmo estado, assim no modelo de Ising ou de Potts em rede. Para temperaturas altas, a magnetização é nula, os estados estão igualmente distribuídos entre os sítios. No entanto, surge uma nova fase entre estas duas, chamada super-paramagnética, em que os spins de sítios pertencentes a um mesmo aglomerado estão fortemente correlacionados, enquanto spins de sítios de aglomerados diferentes estão fracamente correlacionados. Isso faz com que haja um patamar no gráfico da suscetibilidade devido às flutuações causadas pela mudança de estado dos aglomerados.

Podemos também identificar quais sítios pertencem a um mesmo aglomerado através da correlação que apresentam entre si na fase super-paramagnética. Nesta fase, aglomerados distintos estão fracamente correlacionados, mas sítios pertencentes a um aglomerado se comportam como estivessem na fase ferromagnética. Sendo a



correlação entre eles [28]

$$G_{ij} \equiv \langle \delta_{\sigma_i \sigma_j} \rangle = \frac{1}{q}, \quad (2.45)$$

onde  $\delta_{\sigma_i \sigma_j} = 1$  se  $\sigma_i = \sigma_j$  e  $\delta_{\sigma_i \sigma_j} = 0$ , caso contrário;  $\sigma_i$  é o estado do  $i$ -ésimo sítio.

Os spins pertencentes a um mesmo aglomerado estão fortemente correlacionados, sendo esta correlação estimada por [28]

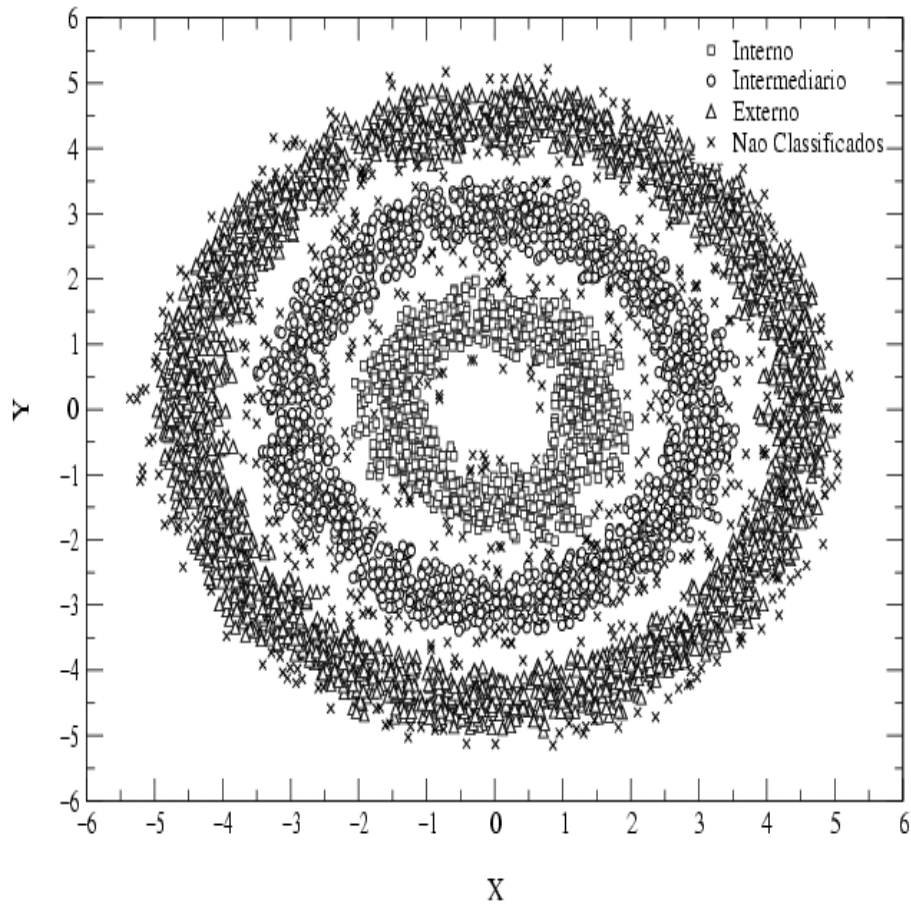
$$G_{ij} = 1 - \frac{2}{q} \quad (2.46)$$

Portanto, calculando a correlação entre os spins podemos identificar se os sítios, associados com esses spins, pertencem a um mesmo aglomerado ou não.

Para tornar mais claro estas propriedades, iremos exemplificá-las com uma distribuição uniforme da componente angular dos pontos e três Gaussianas na distribuição radial, centradas nos raios 1.5, 3.0 e 4.5, com variância 0,87. Foram distribuídos 6000 pontos desta forma, sendo metade deles para a Gaussiana com  $R=4.5$ , um terço para  $R=3.0$  e um sexto para  $R=1.5$  (figura 2.11).

Aplicando essa técnica para a distribuição de pontos, na qual  $a = 1.16$ ,  $\bar{k} = 16$ , sendo o número máximo de vizinhos de um sítio igual a 20, e o número de estados igual a 50, obtemos os gráficos da magnetização e suscetibilidade dados na figura 2.12.

Na fase super-paramagnética calculamos as correlações entre os sítios e, através dela, identificamos os pontos pertencente a cada aglomerado. Classificamos 2778 pontos pertencentes ao anel externo, 1851 pertencente ao anel central, 916 pertencentes ao anel interno e 538 pertencentes a uma série de pequenos grupos (menores que 9 sítios), o que pode ser observado na figura 2.11.



**Figura 2.11:** Distribuição de pontos utilizada para exemplificar a técnica de agrupamento super-paramagnética. Os pontos têm uma distribuição angular uniforme e, uma distribuição radial gaussiana, sendo 1000 pontos distribuidos em torno de  $R = 1.5$ , 2000 em torno de  $R = 3.0$  e 3000 em torno de  $R = 4.5$ . Os símbolos indica a classificação dos grupos na fase super-paramagnética, sendo classificados 2778 pontos como pertencentes ao anel externo, 1851 pertencente ao anel central, 916 pertencentes ao anel interno e 538 pertencentes a uma série de pequenos grupos (menores que 9 sítios).

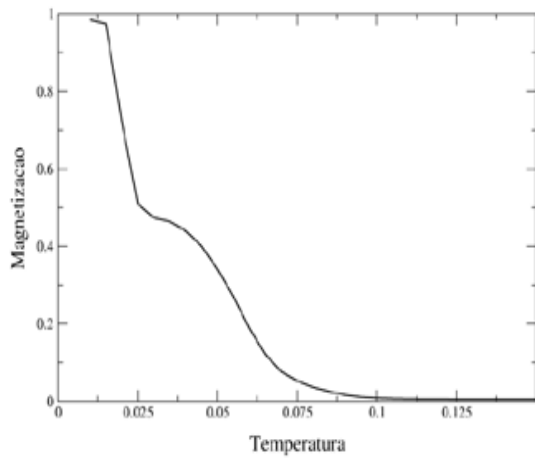


fig.2.12a

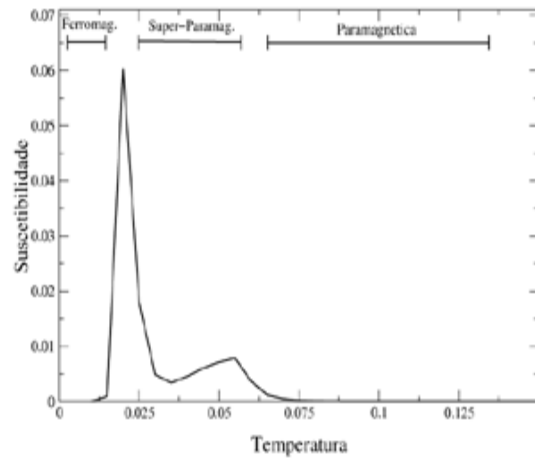


fig.2.12b

**Figura 2.12:** Gráficos da magnetização e suscetibilidade que identificam as três fases do sistema. A fase super-paramagnética é caracterizada por um patamar na suscetibilidade.

# Capítulo 3

## Materiais e Métodos

### 3.1 Experimento

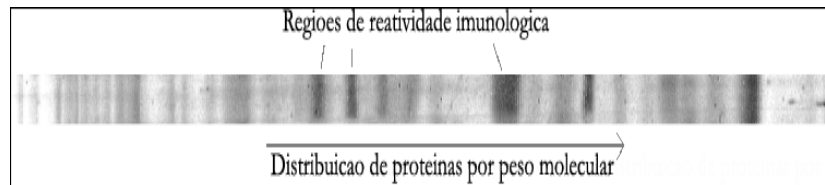
No experimento de obtenção das canaletas, conhecido como *Panama Blot* [33, 34], é necessário que uma população de imunoglobulinas reaja com uma grande quantidade de diferentes proteínas (ou com pedaços delas, os pepitídios). Dentre uma gama de possibilidades, aqui, este extrato de proteínas foi preparado a partir da masseração de fígado ou cérebro humanos. O extrato é então dissolvido em uma solução com Ph adequado para que as proteínas adquiram cargas elétricas<sup>1</sup>.

A solução com as proteínas carregadas é colocada sobre um gel poroso de Poli-acrinamida que está confinado entre duas placas de vidro. O gel é imerso em um campo elétrico constante que faz, devido a suas cargas elétricas, as proteínas se deslocarem para o interior do gel. No entanto, as acelerações adquiridas pelas moléculas dependem das suas massas. Quanto maior a massa, menor a aceleração. Isso faz com que moléculas mais pesadas desloquem-se menos no interior do gel, enquanto as mais leves percorram maiores distâncias. Assim, após desligado o campo elétrico, proteínas com pesos moleculares diferentes estarão em posições diferentes no gel. As mais pesadas estarão mais perto da sua posição inicial, enquanto as mais leves estarão mais distante dela, ocorrendo uma separação entre moléculas com pesos moleculares diferentes.

---

<sup>1</sup>Não será dado enfoque aos processos bio-químicos do experimento, sugerimos ao leitor interessado neste pontos que consulte [1].

Mas, para a reação com o soro, é conveniente se transferir as moléculas imersas no gel para a superfície de uma membrana de nitrocelulose. Para isso, colocamos a membrana sob o gel e aplicamos novamente o campo elétrico até que as proteínas migrem para a nitrocelulose (Figura 3.1).

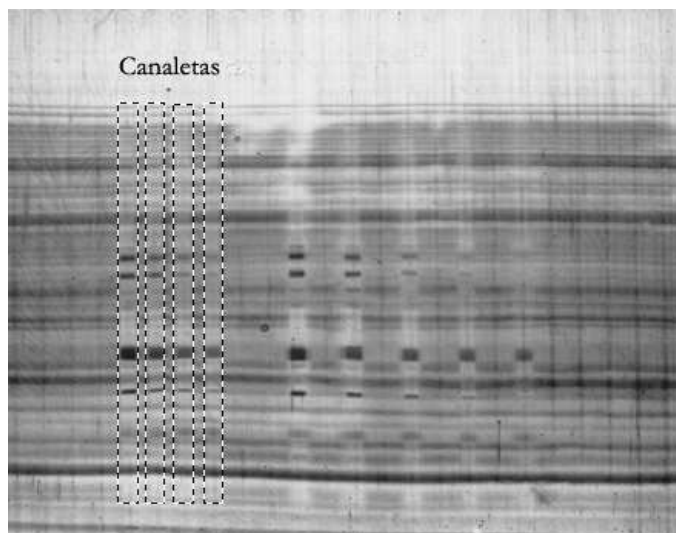


**Figura 3.1:** Imagem das proteínas coradas após o processo de separação e fixação na membrana de nitro-celulose.

A membrana, com as proteínas aderidas à sua superfície, é colocada em uma forma moldada com um série de canaletas paralelas e separadas entre si. Em cada canaleta é colocado o soro de um indivíduo para reagir com as moléculas aderidas à membrana. Após a reação, esta é lavada restando apenas as imunoglobulinas que reagiram com o extrato. Para se observar os locais onde houve a reação e, com qual intensidade, cora-se as imunoglobulinas ligadas à membrana através do extrato. Para isso, é necessário que essas moléculas se liguem com as proteínas que têm uma forma complementar à sua e também contenham uma molécula de fosfatase alcalina ligada a sua estrutura. Após ocorridas essas reações, incuba-se o corante que se liga diretamente à fosfatase alcalina<sup>2</sup>. Agora pode-se visualizar, através do corante, os locais da canaleta onde houve reação e com que intensidade ocorreu. Tradicionalmente, cora-se também todas as proteínas do extrato utilizando *Protogold* (Figura 3.2).

---

<sup>2</sup>Devido a sua funcionalidade na dinâmica do sistema imune, divide-se as imunoglobulinas em várias classes. Pela sua ação na resposta imune contra o parasita da malária, cora-se apenas as imunoglobulinas do tipo M (IgM).



**Figura 3.2:** Membrana resultante do experimento cujas dimensões são 8,5x7,0 cm contendo 24 canaletas. Nela estão coradas as proteínas do extrato (linhas que cruzam todo o perfil) e a imunoglobulinas do tipo M que reagiram (linhas restritas às canaletas).

## 3.2 Tratamentos Digitais

### 3.2.1 Digitalização

Obtidas as membranas temos que digitalizá-las para que possamos tratá-las computacionalmente através de métodos estatísticos. A digitalização é feita por um *scanner* semi-profissional de alta resolução e profundidade. A necessidade de recorrer a este equipamento se deve a distorções provocadas na imagem quando digitalizadas por um *scanner* de mesa comum. Uma das distorções provocadas se deve à resolução do *scanner*, que é a quantidade de pontos por polegada com que ele faz a leitura da imagem. Normalmente, *scanners* não profissionais fazem esta leitura com baixa resolução e adicionam pontos na imagem, já digitalizadas, por algum algoritmo de interpolação, gerando uma imagem de maior resolução. Os pontos adicionados pela interpolação não fazem parte da imagem digitalizada, havendo um corrompimento desta durante o processo de digitalização. Outra distorção provocada se deve à quantidade de tons diferentes de cada cor que o *scanner* consegue distinguir, chamada de profundidade. *Scanners* comuns apresentam uma baixa profundidade, assim, tons diferentes da imagem podem ser lidos como sendo um só tom, provocando uma

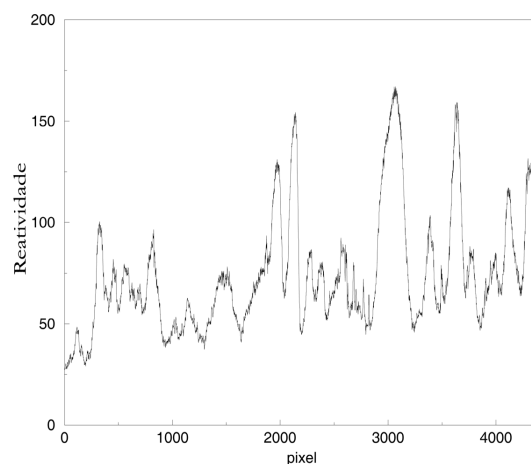
mudança na distribuição de cores. E ainda há indícios de que esta mudança não seja linear, tons claros são deformados de uma forma e tons escuros de outra forma. Para minimizar as distorções inerentes ao processo de digitalização, utilizamos um scanner com resolução real de 2400 pontos por polegadas e com profundidade de 4.6 (numa escala de 0 a 5). Através deste *scanner* as membranas são digitalizadas em tons de cinza e salvas no formato *Post Script* (.eps). Neste formato, a imagem é codificada em um texto que pode ser manipulado por qualquer editor de textos. No cabeçalho do texto existe uma série de instruções que indicam ao programa que lerá a imagem os seus formatos e como deve ser lida. No corpo do texto está a imagem codificada segundo o seguinte algoritmo: cada *pixel* da imagem corresponde a dois caracteres consecutivos no texto, esses caracteres são o tom de cinza (número entre 0 e 255, onde 0 corresponde ao branco e 255 ao preto) daquele *pixel* escrito na base numérica hexadecimal (Figura 3.3).

```
%%Creator: Adobe Photoshop Version 6.0
%%Title: m02_1.eps
%%CreationDate: Thu Oct 31 2002 10:39:36
%%BoundingBox: 0 0 5 142
%%HiResBoundingBox: 0 0 4.681 141.5105
%%SuppressDotGainCompensation
%%DocumentProcessColors: Black
%%EndComments
%%BeginProlog
%%EndProlog
%%BeginSetup
%%EndSetup
%ImageData: 156 4716 8 1 0 156 2 "beginimage"
%BeginPhotoshop: 1652
% 3842494D03ED0A5265736F6C7574696F6E0000000010095F7B7000010002095F
% 7B700001000023842494D040D18465820476C6F62616C204C69676874696E6720
% 416E676C65000000004000000783842494D041912465820476C6F62616C2041
% 6C746974756465000000004000001E3842494D03F30B5072696E7420466C61
% 677300000009000000000000000000001003842494D040ADE436F70797269676874
% 20466C6167000000000100003842494D2710144A6170616E657365205072696E
% 7420466C61677300000000A0001000000000000023842494D03F41C4D6F6E
% 6F6368726F6D652048616C66746F6E652053657474696E677300000000120035
% 00000001002D0000000600000000000013842494D03F71C4D6F6E6F6368726F6D
% 65205472616E736665722053657474696E6773000000001C0000FFFFFFFF
% FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF03E800003842494D0408064775696465
% 730000000100000000100000240000002400000000000000000000000000
% 206F766572726964657300000004000000003842494D041A06536C6963657300
% 00000610000000600000000000000000000000000000000000000000000012
% 6C0000009C0000000000000000000000000000000000000000000000000000
% 0000126C000000000000000000000000000000000000000000000000000000
% 00000000000000000000000000000000000000000000000000000000000000
```

**Figura 3.3:** Parte do arquivo texto no qual a imagem é codificada. A parte superior é o cabeçalho e a inferior o corpo do arquivo.

Conhecida esta regra de codificação, um programa lê o arquivo da imagem e a transforma numa função com o eixo  $x$  corresponde à linha na imagem e o eixo  $y$

ao tom de cinza médio naquela linha (a largura de uma canaleta é de cerca de 100 *pixels*)(Figura 3.4). Este gráfico traduz a reatividade que a proteína apresenta quando exposta às IgMs de um certo indivíduo como função do peso molecular da proteína. Isso retrata, em parte, quais e quantas imunoglobulinas estavam circulando no organismo do doador no momento em que foi feita a coleta do seu sangue.



**Figura 3.4:** Curva obtida da decodificação dos tons de cinza de uma canaleta. No eixo  $y$  o tom de cinza médio da linha correspondente à proteínas de mesmo peso molecular e no eixo  $x$  a posição desta linha na canaleta.

### 3.2.2 Suavização

Toda a análise, que fundamenta este texto, é feita sobre este perfil de reatividade imunológica que, no entanto, deve ser preliminarmente tratado. O perfil obtido diretamente da digitalização, apresenta um ruído que se deve ao experimento pelo qual é obtido e pela grande sensibilidade do *scanner*. Este ruído não retrata nenhum estado de ativação do sistema imune e impossibilita a análise da curva como sendo um fractal, pois o ruído não é invariante por escala, e assim, um gráfico da rugosidade do perfil em função de escala na qual é calculada, não é uma lei de potência. Por isso, é feita uma suavização na curva de modo que seu ruído seja, quase totalmente, retirado. A suavização é feita escolhendo-se um intervalo, de comprimento muito menor que o tamanho do perfil, dentro do qual se faz uma regressão linear da curva,



e atribui-se o valor da melhor reta no ponto médio do intervalo ao valor do perfil suavizado naquele ponto. O intervalo sofre uma pequena translação de modo que somente um ponto saia e um novo ponto entre no intervalo. Novamente é feita a regressão e a translação até que todo o perfil seja percorrido pelo intervalo. O perfil suavizado obtido tem o mesmo número de pontos do original subtraído do número de pontos que cabem dentro do intervalo. Este perfil suavizado apresenta uma lei de potência, com algumas peculiaridades que logo serão estudadas, no gráfico da rugosidade em função da escala. O ruído existente no perfil original pode ser obtido fazendo-se a diferença entre a curva original e a suavizada (Figura 3.5).

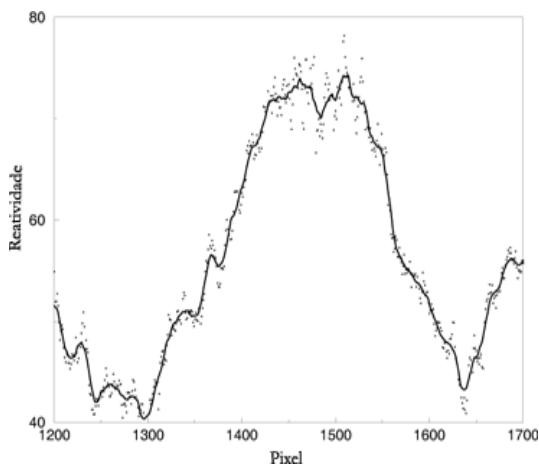


fig.3.5a

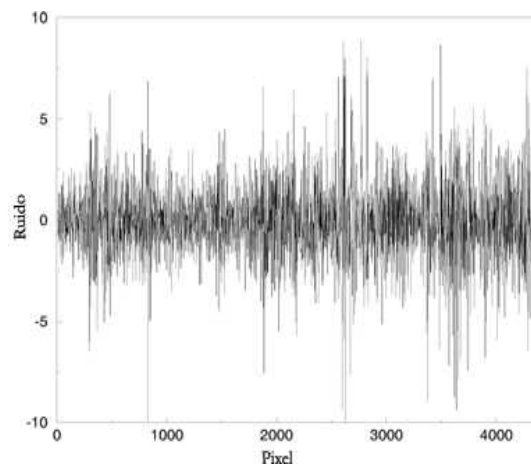
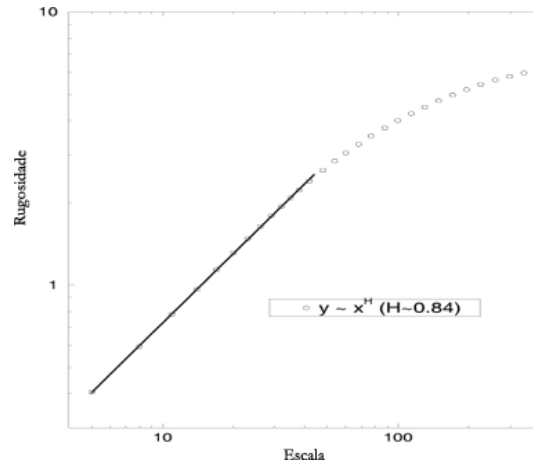


fig.3.5b

**Figura 3.5:** A parte (a) mostra um intervalo do perfil antes da suavização como a curva pontilhada e a curva contínua o mesmo intervalo do perfil após o processo de suavização. E na parte (b) a diferença entre o perfil original e o perfil suavizado. Essa curva é o ruído retirado durante a suavização.

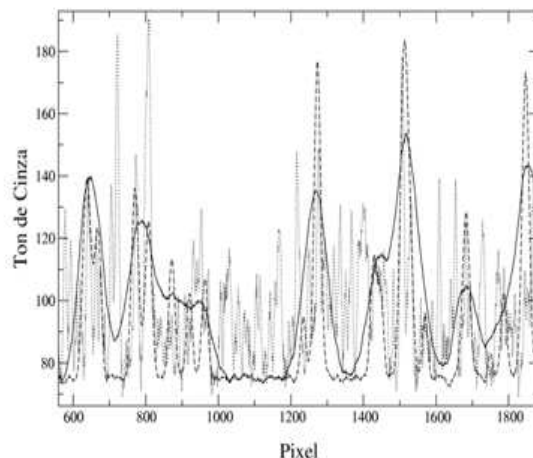
A peculiaridade referida anteriormente se deve ao intervalo de escalas no qual existe a lei de potência. Este intervalo tem comprimento bem menor que o comprimento total do perfil suavizado. Tipicamente, um perfil é formado por cerca de 4500 pontos, enquanto, aproximadamente, o comprimento que essa maior escala alcança é 80 pontos. E ainda, esse intervalo se restringe aos menores valores do comprimento da escala utilizados no cálculo (Figura 3.6). Devido a esse pequeno intervalo de validade da lei de potências, não podemos identificar o perfil como uma curva fractal.



**Figura 3.6:** Gráfico da rugosidade com função da largura da escala. Observa-se que a lei de potência se mantém somente para escalas pequenas.

Para estudar melhor este fenômeno geramos artificialmente uma série de perfis de reatividade. Estes perfis artificiais são uma soma de várias Gaussianas com médias distribuídas aleatoriamente no intervalo de comprimento do perfil, variâncias também distribuídas aleatoriamente dentro de um intervalo e altura de cada Gaussiana proporcional à sua variância. E ainda, para testar o processo de suavização, é somado ao perfil artificial um ruído com distribuição de probabilidades Gaussiana, média nula e variância muito menor que a altura dos picos. A quantidade de Gaussianas, sua variâncias e alturas são ajustadas de modo que reproduzam com grande fidelidade o perfil digitalizado (Figura 3.7). Então, suaviza-se a curva artificial e se compara o histograma do ruído obtido com o ruído que foi adicionado ao gerar o perfil. Com uma escolha coerente do comprimento do intervalo utilizado para a suavização, estes histogramas apresentam apenas pequenas diferenças, indicando que apenas o ruído é retirado no processo de suavização, permanecendo inalterada a informação imunológica, validando, assim, o processo de suavização (Figura 3.7). O gráfico das rugosidades do perfil artificial suavizado apresenta exatamente o mesmo comportamento que o perfil experimental. Alterando o intervalo no qual as variâncias se distribuem e a altura dos picos, geram-se os perfis. O gráfico das rugosidades destes perfis está mostrado na figura 3.9.

Existe uma saturação da rugosidade a partir da escala que corresponde à largura média das Gaussianas, resultado que se repete para centenas de perfis gerados.

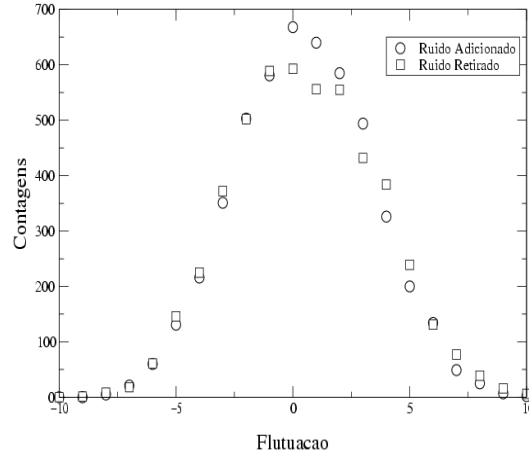


**Figura 3.7:** Três diferentes perfis artificiais, já suavizados, gerados com diferentes variâncias médias das gaussianas.

Assim o expoente de Hurst corresponde à correlação interna média das Gaussianas. A rugosidade de saturação é uma medida da largura do perfil.

### 3.2.3 Normalização

A suavização possibilita a análise via expoente de Hurst. No entanto, para as análises via PCA e agrupamento super-paramagnético é necessário que haja uma normalização dos perfis no eixo  $x$ . Devido a uma série de fatores, a corrida por eletroforese não é homogênea em toda a membrana, provocando discrepâncias entre diferentes canaletas que devem ser corrigidas por esta normalização para que todas as membranas se tornem comparáveis (Figura 3.10a). Para isso escolhemos uma canaleta padrão a partir da qual todas as demais serão normalizadas. Nesta canaleta, o soro que reagiu com o extrato de proteínas é uma mistura dos soros de vários indivíduos utilizados nas outras canaletas. Assim, ela contém, teoricamente, todos os picos de reatividade que podem ser encontrados nas demais canaletas. Um intervalo, de tamanho fixo, é sorteado aleatoriamente no perfil que será normalizado. Este intervalo é deslocado rigidamente, para direita ou para a esquerda, de uma quantidade sorteada aleatoriamente entre  $-d_{max}$  e  $d_{max}$ . Para que não haja sobreposição, nem descontinuidade, entre a parte do perfil que foi deslocada e as partes, vizinhas a essa que não o foram, é necessário que haja uma distorsão (expansão



**Figura 3.8:** Histograma do ruído adicionado ao perfil artificial e histograma do ruído retirado pelo processo de suavização desse mesmo perfil. A semelhança entre as duas curvas valida o processo de suavização.

e compressão) dos intervalos vizinhos ao deslocado. Por exemplo, se o intervalo é deslocado para a direita, o consecutivo a ele é comprimido até que não haja mais sobreposição entre eles; e para que não ocorra uma descontinuidade, o intervalo subsequente é expandido. A área compreendida acima da curva padrão e abaixo da curva sendo normalizada é calculada, no intervalo sorteado, antes e depois do deslocamento. O deslocamento é, então, incorporado à curva com probabilidade

$$P(\delta A) = \frac{1}{1 - e^{-\frac{\delta A}{T}}} \quad (3.1)$$

onde  $\delta A$  é a área após o deslocamento menos a anterior ao deslocamento e  $T$  corresponde à “temperatura do sistema”. Ajustando adequadamente o tamanho do intervalo, o deslocamento máximo e a temperatura, este algoritmo do tipo Metrópolis permite a normalização dos perfis tornando-os comparáveis entre si (Figura 3.10b).

Para a análise via PCA é necessário encontrar a posição dos picos na curva padrão, que, após a normalização, é a mesma posição dos picos nos demais perfis. Definiremos, aqui, um pico da curva como o ponto no qual a inclinação da regressão linear do perfil, num intervalo centrado naquele ponto, tem módulo menor que  $\epsilon_1$ , e a variação da inclinação é menor que  $-\epsilon_2$ , sendo estas duas constantes números positivos. Encontrado os picos da curva padrão, obtemos a altura correspondente a cada pico nos demais perfis como a altura máxima em torno da posição na curva

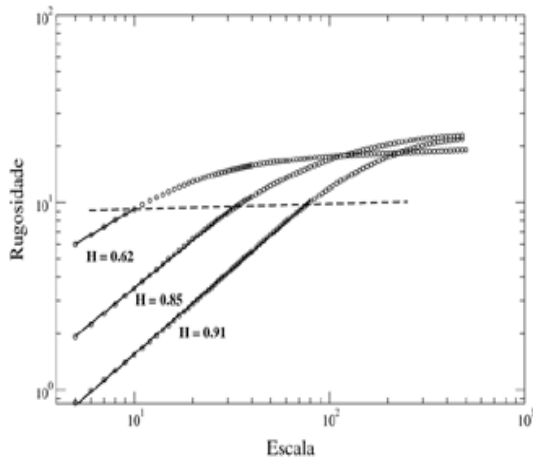


fig.3.9a

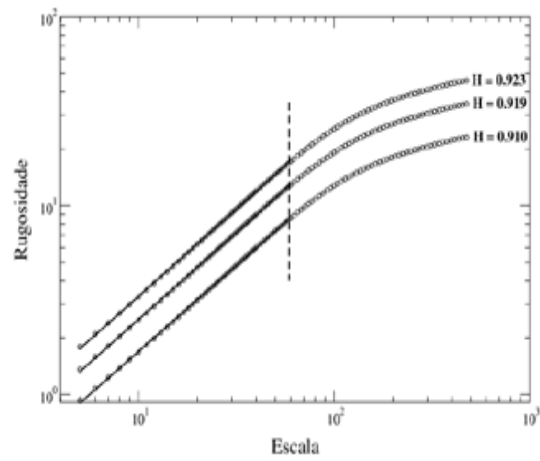


fig.3.9b

**Figura 3.9:** Na parte (a), as rugosidade para os perfis com diferentes variâncias e em (b) as rugosidades para os perfis com diferentes larguras.

padrão daquele pico.

Aplicando a série de tratamentos descritos nesse capítulo podemos analisar os perfis de reatividade imunológica através das técnicas descritas no capítulo 2. Os resultados dessas análises estão mostrados no capítulo seguinte.

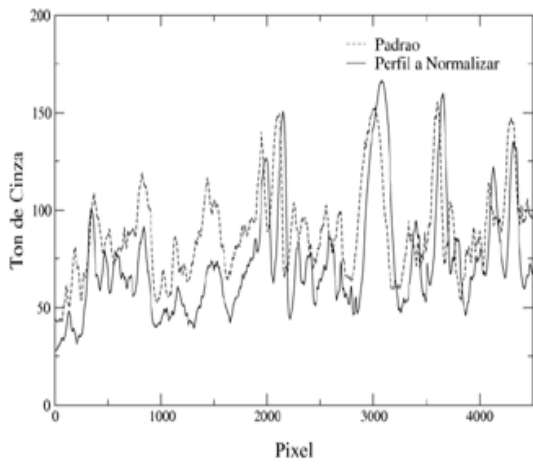


fig.3.10a

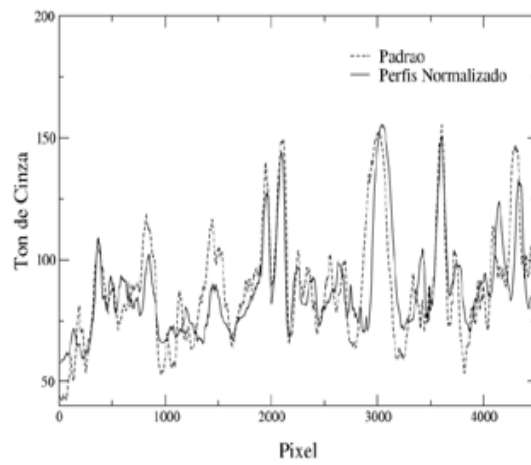


fig.3.10b

**Figura 3.10:** Na parte a, o perfis antes da normalização e em b após a normalização.

# Capítulo 4

## Resultados e Discussões

Os experimentos de *Pamana blot* utilizados neste trabalho para estudar as propriedades do sistema imunológico, foram realizados pelo Laboratório de Imunologia do Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais, que, no momento, parte do interesse se volta para o estudo da endemia da malária. Uma fonte de interesse em estudar essa parasitose é um quadro clínico, peculiar a ela, no qual os indivíduos, ditos assintomáticos, não apresentam nenhum tipo de sintoma da patologia, no entanto estão comprovadamente infectados pelo parasita, ou seja, o contém no sangue.

Para realizar os experimentos, foi colhido o sangue de uma série de indivíduos, com hábitos de vida muito parecidos entre si, residentes na área endêmica de malária no Mato Grosso. Devido a seu quadro clínico, os indivíduos podem ser naturalmente separados nos seguintes grupos:

- Os que, apesar de residir na região endêmica e, por isso, serem expostos ao parasita, nunca comprovadamente o apresentaram no sangue;
- Os que comprovadamente já foram infectados pelo parasita entre 1 e 10 vezes;
- Os que comprovadamente já foram infectados mais de 10 vezes;
- Os sintomáticos, que, no momento da coleta do sangue, apresentavam os sintomas da doença;
- Os assintomáticos, que, no momento da coleta, comprovadamente apresentavam o parasita no sangue, no entanto, não apresentavam os sintomas da doença.

Fora da região endêmica, em Belo Horizonte, também foram feitas coletas de soro sanguíneo. Neste caso, os indivíduos se enquadram no grupo:

- Os que nunca foram expostos ao parasita, por residir numa região não endêmica.

Durante o experimento, o soro de cada indivíduo é colocado em uma das canaletas para reagir com o extrato de proteínas, imobilizadas na membrana. Dois diferentes extratos foram utilizados, um obtido a partir da maceração de fígado humano e outro a partir da maceração de cérebro humano. O extrato é, então, uma mistura de um enorme número de proteínas diferentes em concentrações diferentes. Não há um controle sobre quais proteínas específicas integram o extrato. Assim, não existe uma relação direta entre as proteínas que compõe o parasita da malária e as que compõe o extrato.

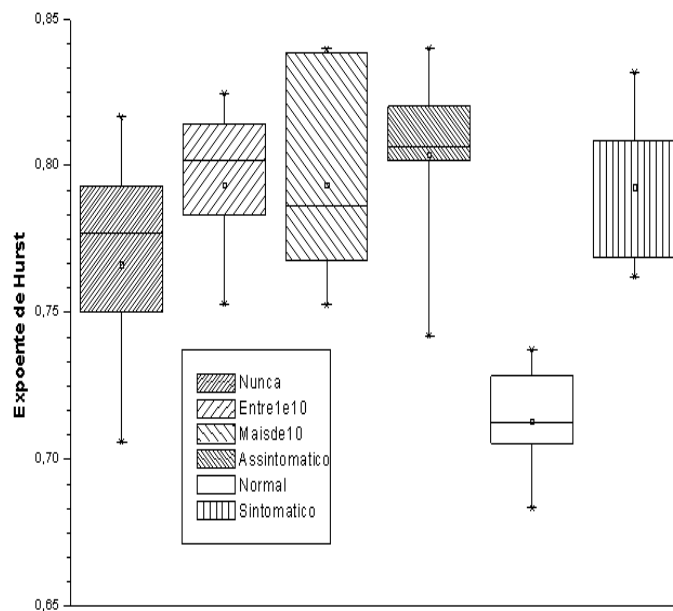
Em cada membrana, além dos soros individuais, incuba-se também, em uma canaleta, um soro formado pela mistura de todos soros individuais. Esta canaleta, que chamaremos padrão, é necessária para o processo de normalização (conforme secção 3.2.3).

Após incubados os soros, revelam-se as imunoglobulinas do tipo M que reagiram e todas proteínas do extrato (conforme seção 3.1). A membrana é digitalizada e separa-se as canaletas de cada indivíduo. Cada canaleta é transformada em um gráfico entre a reatividade que a proteína apresentou quando exposta ao soro daquele indivíduo e seu peso molecular (conforme seção 3.2.1). Este perfil de reatividade é como se fosse uma fotografia do estado de ativação do sistema imunológico no momento da coleta do sangue.

## 4.1 Expoente de Hurst

Cada perfil é suavizado e pelo cálculo da rugosidade em várias escalas (conforme seção 2.1) obtemos o expoente de Hurst de cada um deles, e os agrupamos segundo o quadro clínico do indivíduo correspondente àquele expoente. Dispomos os expoentes de Hurst, separados pelos grupos, no gráfico mostrado na figura 4.1. Para gerar os perfis de reatividade utilizados para obter este gráfico utilizou-se um extrato obtido de fígado humano.

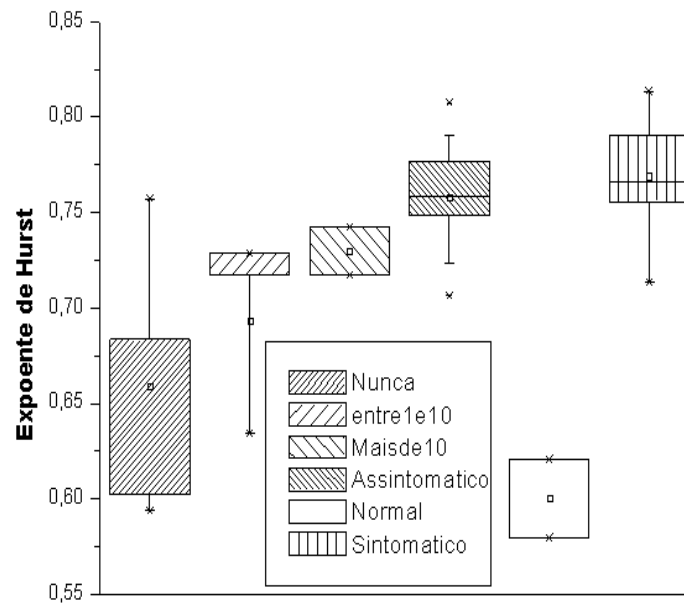




**Figura 4.1:** Expoente de Hurst dos perfis obtidos utilizando extrato de fígado, separados pelo quadro clínico da malária em que se encontravam os indivíduos no momento da coleta do sangue. A barra horizontal indica o maior e menor valor do expoente e a caixa o intervalo onde estão 75% dos dados. Esses se constituem de 7 indivíduos no grupo dos que residem na região endêmica mas nunca apresentaram os sintomas da malária, 8 no grupo dos que apresentaram entre 1 e 10 vezes os sintomas, 8 no grupo dos que apresentaram mais de 10 vezes os sintomas, 9 no grupo dos assintomáticos, 7 no grupo dos que não residem na região endêmica, e 8 no grupo dos sintomáticos. Nota-se que o intervalo em que se encontram os expoentes do grupo do que não residem na região endêmica é diferente dos demais grupos.

Para outra série de pessoas, utilizou-se um extrato obtido de cérebro humano para incubar os soros. Os expoentes de Hurst dos perfis de reatividade destes indivíduos estão dispostos, já separados por grupos, na figura 4.2.

O intervalo no qual encontramos os expoentes correspondentes aos indivíduos que nunca foram expostos ao parasita da malária é diferente do intervalo no qual encontramos os expoentes dos demais grupos, sendo que, em todos eles, os indivíduos já foram expostos ao parasita. Esta diferença no expoente implica em uma diferença no comportamento do perfil de reatividade que implica em uma diferença na localização e intensidade das reações na canaleta. A exposição ao parasita, de alguma forma, modifica as reatividades do sistema imunológico em relação às proteínas presentes



**Figura 4.2:** Expoente de Hurst dos perfis obtidos utilizando extrato de cérebro, separados pelo quadro clínico da malária em que se encontravam os indivíduos no momento da coleta do sangue. Os dados contituem de 4 indivíduos no grupo dos que residem na região endêmica mas nunca apresentaram os sintomas da malária, 5 no grupo dos que apresentaram entre 1 e 10 vezes os sintomas, 5 no grupo dos que apresentaram mais de 10 vezes os sintomas, 12 no grupo dos assintomáticos, 3 no grupo dos que não residem na região endêmica, e 10 no grupo dos sintomáticos. Nota-se que o intervalo em que se encontra os expoentes do grupo do que não residem na região endêmica é diferente dos demais grupos..

no extrato.

## 4.2 Análise de Variância

Para verificar estatisticamente a diferença no comportamento dos grupos, aplicamos uma análise de variância aos dados do expoente de Hurst. Esta análise consiste em testar a hipótese de igualdade entre as médias de cada grupo. Dado que a dispersão dos expoentes de cada grupo não é muito diferente, se o expoente médio dos grupos forem muito próximos, não é possível fazer uma distinção entre eles; no entanto, se houver uma grande diferença na média, podemos fazer a distinção. Testaremos essa hipótese, calculando uma quantidade, que será definida posteriormente, utilizando

os expoentes médios reais de cada grupo, e calculando também a mesma quantidade, supondo, agora, que os expoentes médios dos grupos são os mesmos.

A quantidade calculada, utilizando os expoentes médios reais, denominada soma dos quadrados dentro [35], é definida como

$$SQD = \sum_{i=1}^k \sum_{j=1}^m (Y_{ij} - \bar{Y}_i)^2 = \sum_{i=1}^k \sum_{j=1}^m (Y_{ij})^2 - m \sum_{i=1}^k \bar{Y}_i^2, \quad (4.1)$$

sendo a variável aleatória  $Y_{ij}$  o valor do expoente de Hurst correspondente ao  $j$ -ésimo indivíduo do  $i$ -ésimo grupo,  $\bar{Y}_i$  o expoente médio do  $i$ -ésimo grupo,  $k$  o número de grupos e  $m$  o número de indivíduos por grupos que, por simplicidade, consideramos, aqui, o mesmo para todos os grupos.

Supondo que o expoente médio é o mesmo para todos os grupos, o calculamos como  $\bar{Y} \equiv \frac{1}{mk} \sum_{i=1}^k \sum_{j=1}^m Y_{ij}$ , e neste caso a quantidade utilizada no teste é a soma de quadrados total

$$SQT = \sum_{i=1}^k \sum_{j=1}^m (Y_{ij} - \bar{Y})^2 = \sum_{i=1}^k \sum_{j=1}^m (Y_{ij})^2 - mk\bar{Y}^2. \quad (4.2)$$

A diferença entre estas duas grandezas define a soma de quadrado entre

$$SQE = SQT - SQD \quad (4.3)$$

Utilizando as expressões acima, obtemos

$$SQE = m \left( \sum_{i=1}^k \bar{Y}_i^2 - k\bar{Y}^2 \right). \quad (4.4)$$

Cada uma das somas de quadrados envolve um certo número de quantidades, levando isso em consideração, definimos os correspondentes quadrados médios

$$QMT = \frac{SQT}{km - 1}; \quad (4.5)$$

$$QMD = \frac{SQD}{k(m - 1)}; \quad (4.6)$$

$$QME = \frac{SQE}{k - 1}. \quad (4.7)$$

Quando  $QME$  for grande comparado à  $QMD$  a hipótese de igualdade entre as médias deve se rejeitada. Assim definimos

$$f = \frac{QME}{QMD}. \quad (4.8)$$

Quanto maior o valor de  $f$ , maiores as evidências contra a hipótese. Temos, então que definir um valor crítico  $f_\alpha$  a partir do qual rejeitamos a hipótese. Supondo que, para cada grupo, a distribuição de probabilidades para  $Y_{ij}$  é gaussiana e com mesma variância, pode-se mostrar que a quantidade  $f$  tem uma distribuição de Fisher-Snedecor [36]

$$F_{u,v}(x) = \frac{\Gamma(\frac{u+v}{2})(\frac{u}{v})^{u/2}x^{(u/2)-1}}{\Gamma(\frac{u}{2})\Gamma(\frac{v}{2})[(\frac{u}{v})x + 1]^{(u+v)/2}}, \quad 0 < x < \infty \quad (4.9)$$

sendo

$$\Gamma(r) = \int_0^\infty x^{r-1}e^{-x}dx, \quad \text{para } r > 0 \quad (4.10)$$

a função Gama e os parâmetros  $u$  e  $v$ , neste caso, iguais à  $k - 1$  e  $k(m - 1)$  respectivamente.

Definimos  $f_\alpha$  como

$$P(F > f_\alpha) = \int_{f_\alpha}^\infty f(x)dx = \alpha \quad (4.11)$$

Assim, a quantidade, pré-determinada,  $\alpha$ , conhecida como nível de significância, é a probabilidade de  $F$  estar acima do valor crítico  $f_c$ .

Para o conjunto de expoentes obtidos utilizando tanto o extrato de fígado quanto o extrato de cérebro, aplicamos a análise de variância, utilizando o programa *Origin*. Em ambos os casos, obtivemos que, com nível de significância de 0,01, a média do grupo dos indivíduos que nunca foram expostos ao parasita é distinta das médias dos demais grupos, confirmando, estatisticamente, a diferença no comportamento entre esses grupos. No experimento com extrato de cérebro, não há distinção entre as médias dos grupos dos que nunca foram expostos e dos que moram na região endêmica mas nunca apresentaram os sintomas da doença.

### 4.3 Análise por PCA

Para que possamos aplicar as duas análises que ainda restam, temos que representar cada perfil de reatividade como um ponto em um espaço  $d$ -dimensional. Uma forma de se fazer esta representação é simplificando o perfil de reatividade. Para isso, encontramos os picos (denificados na secção 3.2.3) que compõem o perfil e calculamos a altura média de cada um deles. Dispondo esses valores na ordem em que os picos aparecem no perfil, obtemos uma versão simplificada dele, pois, ao invés de escrevermos todos os seus pontos, escrevemos apenas a altura média de cada um de seus picos. Com esta simplificação podemos representar cada perfil como um ponto no espaço onde cada eixo, que forma sua base, corresponde à altura média de um dos picos. Assim, o número de dimensões deste espaço é o número de picos que o perfil apresenta. Na prática, o número de picos para os experimentos com extrato de fígado foi 19 e com extrato de cérebro 15.

Para distribuímos todos os perfis neste espaço, primeiro temos que normalizá-los (conforme secção 3.2.3) de modo que a posição de um dado pico seja a mesma em todos perfis, incluindo o perfil padrão. Feita a normalização, encontramos a posição de cada um dos picos da curva padrão, que, teoricamente, contém todos picos que podem aparecer nos demais perfis. E, para cada pico dessa curva, encontramos, nos demais perfis, os picos que têm a mesma posição que aquele da curva padrão, e calculamos as suas alturas médias. Distribuímos, então, os perfis no espaço onde o primeiro eixo corresponde à altura do pico com a mesma posição do primeiro pico da curva padrão. E a mesma idéia para os demais eixos.

De posse dessa distribuição de pontos, utilizando o programa Maple, calculamos a sua matriz de covariâncias (conforme secção 2.2), e projetamos os pontos nos auto-vetores correspondentes ao maior e terceiro maior auto-valores, sendo que a projeção no segundo eixo principal não forneceu resultado satisfatório. Para os experimentos utilizando extrato de fígado obtemos o gráfico da figura 4.3a. E para os experimentos utilizando extrato de cérebro obtemos o gráfico da figura 4.3b.

Para obter esses gráficos, é necessário fazer uma segunda normalização dos perfis. Isso se deve ao fato de que, para um mesmo extrato, o experimento foi feito em diferentes membranas, pois o número de soros que foram incubados é maior que o

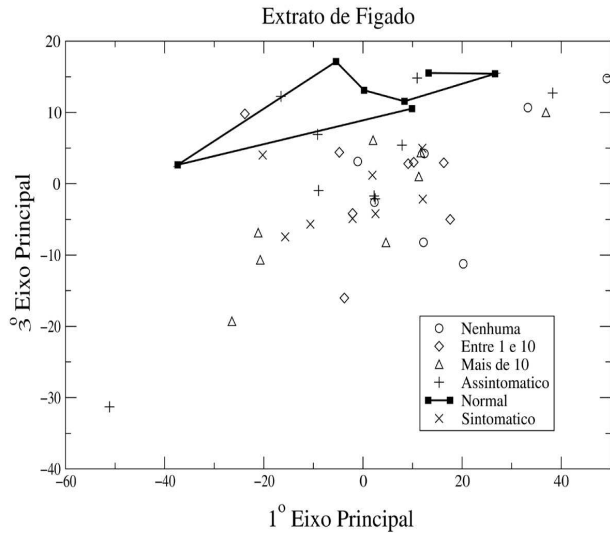


fig.4.3a

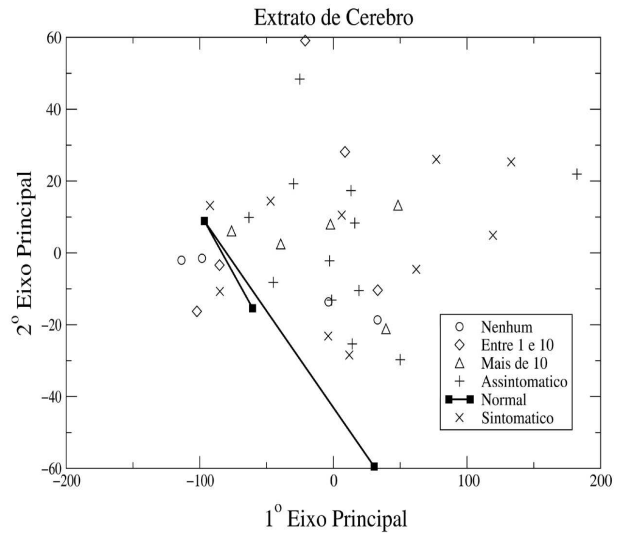


fig.4.3b

**Figura 4.3:** Na parte (a), a projeção dos pontos correspondentes aos perfis obtidos através do extrato de fígado nos 1º e 3º eixos principais. Nota-se a tendência dos pontos correspondentes ao grupo dos que não residem na região endêmica a se concentrarem na parte superior, enquanto os pontos dos demais grupos de distribuem por todo o gráfico. Na parte (b), a projeção dos pontos correspondentes aos perfis obtidos através do extrato de cérebro nos 1º e 2º eixos principais. Nota-se a tendência dos pontos correspondentes ao grupo dos que não residem na região endêmica a se concentrarem na parte inferior esquerda, enquanto os pontos dos demais grupos se distribuem por todo o gráfico.

número de canaletas de uma membrana. Devido a diferentes condições experimentais, existe uma diferença na coloração dessas membranas, e também, há um ajuste, inerente ao processo de digitalização, do brilho e contraste da imagem que, devido a coloração, é diferente para cada membrana. Para que possamos comparar os perfis provenientes de diferentes membranas, temos que eliminar essas discrepâncias. Para isso, deslocamos os pontos, que representam os perfis no espaço, de modo que a média e a dispersão do grupo de todos pontos que pertencem a uma membrana, sejam as mesmas para todas as membranas.

Observamos, novamente, a mesma tendência de distinção do comportamento do grupo dos que nunca foram expostos ao parasita dos demais grupos.

## 4.4 Análise por Agrupamento Super-paramagnético

Finalizando o trabalho, aplicamos a técnica de agrupamento super-paramagnético, à distribuição de pontos, no espaço  $d$ -dimensional, onde cada um deles corresponde a um perfil simplificado. Podemos visualizar a projeção desses pontos num sub-espaço bi-dimensional através da projeção nos componentes principais, dada pelos gráficos 4.3.

Para a distribuição obtida utilizando extrato de fígado, definimos o número de estados como  $q = 50$  e o número máximo de vizinhos de um sítio como  $K = 10$ . O gráfico da suscetibilidade 4.4a mostra a não existência da fase super-paramagnética, pois não há um patamar na suscetibilidade. O sistema passa da fase ferromagnética diretamente para a paramagnética, visto que há somente um pico nesse gráfico. Portanto não se formam grupos de sítios próximos, o que pode ser observado no gráfico 4.4b. Essa não formação de grupos se deve ao fato da distribuição de pontos ser bastante homogênea, não havendo grupos de diferentes densidades.

Para a distribuição obtida com extrato de cérebro, utilizamos novamente  $q = 50$  e  $K = 10$ . O gráfico da suscetibilidade 4.5a mostra, agora, a existência da fase super-paramagnética, pois há um patamar na suscetibilidade. Nessa fase, formam-se grupos correspondentes às regiões de diferentes densidades e o tamanho desses grupos pode ser visto no gráfico 4.5b.

Na temperatura 0.095, os grupos identificados estão mostrados no gráfico 4.6. Estes grupos são, aproximadamente, aqueles formados pela separação que existe na distribuição de pontos, entre o grupo do lado esquerdo e o grupo do lado direito. No entanto, esses grupos não correspondem aos grupos do quadro clínico da malária. Há suspeitas, baseadas em resultados não publicados do Laboratório de Imunologia de Universidade de Minas Gerais, de que esses grupos correspondam aos grupos dos indivíduos que foram infectados ou não pela hepatite B. A possibilidade de identificação desses grupos aqui, e a não identificação no experimento com extrato de fígado, se deve, justamente, aos diferentes extratos utilizados. A malária modifica a reatividade do sistema imune com relação a um certo conjunto de proteínas, e uma outra patologia modifica com relação a um outro conjunto. Assim, utilizando

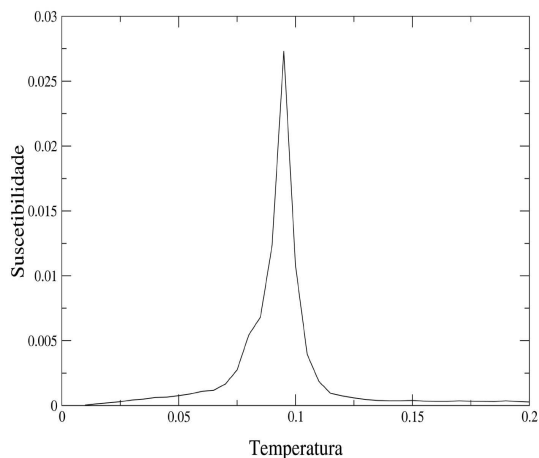


fig.4.4a

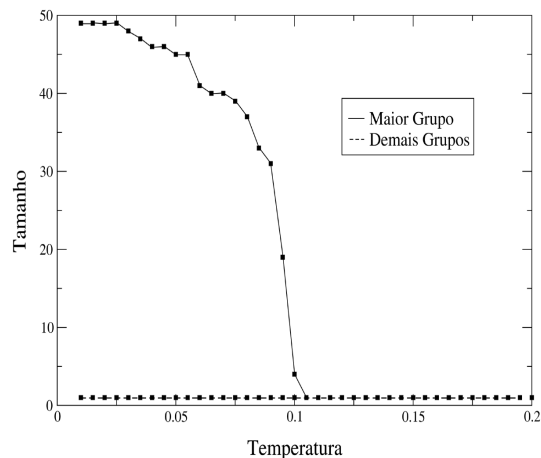


fig.4.4b

**Figura 4.4:** A parte (a) mostra a suscetibilidade obtida aplicando essa técnica. Observa-se apenas um pico, que corresponde a uma transição de segunda ordem da fase ferromagnética para a fase paramagnética, indicando a não existência da fase superparamagnética. Na parte (b), os tamanhos dos dois maiores agrupamentos identificados, o maior deles, durante a transição de fase, passa de um grupo, praticamente do tamanho da rede, para um grande número de grupos muito pequenos.

proteínas diferentes, ou seja, extratos diferentes, podemos identificar quadros clínicos distintos correspondentes a patologias distintas.



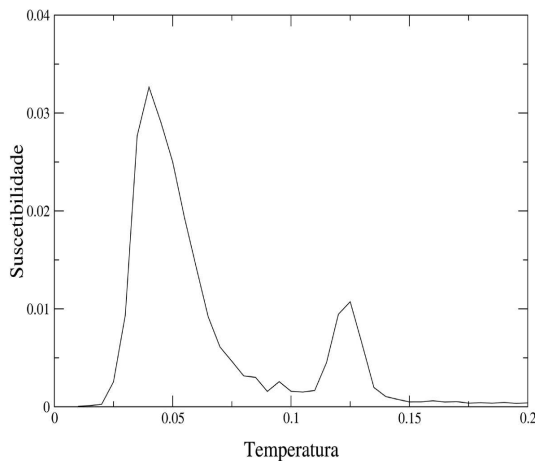


fig.4.5a

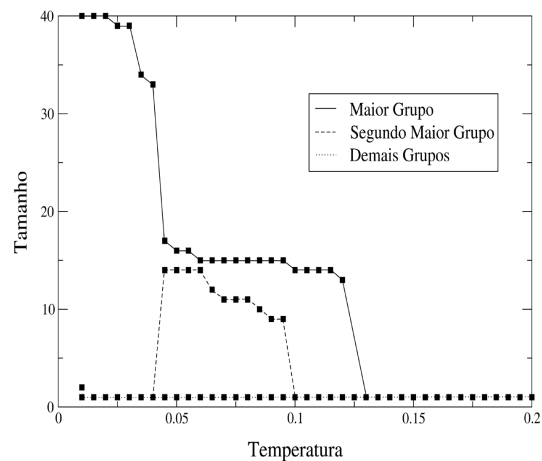
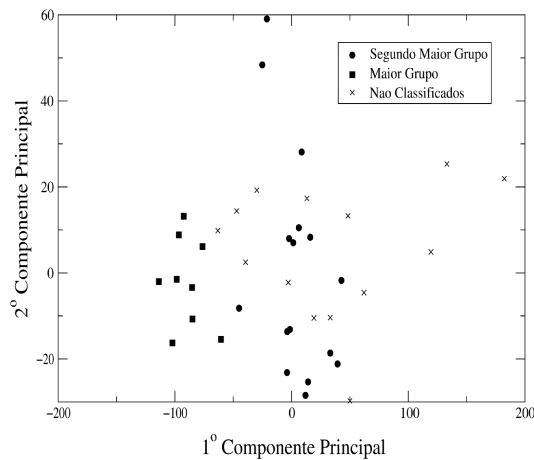


fig.4.5b

**Figura 4.5:** A parte (a) mostra a suscetibilidade obtida aplicando essa técnica. Observa-se o pico seguido de um patamar, o que corresponde à transição fase ferromagnética para a fase super-paramagnética. O segundo pico indica a transição, interna a um dos grupos formados na fase super-paramagnética, para a fase paramagnética. Na parte (b), os tamanhos dos três maiores agrupamentos identificados, o maior deles, durante a transição de fase, se divide de dois grupos menores de tamanhos, aproximadamente, 16 e 9, que persistem durante a fase super-paramagnética.



**Figura 4.6:** Projeção dos pontos correspondentes aos perfis obtidos através do extrato de cérebro nos 1º e 2º eixos principais, separados no grupos identificados pela técnica de agrupamento super-paramagnético. Essa separação corresponde, aproximadamente, aos grupos do lado esquerdo e do lado direito.

# Capítulo 5

## Conclusões

A introdução da interação entre os linfócitos possibilita o estudo do sistema imunológico como um sistema complexo. O comportamento desse tipo de sistema difere do comportamento de seus componentes quando isolados uns dos outros. Assim, podemos obter novos resultados quando analisamos experimentos que possibilitem a interação inter-lifocitária, como não é do caso da medida da reatividade de cada imunoglobulina, o que tradicionalmente é feito na imunologia. Um experimento que possibilita as interações, medindo a reatividade de uma população de proteínas quando reage com soro humano, é o *Panama blot*, cujo resultado pode ser digitalizado e transformado em um gráfico e devidamente tratado, sendo o perfil de reatividade.

Apesar de não podermos dizer que esse perfil é uma curva fractal, podemos calcular seu expoente de Hurst para pequenas escalas. O intervalo no qual se encontraram os expoentes correspondentes aos indivíduos que residem numa região não endêmica de malária e, por isso, nunca foram expostos ao parasita, é diferente do intervalo em que se encontram os expoentes dos indivíduos que, de alguma forma, foram expostos ao parasita. A exposição ao parasita modifica a reatividade do sistema imunológico com relação a uma série de proteínas que não estão diretamente relacionadas com as encontradas no parasita, modificação esta, conseqüente da forma de interação entre os linfócitos.

Transformando cada perfil em um ponto num certo espaço, podemos visualizá-los, adequadamente, através da análise de componentes principais. Projetando os pontos

nas componentes principais com maiores variâncias, os perfis correspondentes aos indivíduos que moram numa região não endêmica tendem a se concentrar na lateral do gráfico, numa região pequena e bem definida, enquanto os demais perfis se distribuem por todo o gráfico, novamente, refletindo a diferença na resposta imunológica dos grupos.

Aplicando uma análise de agrupamento super-paramagnético, não há a formação de grupos nos perfis obtidos utilizando extrato de fígado, o que se deve à homogeneidade da distribuição de pontos. Para os perfis obtidos com extrato de cérebro, há a formação, basicamente, de dois grupos que, no entanto, não correspondem aos grupos do quadro clínico da malária, podendo eles corresponder a grupos de diferentes quadros clínicos de outra patologia.

# Bibliografia

- [1] B. Stransky, *Tese de Doutorado* Instituto de Microbiologia, Universidade Federal do Rio de Janeiro (2003).
- [2] L. N. Castro, *Introduction to the Artificial Immune System* (ICANNGA, 2001).
- [3] A. S. Perelson, G. Weisbuch, *Reviews of Modern Physics* **64**, 1219 (1997).
- [4] A. M. C. Faria, N. M. Vaz, *Guia Incompleto de Imunobiologia* (Coopmed, Universidade Federal de Minas Gerais, 1993).
- [5] N. K. Jerne, *Annual Immunology* **125C**, 373 (1974).
- [6] R. M. Z. dos Santos, A. T. Bernardes, *Physical Review Letters* **81**, 3034 (1998).
- [7] M. Plischke, B. Bergersen, *Equilibrium Statistical Physics* (World Scientific, Singapore, 1994).
- [8] P. Bak, *How Nature Works* (Corpenicus, New York, 1996).
- [9] S. Wolfram, *Reviews of Modern Physics* **55**, 601 (1984).
- [10] S. Wolfram, *Theory and Applications of Cellular Automata* (World Scientific, Singapore, 1986).
- [11] R. Dickman, M. A. Muñoz, A. Vespignani, S. Zapperi, *Brazilian Journal of Physics* **30**, 27 (2000).
- [12] R. Frigg, *Studies in History and Philosophy of Science* **61**, 5678 (2000).
- [13] A. T. Bernardes, R. M. Z. dos Santos, *Journal of Theoretical Biology* **186**, 173 (1997).

- [14] A. T. Bernardes, R. M. Z. dos Santos, *International Journal of Modern Physics C* **12**, 1 (2001).
- [15] R. M. Z. dos Santos, *International Journal of Modern Physics C* **9**, 193 (1998).
- [16] M. E. R. Dotto, *Tese de Doutorado* Instituto de Física, Universidade Federal Campinas (2003).
- [17] A. V. C. Castro-e-Silva, *Tese de Mestrado* Departamento de Física, Universidade Federal de Minas Gerais (1995).
- [18] T. Vicsek, *Fractal Growth Phenomena* (World Scientific, Singapore, 1992).
- [19] H. M. Nussenzveig, *Complexidade & Caos* (Copea, Universidade Federal do Rio de Janeiro, 1999).
- [20] J. G. Moreira, J. K. L. Silva, S. O. Kamphorst, *Journal of Physics A* **27**, 8079 (1994).
- [21] K. Ivanova, M. Auloo, *Physica A* **274**, 349 (1999).
- [22] R. B. Govindan, D. Vjushin, S. Brenner, A. Bunde, S. Havlin, H. Schellnhuber, *Physica A* **294**, 239 (2001).
- [23] R. L. Costa, G. L. Vasconcelos, *Physica A* **329**, 231 (2003).
- [24] J. A. Sales, M. L. Martins, J. G. Moreira, *Physica A* **245**, 461 (1997).
- [25] W. R. Dillon, M. Goldstein, *Multivariate Analysis: methods and applications* (John Wiley, New York, 1984)
- [26] M. Blatt, S. Wiseman, E. Domany, *Neural Computacion* **9** **8**, 1805 (1997).
- [27] M. Blatt, S. Wiseman, E. Domany, *Physical Review Letters* **76**, 3251 (1996).
- [28] M. Blatt, S. Wiseman, E. Domany, *Physical Review E* **57**, 3767 (1998).
- [29] S. R. A. Salinas, *Introdução à física estatística* (EdUsp, Universidade de São Paulo, 1997).

- [30] N. Metropolis, S. Ulam, *Journal of the American Statistical Association* **44**, 335 (1949).
- [31] R. H. Swendsen, S. Wang, A. M. Ferrenberg, *The Monte Carlo Method in Condensed Matter Physics* (Springer-Verlag, Berlin, 1992).
- [32] J. Hoshen, R. Kopelman, *Physical Review B* **14**, 3438 (1976).
- [33] A. Nobrega, M. Haury, A. Gradien, E. Malanchere, A. Sundblad, A. Coutinho, *European Journal of Immunology* **23**, 2851 (1993).
- [34] M. Haury, A. Gradien, A. Sundblad, A. Coutinho, A. Nobrega, *Scandinavian Journal of Immunology* **39**, 79 (1994).
- [35] M. N. Magalhães, A. C. P. Lima, *Noções de Probabilidade e Estatística* (EdUsp, Universidade de São Paulo, 2002).
- [36] D. C. Montgomery, G. C. Runger, *Applied Statistics and Probability for Engineers* (John Wiley & Sons, New York, 1999).