



UNIVERSIDADE ESTADUAL PAULISTA

Instituto de Biociências, Letras e Ciências Exatas

DEPARTAMENTO DE FÍSICA

**Estudo da dinâmica do modelo de
Peyrard-Bishop não homogêneo
para o DNA**

Joaquim Manoel da Silva

Dissertação de Mestrado
Pós-Graduação em Biofísica Molecular

Rua Cristovão Colombo, 2265

Caixa Postal 136

15054-000 - São José do Rio Preto - SP - Brasil

Telefone: (017) 3221-2200

Fax: (017) 3221-2200

**Estudo da dinâmica do modelo de Peyrard-Bishop não
homogêneo para o DNA**

Joaquim Manoel da Silva

Dissertação apresentada ao Instituto de Biociências, Letras e Ciências Exatas da
Universidade Estadual Paulista "Júlio de Mesquita Filho", Câmpus de São José
do Rio Preto, São Paulo, como parte exigida para a obtenção do título de Mestre
em Biofísica Molecular.

Orientador: Prof. Dr. José Roberto Ruggiero

Co-orientador: Prof. Dr. Elso Drigo Filho

São José do Rio Preto

Fevereiro 2006

Dedicado a toda minha família e em especial
a minha mãe Marina e minha namorada
Karina.

"O curso do destino de um guerreiro é inalterável. O desafio é quão longe ele pode ir dentro destes limites rígidos, o quão impecável ele pode ser dentro destes limites. Se há obstáculos no seu caminho, o guerreiro luta impecavelmente para ultrapassá-los. Se acha dificuldades e dores insuportáveis no seu caminho ele chora, mas todas as suas lágrimas juntas não movem a linha do destino nem um milímetro".

C. Castanheda em "O presente da Águia"

Agradecimentos

Início agradecendo aos meus amigos e amigas da moradia estudantil da Unesp São José do Rio Preto: Serginho (da cadeira), Alfredo, Azul, Cristiane, Fernanda Gonçalves, Marcela, Junior Leal, Marcos de Jesus, Chico, Carla, Ademar, Ricardo, Regis, Minduim, em especial a um irmão que encontrei na graduação, João Gabriel, amigo de todas as horas e ao meu grande amigo Marcus Augusto, sempre participando nos momentos de estudo e lazer. Meus amigos e amigas de graduação, João César, Cristian, Eduardo, Fernanda Tomé, Lucelene, Elen, Cátia, Júlio, Tatiane, as outras Fernandas, Gi, minha grande amiga Karolina Barone. Esses dentre outros amigos e amigas tornaram meu aprendizado mais agradável e proveitoso.

Gostaria também de agradecer aos meus professores de graduação que me iniciaram nesta jornada: Toninho, Clotilzio, Cláudio, Jefferson, Walter Júlio, Gorete, Ermínia, Chahine, Johnny, Baida, Elso, Ramon, Masayoshi, Borges e Maurílio, vocês foram fundamentais na minha formação e responsáveis para eu chegar até aqui.

No departamento de Física no campus da UNESP de São José do Rio Preto, fiz muitos amigos (Carlos, Ésio, Sidney, Denis, Sabrina, Márcia, Glaucia, Fausto, Priscila, André, Diogo, Leandro, Luciana, Ricardo, Patrícia, Henrique, Nelson, Ana Helena, Helen, entre outros). Gostaria de dizer a todos “MUITO OBRIGADO”. As discussões de física com o Carlos e o Ésio ajudaram muito a entender

um pouco mais de física. O Carlos ajudou muito com a parte computacional, pois com a sua ajuda comprei o computador e tudo ficou funcionando maravilhosamente bem. Também ajudou muito com o Linux. A Sabrina e a Márcia foram minhas companheiras de sala, sempre foram extremamente prestativas e me ajudaram sempre a ver as coisas com mais clareza. Na etapa final passei a dividir sala com André, Ricardo e Carlos, uma convivência muito agradável e divertida.

Este trabalho não teria sido possível sem o ensinamento dos meus mestres da pós-graduação: João, Masayoshi, Chahine e Márcio: obrigado pela paciência e dedicação que sempre demonstraram.

Gostaria de agradecer também a todos os docentes e funcionários (Marcelo, Chahine, Elso, Eloi, Ramon, Paulinho, Barbosa e Ilva), do departamento de física que sempre foram bastante prestativos.

Nesta vitória estão envolvidas muitas pessoas que acreditaram no meu trabalho e dedicaram seu tempo e trabalho para tornar minha vida cada vez melhor: Minha mãe, uma pessoa maravilhosa que não mediu esforços para que conseguisse estudar. Trabalhadora incansável, uma fortaleza. Ela é fundamental na minha vida e grande responsável por esta vitória. Minha tia Teresa que sempre me ajudou em tudo desde que cheguei em Rio Preto. E todos da família, obrigado pela dedicação e paciência.

Neste último ano tive uma mudança radical na minha vida e achei que talvez não fosse possível terminar este trabalho. Fui trabalhar no estado do Mato Grosso na cidade de Nova Xavantina na Unemat. Então gostaria de agradecer em especial

ao Prof. Chahine, coordenador da pós-graduação, a minha chefe de departamento em Nova Xavantina, Profa. Giselma, que acreditaram no meu trabalho e a todos os professores e amigos de Nova Xavantina.

A vida nos traz muitas surpresas. Ela me trouxe uma pessoa maravilhosa, companheira incondicional, que me dá forças nos momentos em que parece que os obstáculos são intransponíveis. Uma pessoa que enche minha vida de alegria: minha namorada Karina, obrigado por tudo. Você com certeza é responsável por cada vitória e cada momento de felicidade na minha vida.

Bom, ao meu orientador Prof. José Roberto Ruggiero, nem sei como falar de você chefe. Na vida acho que poucas pessoas são iluminadas como você. Agradeço, muito a Deus por tê-lo como orientador e amigo. Agradeço a paciência e a dedicação que teve comigo desde o princípio e por acreditar na minha capacidade. Quanto mais amadureço e convivo com você percebo a importância que teve e tem pra mim. **MUITO OBRIGADO** pelos ensinamentos, por saber “puxar minha orelha” nas horas certas, por ser compreensivo, por dar-me o apoio necessário em todas as situações, por me incentivar, enfim, por TUDO. Se algum dia conseguir ser um pouco do ser humano que o você é, ficarei muito satisfeito.

Ao Elso, meu co-orientador, uma pessoa que aprendi a admirar ainda mais nestes dois últimos anos. Sua contribuição foi muito importante para minha formação.

Agradeço também ao Prof. Dr. Jayme de Luca por ter cedido o programa computacional que usei para iniciar meus cálculos numéricos.

Gostaria de agradecer ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ) pelo suporte financeiro, sem o qual este trabalho não poderia ser desenvolvido.

Resumo

Usando o modelo de Peyrard-Bishop [2] exploramos o limiar de energia para existência de localização de energia, assim como a dinâmica de um par de cadeias de osciladores representando o DNA. Em trabalho recente [9], obteve-se o limiar de energia para uma cadeia homogênea. Neste trabalho os resultados para cadeias homogêneas e não homogêneas serão considerados. A não homogeneidade é tratada na forma de blocos com diferentes parâmetros para o potencial de Morse. O potencial de Morse é usado para simular as ligações de hidrogênio que estabilizam a dupla hélice do DNA. Esta é uma primeira aproximação para a molécula real. Os resultados mostram que o valor da energia crítica, a energia mínima para haver localização de energia, é uma função da presença de interfaces entre os blocos e a dinâmica revela que essa localização se restringe ao bloco inicialmente excitado.

Abstract

Using the Peyrard-Bishop model [2] we explore the threshold energy for energy localization as well as the dynamics of a DNA chain. In a recent work [9], the threshold energy for localization in homogeneous chain have been obtained. Here results for homogeneous and nonhomogeneous chains are considered. This nonhomogeneity is treated as blocks with different values for the parameters in the Morse Potential, the usually used potential function for simulating the H-bonds in the DNA double helix. This is a suitable first approximation for the real molecule. The results show the threshold energy depends on the presence of block interfaces and the dynamics shows that localization is restricted to the block initially excited.

Sumário

| | |
|---|-----------|
| Introdução | 1 |
| 1 Aspectos estruturais da molécula de DNA | 6 |
| 1.1 Aspectos estruturais da molécula de DNA | 6 |
| 1.1.1 DNA: composição e estrutura | 7 |
| 1.1.2 A estrutura do DNA | 9 |
| 2 Modelo dinâmico para o DNA | 16 |
| 2.1 O modelo vibracional planar ou Modelo Peyrard-Bishop (PB) . . | 16 |
| 2.2 Critério para localização de energia | 22 |
| 3 Análise de Cadeias Homogêneas | 25 |
| 3.1 Preliminares | 25 |
| 3.2 Tratamento numérico das equações de movimento | 27 |
| 3.3 Resultados | 29 |
| 4 Cadeias não-homogêneas | 39 |

| | | |
|----------|---|-----------|
| 4.1 | Preliminares | 39 |
| 4.2 | Resultados | 40 |
| 4.3 | Estudo da dinâmica da cadeia de osciladores não-homogêneos . . . | 45 |
| 5 | Deslocalização de energia | 51 |
| 5.1 | Preliminares | 51 |
| 5.2 | Cadeias homogêneas | 52 |
| 6 | Conclusões | 58 |
| A | Transformação da lagrangiana L_{PB} em termos dos modos Normais | 61 |
| B | Adimensionamento da Lagrangiana L_Y | 63 |
| C | Espectro de frequência dos modos normais | 65 |

Lista de Figuras

| | | |
|-----|---|----|
| 1 | Representação esquemática da replicação e da transcrição do DNA. | 3 |
| 1.1 | Esquema da estrutura do DNA. | 7 |
| 1.2 | Fórmula estrutural do grupo fosfato. | 8 |
| 1.3 | Fórmula estrutural da β -D-2-desoxirribose e da ribose | 8 |
| 1.4 | Fórmula estrutural das duas pirimidinas e duas purinas presentes no DNA. | 9 |
| 1.5 | Modelo de uma dupla hélice de DNA na sua forma B. | 12 |
| 1.6 | Foto de difração de raios X de uma fibra de DNA hidratada (forma B). | 13 |
| 1.7 | Diagrama de um dos filamentos da dupla hélice de DNA visto pelo eixo da hélice. | 13 |
| 1.8 | Parte de uma cadeia polinucleotídica linear. | 14 |
| 1.9 | Modelos dos pares de bases emparelhadas, timina com adenina e citosina com guanina. | 15 |
| 2.1 | Modelo vibracional planar | 18 |

| | | |
|-----|---|----|
| 2.2 | Representação gráfica para o potencial de Morse. | 19 |
| 2.3 | Representação da cadeia de osciladores da lagrangiana (2.6) . . . | 21 |
| 3.1 | Gráfico da Energia Crítica para localização em função do tamanho da cadeia N. | 30 |
| 3.2 | Gráfico da Energia Crítica para localização em função do valor de μ . | 31 |
| 3.3 | T_{max} como função da energia de excitação E com $N=200$, as unidades são arbitrárias com $\mu_{A-T} = 2$, (a) $\eta_L = 0.05$, (b) $\eta_L = 0.10$, (c) $\eta_L = 0.15$, (d) $\eta_L = 0.20$, (e) $\eta_L = 0.25$ e (f) $\eta_L = 0.30$ | 36 |
| 3.4 | Energia do j-ésimo par de base contra o n^o do par de base | 37 |
| 3.5 | Instantâneos da distribuição de energia por oscilador, cadeia homogênea com $N = 200$, $\mu_{C-G} = 3$, excitamos os osciladores 101, 102 e 103 | 38 |
| 4.1 | T_{max} como função da energia de excitação E com $N=200$, as unidades são arbitrárias, $\mu_{C-G}(1 : 100) = 3$ | 41 |
| 4.2 | T_{max} como função da energia de excitação E com $N=200$, as unidades são arbitrárias, tem-se uma cadeia do tipo $\mu_{C-G}(1 : 50)$, $\mu_{A-T}(51 : 100)$, $\mu_{C-G}(101 : 150)$ e $\mu_{A-T}(151 : 200)$ | 44 |
| 4.3 | Instantâneos da distribuição de energia por oscilador. Cadeia homogênea com dois blocos, $\mu_{C-G}(1 : 100) = 3$ e $\mu_{A-T}(101 : 200) = 2$, pulso inicial dado nos osciladores 49, 50 e 51. | 47 |

| | | |
|-----|--|----|
| 4.4 | Instantâneos da distribuição de energia por oscilador. Cadeia com dois blocos, $\mu_{C-G}(1 : 100) = 3$ e $\mu_{A-T}(101 : 200) = 2$, pulso inicial dado nos osciladores 101, 102 e 103. | 48 |
| 4.5 | Instantâneos do deslocamento ξ_i | 49 |
| 4.6 | Instantâneos da distribuição de energia por oscilador. Cadeia com quatro blocos, $\mu_{C-G}(1 : 50) = 3$, $\mu_{A-T}(51 : 100) = 2$, $\mu_{C-G}(101 : 150) = 3$, $\mu_{A-T}(151 : 200) = 2$ | 50 |
| 5.1 | T_{max} como função da energia de excitação E com $N=200$, para uma cadeia homogênea do tipo $\mu_{A-T}(1 : 200)$ | 55 |
| 5.2 | Transformada de fourier com o pulso inicial dado nos osciladores 1, 2 e 3, cadeia homogênea com $\mu = 2$ | 56 |
| 5.3 | T_{max} como função da energia de excitação E com $N=200$, $\mu_{C-G}(1 : 100) = 3$ | 57 |
| 5.4 | T_{max} como função da energia de excitação E com $N=200$, para uma cadeia do tipo $\mu_{C-G}(1 : 50)$, $\mu_{A-T}(51 : 100)$, $\mu_{C-G}(101 : 150)$, $\mu_{A-T}(151 : 200)$ | 57 |

Lista de Tabelas

| | | |
|-----|--|----|
| 3.1 | Energia Crítica (E_c) necessária para que haja localização. | 33 |
| 4.1 | Energia crítica de localização em função do tamanho dos blocos e local da excitação inicial. | 42 |

Introdução

A seqüência de pares de bases no DNA codifica a informação gênica para síntese de proteínas, que são polímeros formados por 20 aminoácidos diferentes. Como existem apenas quatro bases possíveis no DNA, uma base não pode codificar um aminoácido. Com dois pares de bases poderiam ser formados 4^2 combinações e somente 16 aminoácidos. O código genético mostra que uma trinca de pares de bases, denominada códon, codifica um aminoácido. Como existem $4^3 (= 64)$ combinações, o código é redundante. Por exemplo, GAA e GAG codificam ambos para o aminoácido glutamato. Os códons ATG e GTG são códons de iniciação da cadeia que determinam o ponto de início para a síntese da cadeia polipeptídica. Entretanto, eles também determinam os aminoácidos metionina e valina, respectivamente, em posições internas nas cadeias polipeptídicas. Os códons TAA, TGA e TAG, não codificam nenhum aminoácido. São, contudo, de extrema importância para a célula, pois são responsáveis por indicar onde deve parar a leitura da proteína que é codificada a partir do DNA. Estes códons são chamados de códon de terminação (*stop codon*).

Além dos genes, o DNA contém regiões de não codificação, cujo papel ainda

não é completamente compreendido. Alguns deles podem ser simples remanescentes da evolução dos organismos.

A leitura do código genético é uma operação complexa porque, como mostrado na figura (1.5), as bases no DNA estão localizadas no interior da molécula, resguardadas pelo esqueleto de desoxirribose-fosfato. Portanto, tem-se um obstáculo a transpor. É necessário que ocorra uma exposição das bases para haver interação entre a proteína (por ex. RNA-polimerase) e o DNA. Dessa forma, o acesso ao código genético requer grandes distorções da estrutura do DNA, para que as bases possam ser expostas para fora da pilha. Na replicação do DNA, fenômeno que ocorre durante a divisão celular por exemplo, a molécula é toda copiada em face a sua abertura, semelhante a um zíper. A transcrição do DNA é a leitura de um único gene para sintetizar uma proteína. Isto acontece com uma frequência muito maior que a replicação e é realizado pela RNA polimerase. Esta enzima executa o controle, varrendo a enorme molécula de DNA até encontrar o local de início do gene. Feito isso, a enzima abre a dupla hélice numa extensão de cerca de 20 pares de bases, como esquematizado na figura (1). As bases expostas são copiadas numa molécula de RNA que servirá de modelo para a síntese da proteína, e a "bolha de transcrição" se move ao longo do DNA, junto com a enzima, fechando os pares de bases já lidos e abrindo outros até alcançar o fim do gene.

O processo de transcrição envolve distorções na molécula, de tal magnitude que necessariamente testam o caráter não-linear das forças que ligam os vários grupos. Este processo ainda está longe do alcance de uma descrição física, devido

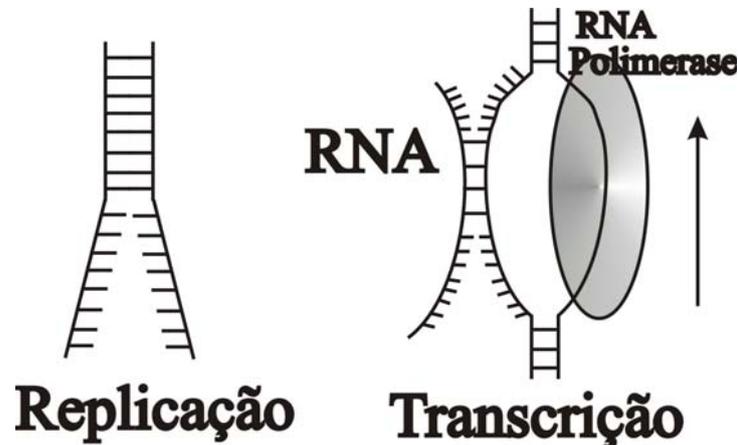


Figura 1: Representação esquemática da replicação e da transcrição do DNA.

à presença da enzima, que por si só é uma molécula muito complexa. No entanto, como será discutido mais adiante, a desnaturação térmica do DNA possui algumas similaridades com a transcrição, e pode ser descrita do ponto de vista físico. Esses estudos podem ser vistos como um primeiro passo de modelagem da transcrição.

Todos esses aspectos funcionais evidenciam grande flexibilidade e mobilidade da molécula de DNA e fundamentam uma modelagem dinâmica. Sendo assim, o DNA não é um objeto de estudo apenas para biólogos, mas apresenta várias questões de interesse para físicos, químicos, matemáticos. Devido a riqueza da molécula e aos fatores envolvidos nos processos biológicos, uma modelagem matemática da dinâmica é complexa e de custo computacional alto. Para modelagem matemática, o problema deve ser simplificado e levado em con-

sideração apenas as contribuições dominantes para o fenômeno.

O modelo vibracional de *Peyrard e Bishop* [2], o modelo PB, foi inicialmente empregado pelos autores no estudo da desnaturação térmica do DNA, ou seja, relaciona-se com a quebra das ligações de hidrogênio e a separação das duas fitas. A partir desse modelo proposto foram desenvolvidos vários estudos [3, 4] abordando a desnaturação térmica do DNA. O fato do modelo relacionar-se com a quebra das ligações de hidrogênio, é um aspecto interessante, pois para nosso estudo estamos interessados em observar movimento de grandes amplitudes, que serão relacionados a quebra das ligações de hidrogênio.

Além do tratamento estatístico de variações do modelo PB [5, 6] e do tratamento estatístico do modelo rotacional [7] para a desnaturação térmica do DNA, os aspectos dinâmicos do modelo, especialmente com relação à localização de energia [8, 9] também, vêm sendo estudado. A localização de energia implica no movimento local, restrito a um pequeno número de pares de base, com amplitudes significativamente maiores do que o restante da cadeia. A idéia é que esse movimento possa atuar como sinalizador para a maquinaria bioquímica.

Seguindo os passos da referência [9], o objetivo deste trabalho é estudar as condições para localização de energia numa cadeia de DNA não homogênea, usando o modelo PB. Neste trabalho, cujos resultados serão representados a seguir, a não homogeneidade foi introduzida através de blocos de pares de bases do tipo A-T e blocos do tipo C-G. Que variações a presença de blocos introduz nos valores da energia crítica para que ocorra a localização e como é a dinâmica dessa

localização, são perguntas que iremos responder.

Capítulo 1

Aspectos estruturais da molécula de DNA

1.1 Aspectos estruturais da molécula de DNA

No ano de 1953, James Watson e Francis Crick [1] deduziram a estrutura tridimensional do DNA, e dela seguiu o mecanismo de replicação. Os autores terminam o artigo com a afirmação "Não escapou à nossa atenção que o emparelhamento específico por nós postulado imediatamente sugere um possível mecanismo de cópia do material genético". Este feito magnífico está incluído também entre um dos mais significativos da história da biologia, pois levou à compreensão do funcionamento do gene em termos moleculares.

1.1.1 DNA: composição e estrutura

O DNA é uma macromolécula filamentar muito longa, feita de um grande número de desoxirribonucleotídeos, cada um composto por uma base, um açúcar e um grupamento fosfato. As bases das moléculas de DNA portam informação genética, enquanto o açúcar e o fosfato têm um papel estrutural, como mostra a figura (1.1).

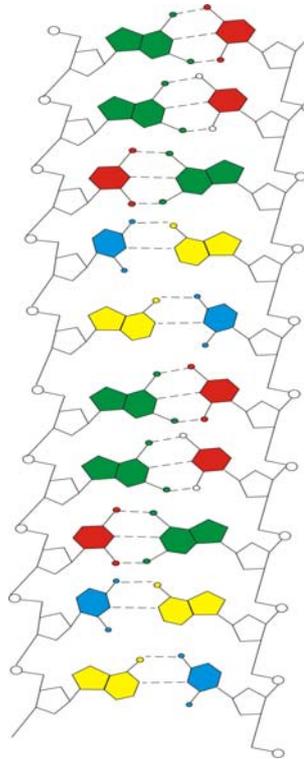


Figura 1.1: Esquema da estrutura do DNA, a seqüência do açúcar e o fosfato é mostrado em preto, e as bases purinas e pirimidinas são mostrados em cores.

Um nucleotídeo consiste de uma base nitrogenada derivada de uma *purina* ou *pirimidina*, um açúcar (*desoxirribose*, o prefixo *desoxi* indica que esse açúcar tem um átomo de oxigênio a menos que a ribose seu composto correspondente) e um

ou mais radicais fosfato.

O ácido fosfórico, responsável por uma carga residual negativa no esqueleto da molécula, é apresentado na figura (1.2).

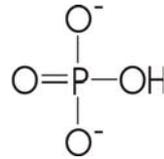


Figura 1.2: Fórmula estrutural do grupo fosfato, carregado negativamente, presente no esqueleto da molécula.

No DNA o açúcar presente é uma pentose, a desoxirribose. Ela é uma ribose modificada pela substituição do grupo hidroxila (OH) ligado ao carbono da posição 2' por um hidrogênio (H), como mostra a figura (1.3).

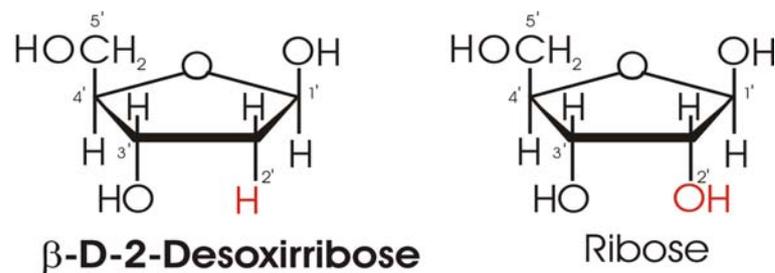


Figura 1.3: Fórmula estrutural da β -D-2-desoxirribose e da ribose

No DNA as bases purinas são a *adenina* (A) e a *guanina* (G), e as pirimidinas são *timina* (T) e a *citocina* (C). A timina apresenta um grupo metil (CH_3) ligado ao carbono C_5 e grupos carbonila ($C=O$) nas posições C_4 e C_2 . A citosina apresenta um grupo amina (NH_2) ligado ao carbono C_4 e um grupo carbonila na

posição C_2 . A adenina possui um grupo amino ligado ao carbono C_6 . A guanina possui um grupo amino ligado ao carbono C_2 e um grupo carbonila na posição C_6 . A figura (1.4) apresenta a forma estrutural destas bases.

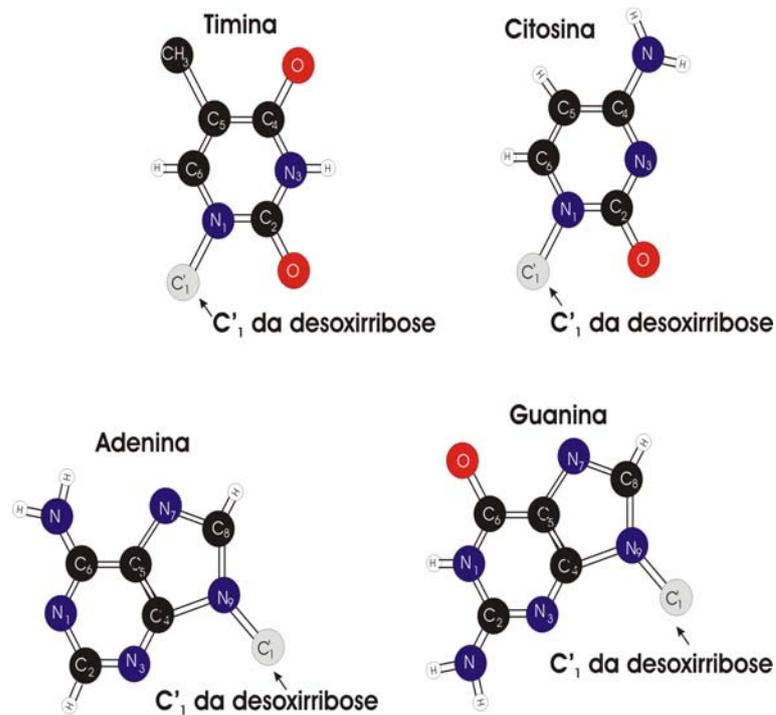


Figura 1.4: Fórmula estrutural das duas pirimidinas e duas purinas presentes no DNA.

1.1.2 A estrutura do DNA

Como mencionado no início do capítulo, em 1953 James Watson e Francis Crick [1], estudando fotografias de difração de raios X de fibras de DNA feitas por Rosalind Franklin e Maurice Wilkins deduziram a estrutura tridimensional do

DNA, da qual descrevem-se brevemente as características mais importantes:

1. Duas cadeias polinucleotídicas helicoidais estão enroladas em torno de um eixo comum. As cadeias correm no sentido oposto (figuras (1.5) e (1.6)).
2. As bases purinas e pirimidinas estão do lado de dentro da hélice, enquanto as unidades fosfato e desoxirribose estão fora (figura (1.7)). Os planos das bases são perpendiculares ao eixo da hélice. Os planos das desoxirriboses são quase perpendiculares aos das bases.
3. Os grupos fosfatos e desoxirriboses dos nucleotídeos concatenados formam uma estrutura de esqueleto conforme a figura (1.8). A ligação glicosídica é responsável pela conexão entre a base nitrogenada e o açúcar. A ligação que junta dois nucleotídeos a partir do carbono $C_{3'}$ do açúcar de um nucleotídeo ao carbono $C_{5'}$ do nucleotídeo adjacente é a ligação fosfodiéster formada durante a síntese bioquímica do DNA.
4. Na estrutura do polinucleotídeo existem duas extremidades distintas, uma dita de 5' e outra 3'. Essas extremidades são bem diferentes, química e biologicamente. Na extremidade 3' temos um grupo hidroxila e na 5' um grupo fosfato, o que atribui uma polaridade a cada fita.
5. O diâmetro da hélice é de 20 Å. Bases adjacentes são separadas por 3,4 Å, ao longo do eixo da hélice, e com uma rotação de 36 graus. Dessa forma, a estrutura helicoidal repete-se após 10 pares de bases em cada cadeia, isto

é, a intervalos de 34 Å. Observe a figura (1.7).

6. As duas cadeias são mantidas juntas por ligações de hidrogênio entre os pares de bases. A adenina está sempre emparelhada com timina formando duas ligações de hidrogênio. A guanina sempre emparelhada com citosina formando três ligações de hidrogênio como é apresentado na figura (1.9).
7. A seqüência de bases ao longo de uma cadeia polinucleotídica não é de nenhuma forma restrita. A seqüência das bases é portadora da informação genética.

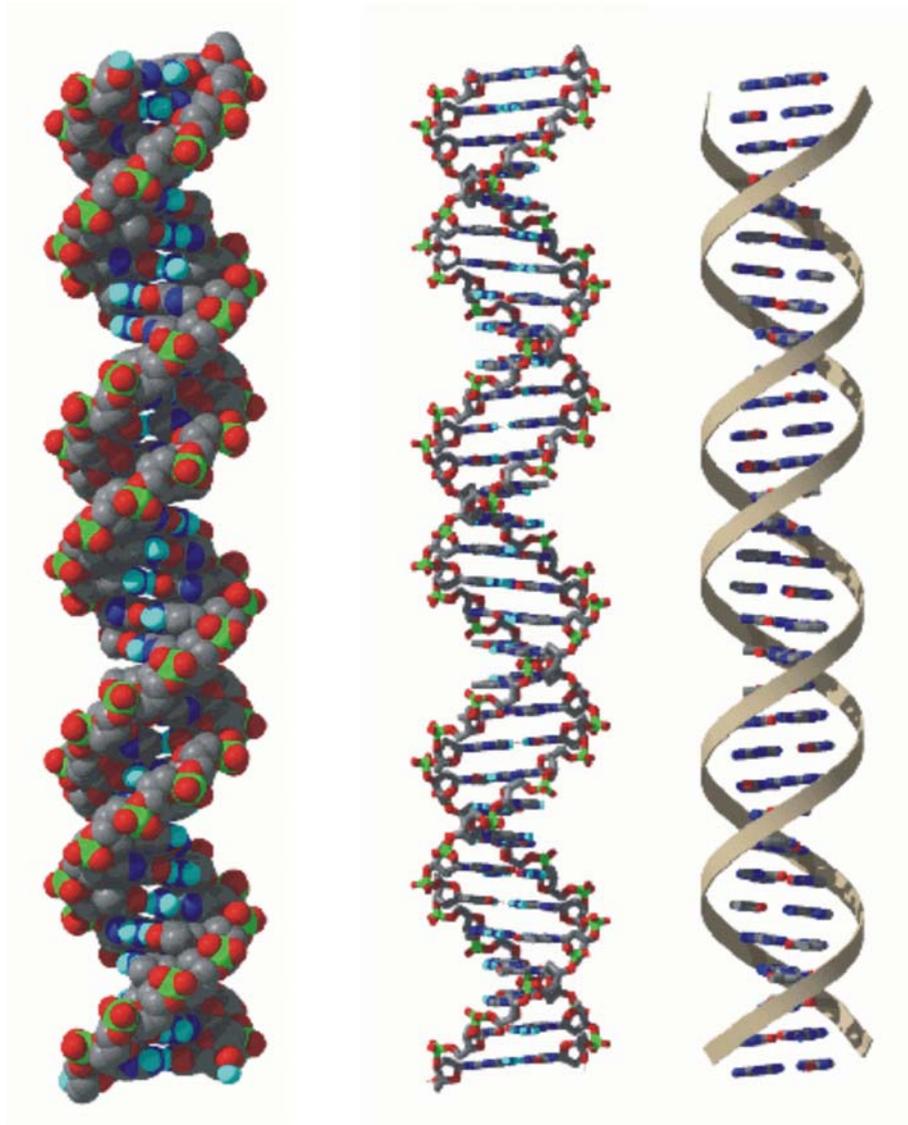


Figura 1.5: Modelo de uma dupla hélice de DNA na sua forma B. A estrutura se repete em intervalos de 34 Å, que correspondem a 10 resíduos em cada cadeia.

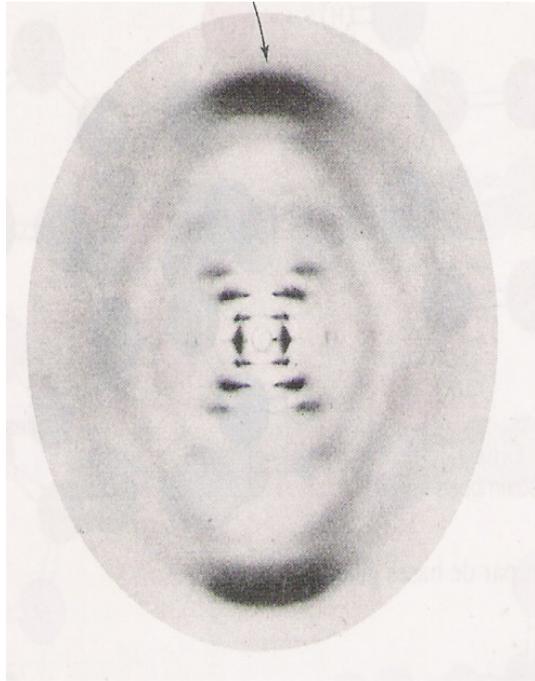


Figura 1.6: Foto de difração de raios X de uma fibra de DNA hidratada (forma B). A cruz central é diagnóstica de uma estrutura helicoidal. Os fatores arcos no meridiano surgem do empilhamento de bases, que estão separados em $3,4 \text{ \AA}$.

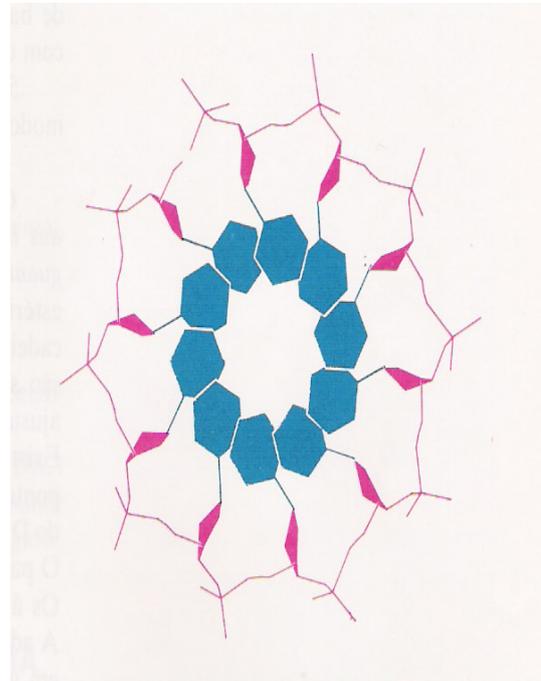


Figura 1.7: Diagrama de um dos filamentos da dupla hélice de DNA visto pelo eixo da hélice. As bases (aqui, todas pirimidinas) estão do lado interno, enquanto a seqüência desoxirribose e fosfato está externa. A simetria de 10 em 10 é evidente. As bases são mostradas em azul e as desoxirriboses em vermelho.

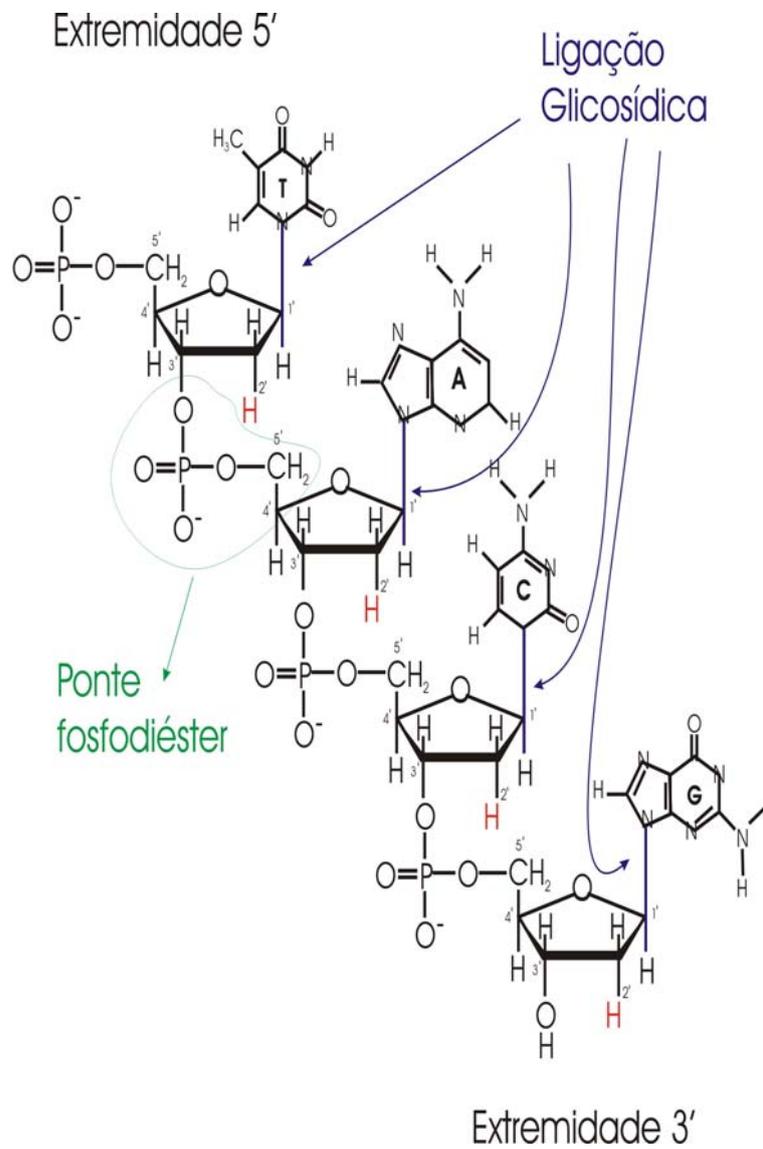


Figura 1.8: Parte de uma cadeia polinucleotídica linear, com as ligações glicosídicas indicadas pelas setas azuis, e a fosfodiéster pela linha verde.

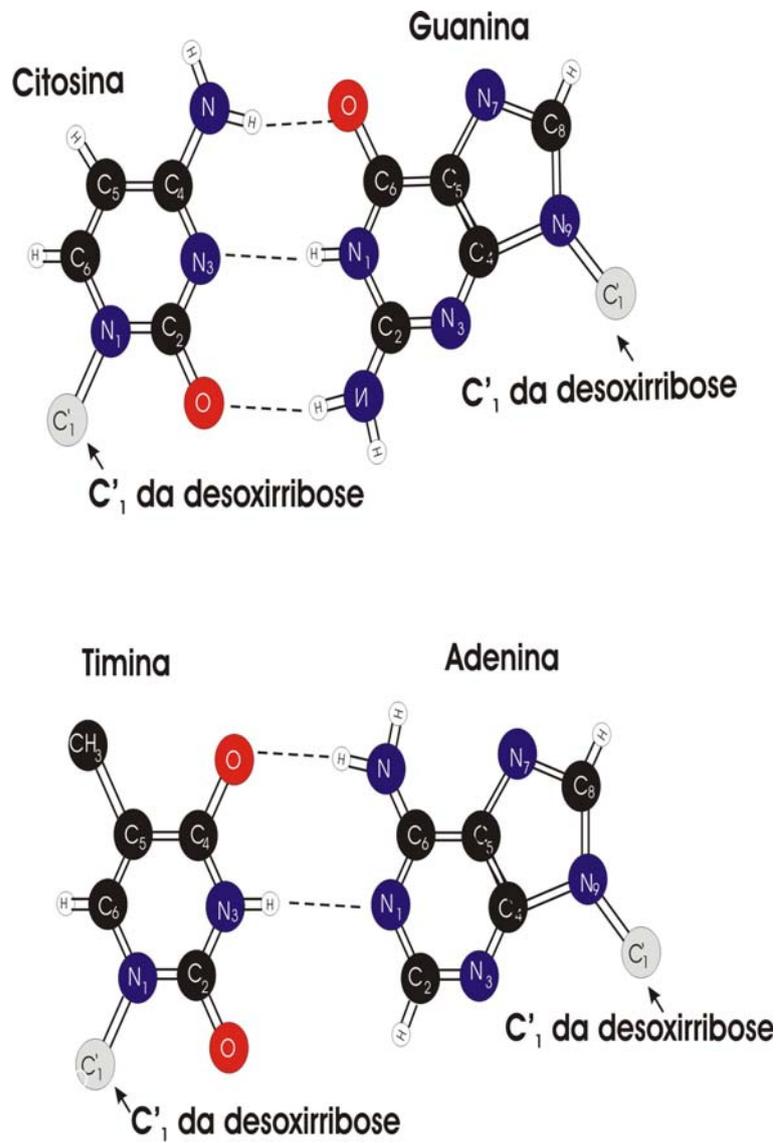


Figura 1.9: Modelos dos pares de bases emparelhadas, timina com adenina e citosina com guanina.

Capítulo 2

Modelo dinâmico para o DNA

2.1 O modelo vibracional planar ou Modelo Peyrard-Bishop (PB)

No modelo vibracional planar proposto por Peyrard e Bishop [2] e ilustrado na figura (2.1), a dupla hélice é reduzida a uma dupla cadeia de osciladores harmônicos acoplados por um potencial V_j . A massa de cada oscilador (m_{u_j} ou m_{v_j}) representa um nucleotídeo e cada cadeia, u e v , uma fita do DNA. Cada nucleotídeo u_j ou (v_j) está ligado aos seus vizinhos ao longo da cadeia por potencial elástico (mola) de constante (k_{u_j} ou k_{v_j}) simulando a estrutura (esqueleto) da fita. O potencial V_j acopla os nucleotídeos adjacentes de cada uma das fitas, ou seja, u_j com v_j . Considerando a rigidez implicada pela estrutura química do esqueleto das fitas e o interesse em estudar a dinâmica da abertura das ligações

de hidrogênio entre as cadeias, os movimentos são restritos às vibrações na direção perpendicular às cadeias (hélice). Dessa forma, o modelo não leva em conta a estrutura helicoidal do DNA (para o modelo helicoidal ver ref. [4, 7]), mas restringe a análise a uma estrutura planar.

Uma função que tem sido utilizada para simular as ligações de hidrogênio é o potencial de Morse, equação (2.1). No caso, essa função depende do deslocamento relativo ($u_j - v_j$) entre os nucleotídeos (pares de bases), ou seja, $V_j = V(u_j - v_j)$. A forma funcional desse potencial é

$$V(y) = D(e^{-ay} - 1)^2 \quad (2.1)$$

e o gráfico dessa função está mostrado na figura (2.2). O parâmetro D representa a energia de dissociação (quebra da ligação) e o parâmetro a , que tem dimensão de inverso de comprimento, estabelece a escala espacial do potencial. Esse potencial é conveniente porque quando $y = 0$, mínimo da função, a ponte de hidrogênio está na sua configuração ótima. Para $y < 0$, ocorre uma aproximação dos átomos/grupos que formam a ligação de hidrogênio, o potencial é fortemente repulsivo representando os impedimentos estéricos associados a esse tipo de movimento e, para $y > 0$, um afastamento dos átomos que compõe a ponte, ocorre uma atração que desaparece para valores muito grandes de y , ou seja, a ligação se rompe.

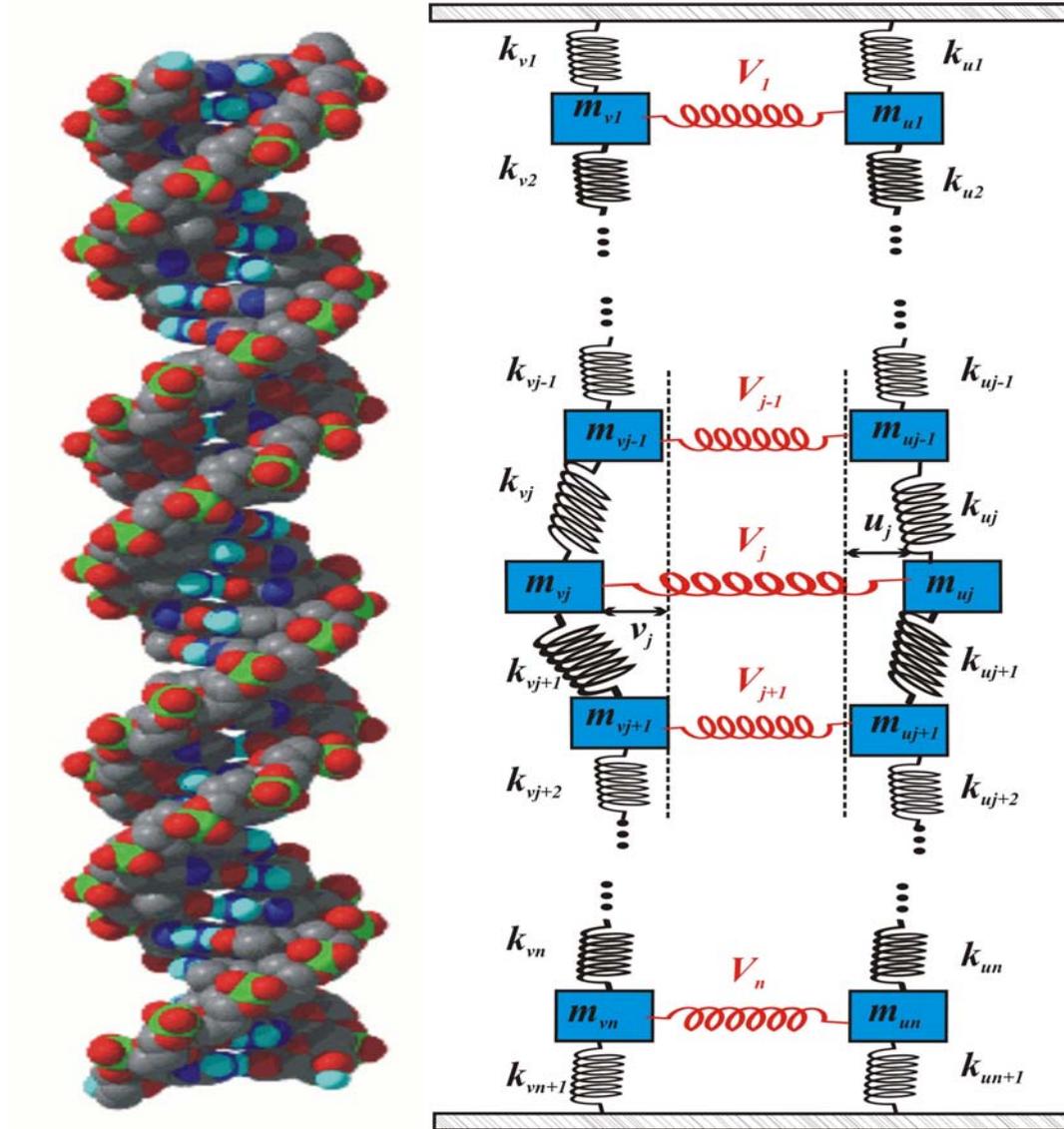


Figura 2.1: À esquerda temos modelo de uma dupla hélice de DNA na sua forma B. À direita o modelo planar, que consiste de um sistema mecânico que representa o modelo simplificado bidimensional para a dinâmica do DNA, dado pela Lagrangiana (2.2). Consiste de duas cadeias de osciladores massa-mola com extremidades fixas. Notemos que as massas estão acopladas entre si por um potencial harmônico e um potencial V_j que vai depender do emparelhamento (A-T) ou (C-G) que acopla as duas cadeias.

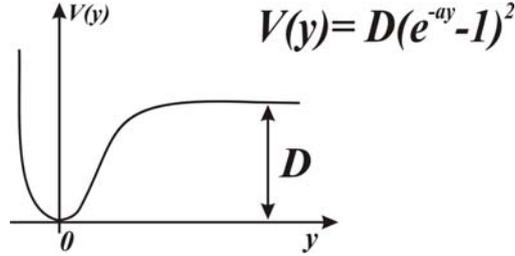


Figura 2.2: Representação gráfica para o potencial de Morse. Potencial usado para simular as ligações de hidrogênio.

A Lagrangiana para o modelo da macromolécula DNA, proposto por Peyrard-Bishop (PB) é escrita como

$$L_{PB} = \sum_{j=1}^N \frac{m}{2} (\dot{u}_j^2 + \dot{v}_j^2) - \frac{k}{2} (u_{j+1} - u_j)^2 - \frac{k}{2} (v_{j+1} - v_j)^2 - D [\exp a(v_j - u_j) - 1]^2, \quad (2.2)$$

onde $(u_j$ e $v_j)$ denotam os deslocamentos relativos das bases dos nucleotídeos, com relação a cada fita de DNA, e o ponto representa a derivada em relação ao tempo. O número N indica a quantidade de pares de base da cadeia de DNA. A equação (2.2) está escrita com uma simplificação introduzida em [2, 10], que considera todas as massas homogêneas, de forma que $m_{u_j} = m_{v_j} = m$ e que as interações ao longo de cada fita são indistinguíveis [10, 11], ou seja, $k_{u_j} = k_{v_j} = k$. Os valores experimentais para estes parâmetros estão discutidos na literatura [3, 10]: a massa dos pares de bases é aproximadamente $300u.m.a. = 5.0 \times 10^{-25}kg$ e a constante linear da mola é $1.50eV/\text{Å}^2$. As ligações de hidrogênio

são simuladas pelo potencial de Morse com $a=2.20\text{\AA}^{-1}$ e para D [12] temos no caso emparelhamento do tipo (A-T), $D_{A-T} \approx 0.304eV$ e no caso emparelhamento do tipo (C-G), $D_{C-G} \approx 0.728eV$. É importante destacar que há uma grande variação nos parâmetros segundo cada autor, ver ref. [2, 3, 10, 11].

Podemos reescrever a Lagrangiana (2.2) em termos dos modos normais, definidos por (para detalhes ver apêndice A)

$$x_j = \frac{(u_j + v_j)}{\sqrt{2}} \quad y_j = \frac{(u_j - v_j)}{\sqrt{2}}. \quad (2.3)$$

Com isso, a Lagrangiana PB (2.2) é transformada em $L_{PB} = L_X + L_Y$ com

$$L_X = \sum_{j=1}^N \frac{m}{2} \dot{x}_j^2 - \frac{k}{2} (x_{j+1} - x_j)^2, \quad (2.4)$$

que depende apenas da coordenada x , e L_Y , que depende somente da coordenada y , como

$$L_Y = \sum_{j=1}^N \frac{m}{2} \dot{y}_j^2 - \frac{k}{2} (y_{j+1} - y_j)^2 - D[\exp(-\sqrt{2}ay_j) - 1]^2. \quad (2.5)$$

Para efeito de simplificação dos cálculos, introduziu-se um fator de adimensionamento temporal $\tau \equiv \sqrt{\frac{k}{m}}t$ e um adimensionamento nas coordenadas $\xi_j \equiv \sqrt{2}ay_j$ (para detalhes ver apêndice B). Este procedimento transforma a lagrangiana L_Y da eq.(2.5) em uma forma normal $L_Y \equiv \frac{kL}{2a^2}$ com

$$L = \sum_{j=1}^N \frac{1}{2} \dot{\xi}_j^2 - \frac{1}{2} (\xi_{j+1} - \xi_j)^2 - \frac{\mu^2}{2} [\exp(-\xi_j) - 1]^2, \quad (2.6)$$

onde o ponto denota, agora, a derivada com relação a τ e $\mu^2 = \frac{4Da^2}{k}$.

A Lagrangiana da equação (2.6) pode ser mapeada no problema de uma cadeia unidimensional de osciladores harmônicos, sendo que cada oscilador é acoplado a vínculos externos à cadeia pelo termo de potencial não linear na equação (vide fig.(2.3)).

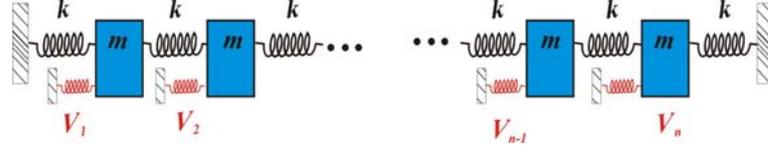


Figura 2.3: Representação da cadeia de osciladores da lagrangiana (2.6)

Para pequenas amplitudes, o espectro de frequências dos modos normais da equação (2.6) é (ver Apêndice C)

$$\omega(l) = \sqrt{\mu^2 + 4 \sin^2(l/2)}, \quad (2.7)$$

onde $l = (j\pi/N + 1)$ com $j = 1, \dots, N$. A faixa de frequências dos modos normais está entre $\mu \leq \omega \leq \sqrt{\mu^2 + 4}$ e constitui a banda de dispersão, enquanto as duas relações $0 < \omega < \mu$ e $\omega > \sqrt{\mu^2 + 4}$, representam as regiões abaixo e acima da banda de dispersão, respectivamente. Para que exista a localização é esperado que as frequências que compõe o movimento estejam localizadas fora da faixa de frequências dos modos normais, ou seja, estejam localizadas acima ou abaixo da banda de dispersão. Para frequências na banda de dispersão, a energia inicialmente localizada espalha-se pela rede. Existem estudos para cadeias unidimensionais, que investigam essa troca de energia entre os modos de vibração do

sistema linearizado (ver ref. [19] e bibliografia citada).

2.2 Critério para localização de energia

O conceito de localização usado neste trabalho, segue aquele de Anderson [22], que mostrou que numa cadeia de osciladores harmônicos desordenados, alguns modos são localizados, ou seja, somente alguns osciladores se movem significativamente, os demais permanecendo essencialmente em repouso. Tipicamente existe um limiar de energia a partir do qual isso ocorre. Para valores menores a energia é distribuída por todos os modos.

A energia do sistema descrito pela lagrangiana (2.6) é dada por $H = \sum_j p_j \dot{\xi}_j - L$, onde $p_j = \frac{\partial L}{\partial \dot{\xi}_j}$, ou seja,

$$H = \sum_{j=1}^N \frac{1}{2} p_j^2 + \frac{1}{2} (\xi_{j+1} - \xi_j)^2 + \frac{1}{2} \mu^2 [\exp(-\xi_j) - 1]^2. \quad (2.8)$$

Vamos associar a cada oscilador a energia E_j , $j = 1, \dots, N$, escrita como

$$E_j = \frac{1}{2} p_j^2 + \frac{1}{4} (\xi_{j+1} - \xi_j)^2 + \frac{1}{4} (\xi_j - \xi_{j-1})^2 + \frac{1}{2} \mu^2 [\exp(-\xi_j) - 1]^2. \quad (2.9)$$

onde incluímos 50% da interação com o vizinho mais próximo de cada lado, tal que a soma dos E_j , $j = 1, \dots, N$ seja a energia total constante. A entropia informacional [20] é definida por:

$$S = - \sum_{j=1}^N e_j \ln e_j, \quad (2.10)$$

onde $e_j = \frac{E_j}{\sum_{j=1}^N E_j}$, são as energias instantâneas normalizadas. Normalmente a energia está distribuída por cerca de r osciladores onde $r < N$, dessa forma r dos osciladores tem energia e_j . Essa energia é da ordem de $\frac{1}{r}$ e a energia dos $N - r$ osciladores restantes pode ser desprezada. Assim, a eq. (2.10) nos diz que $S \simeq \ln(r)$. Com essa aproximação somos motivados a definição de

$$N_{osc} \equiv \exp(S), \quad (2.11)$$

como o número efetivo de osciladores que compartilham a mesma energia. Como estamos trabalhando sempre com parâmetros normalizados é conveniente a normalização deste também

$$n_{osc} \equiv \frac{N_{osc}}{N}. \quad (2.12)$$

O parâmetro normalizado é um número entre 0 e 1, uma vez que a entropia eq. (2.10) é sempre menor que $\ln N$. Devido a flutuações na energia n_{osc} não assume seu valor máximo 1. Para o cálculo das flutuações, consideramos δe_j o desvio da equipartição da energia, logo temos $e_j = \bar{e} + \delta e_j$. Substituindo esse valor na eq. (2.10), fazendo a expansão em série de Taylor da função logarítmica como sendo $\ln\left(1 + \frac{\delta e_j}{\bar{e}}\right) = \frac{\delta e_j}{\bar{e}} - \left(\frac{1}{2}\right)\left(\frac{\delta e_j}{\bar{e}}\right)^2$ e somando sobre todos os j obtemos:

$$n_{osc} = \frac{1}{N} \exp\left\{-N\bar{e} \ln(\bar{e}) - N\frac{(\overline{\delta e})^2}{2\bar{e}}\right\} = \exp\left\{-N\frac{(\overline{\delta e})^2}{2\bar{e}}\right\} \quad (2.13)$$

Tomando $\bar{e} = \frac{1}{N}$ e supondo que da mecânica estatística podemos assumir que para cada modo normal de vibração temos $(\overline{\delta e})^2 = \bar{e}^2$ (isto é verdade apenas para dinâmica de uma cadeia linear), assim os N se cancelam dando um valor máximo

$n_{osc} = \exp(-0.5) \simeq 0.61$. Os cálculos mostram que o resultado não depende do número de osciladores, caso N seja grande, e mostra também que o valor é diferente de 1, como havíamos mencionado anteriormente. Cálculos mais precisos foram feitos para uma cadeia de osciladores com termos não lineares [19]

$$n_{osc} = 0.74 \tag{2.14}$$

para equipartição da energia entre os osciladores da cadeia. Este valor tem sido checado numericamente e apresenta boa concordância [19].

Para uma escolha aleatória da localização inicial, testes numéricos mostram que variando o valor de n_{osc} , o tempo necessário para que a energia espalhe pela rede é da ordem de N [9], logo N foi considerado um tempo típico para o espalhamento. O critério de localização utilizado [9], estabelece que um estado é localizado quando n_{osc} é significativamente menor que o valor observado para equipartição de energia 0.61, e o tempo computacional é maior que 50 vezes o tempo de espalhamento. Então, o critério prático adotado para considerar que houve localização de energia é:

- $n_{osc} < \eta_L \equiv (N_{max}/N)$;
- $t_{lim} > 50N$;

onde N_{max} é o número de osciladores com energia maior que $\frac{E}{N}$ e t_{lim} é o tempo limite (do ponto de vista computacional) para o qual considera-se que o sistema permaneça com a energia localizada em um número de osciladores menor do que N_{max} .

Capítulo 3

Análise de Cadeias Homogêneas

3.1 Preliminares

O problema da equipartição de energia em cadeias de osciladores vem sendo estudado há muito tempo. Para cadeia de osciladores não lineares, o marco histórico é o problema de Fermi-Pasta-Ulam (FPU) [13], no qual se discute a equipartição de energia nos modos normais de vibração para um potencial do tipo $\frac{1}{2}kx^2 + \frac{1}{4}\beta x^4$. Usando, pela primeira vez um computador, o Maniac (Mathematical Analyser, Numerical Integrator and Computer) construído em 1952, em Los Alamos [14], FPU pensaram que, devido ao termo não linear, a energia introduzida em um único modo de vibração em suas simulações deveria lentamente ser transmitida aos outros modos de vibração até que a equipartição de energia fosse alcançada, como previsto pela física estatística. No início dos cálculos, os resultados sugeriam que certamente este seria o caso, os outros modos de vibração estavam sendo

sucessivamente excitados. Entretanto, por acidente [15], um dia deixaram o programa funcionar por muito tempo além daquele no qual o sistema mostrava um comportamento estacionário. Quando perceberam o engano voltaram à sala e observaram que o sistema estava muito próximo do estado inicial.

O estado inicial parecia quase perfeitamente recuperado depois de um tempo correspondente a cerca de 157 vezes o período do modo inicialmente excitado. Deste então, foram então executados cálculos com computadores mais rápidos e com maior precisão, mostrando que o fenômeno se repetia muitas vezes, e que o período de "super retorno" existia e o estado inicial era recuperado com grande precisão. Ao contrário da expectativa dos autores, a energia do modo inicialmente excitado era transmitida em direção aos outros modos, mas não se alcançava o estado de equipartição, com a energia retornando ao estado inicial. Este resultado altamente notável ficou conhecido como *Paradoxo de FPU*, mostrando que a não linearidade não é suficiente para garantir a equipartição da energia, dando origem a uma linha de estudo de sistemas não lineares e localização de energia em cadeias de osciladores.

Como foi dito anteriormente, vamos discutir aqui a questão da localização espacial da energia para o DNA, representado no modelo discutido no capítulo anterior. Diferentemente do problema de FPU que, ao excitar um modo de vibração, uma única frequência distribuía a energia por toda a rede, nosso objetivo é discutir a localização espacial, ou seja, em que condições a energia, colocada inicialmente em poucos osciladores (no caso três) permanece por um tempo in-

definido (computacionalmente longo) num número restrito de osciladores, dados pelos critérios discutidos no capítulo precedente.

3.2 Tratamento numérico das equações de movimento

As equações de movimento para o sistema descrito pela lagrangiana (2.6) são:

$$\ddot{\xi}_j - \xi_{j+1} + 2\xi_j - \xi_{j-1} - \mu^2[\exp(-\xi_j) - 1] \exp(-\xi_j) = 0 \quad i = 1, 2, \dots, N. \quad (3.1)$$

Vamos resolver esse sistema de equações sujeitas a condições de contorno periódicas, no sentido de simularmos uma cadeia infinita.

A integração dessas equações foi realizada usando-se o método simplético de Runge-Kutta-Nystrom de décima ordem [16]. Métodos simpléticos são importantes porque são construídos para preservar invariantes do problema [17, 18]. No presente caso, o método preserva invariante a estrutura do hamiltoniano, o que implica na conservação de energia. Testes realizados por Sanz-Serna e Calvo [18] mostraram uma excelente precisão para integrações por períodos de tempo longos, mesmo para passos de integração grandes, por exemplo, para hamiltonianos do tipo Hénon-Heiles [23], e o problema de Kepler [23, 16]. A alta ordem do integrador permite que para um passo 0.6 a energia seja conservada com uma precisão de 10^{-10} , para um período de integração da ordem de 10^{10} [9].

É importante destacar que a localização que queremos estudar depende da não linearidade existente nas ligações de hidrogênio, representados no modelo pelo potencial de Morse. Porém, a não linearidade é uma condição necessária mas não suficiente. Em outras palavras, não basta excitar o sistema com uma certa energia para que a localização se verifique.

Para cada valor da energia, a excitação inicial é colocada sempre em três osciladores, todos os demais sendo mantidos em repouso, nas suas posições de equilíbrio ($\dot{\xi} = 0$, $\xi = 0$). Admitimos que os osciladores $j-1$, j , $j+1$ sejam excitados, o critério usado foi sempre que ($\dot{\xi}_{j+1} = \dot{\xi}_{j-1}$, $\xi_{j+1} = \xi_{j-1}$), a excitação é simétrica e $\xi_j = 0$. Quarenta e nove combinações de $\dot{\xi}_{j+1}$, ξ_{j+1} e $\dot{\xi}_j$ compatíveis com a energia dada ao sistema, ou seja, que satisfazem a equação

$$E = \frac{1}{2}\dot{\xi}_j^2 + \dot{\xi}_{j+1}^2 + \xi_{j+1}^2 + \mu^2[\exp(-\xi_{j+1}) - 1]^2. \quad (3.2)$$

foram testados.

Dois critérios coerentes com as discussões da seção (2.2) foram utilizados para interromper o processamento:

1. quando $n_{osc} > \eta_L$ o que significa que não houve localização, ou
2. quando $n_{osc} < \eta_L$ e o tempo t de integração for maior que t_{lim} .

3.3 Resultados

Na figura (3.1) mostramos um gráfico da energia crítica necessária para localização em função do tamanho da cadeia N , mantendo fixo $N_{max} = 20$, e considerando $\mu = 3$. Deste gráfico, pode-se observar que a energia crítica aumenta com o crescimento da cadeia de uma maneira geométrica, com E_c tendendo a um patamar constante com N . Por razões computacionais (tempo de processamento) limitamos $N = 200$ por entender que, qualitativamente, os resultados seriam os mesmos e, pelo fato de que no processo de transcrição do DNA, uma região com cerca de 20 pares de base, o TATA-box e bases vizinhas, desempenha um papel importante no processo. Além disso 20 pares de base representariam somente dez por cento da cadeia, ou seja, $\eta_L = 0.1$, o que serviria para caracterizar bem a localização, lembrando que para uma cadeia não linear a equipartição de energia ocorre para $\eta = 0.74$ (capítulo 2).

Para cadeias com $N = 200$, os parâmetros de controle discutidos no capítulo anterior, têm valores: tempo de espalhamento $T_{esp} \approx 200$, tempo limite $t_{lim} \approx 53N > 10600$ unidades de tempo. Formalmente, esse tempo deveria ser infinitamente grande, $t_{lim} \rightarrow \infty$, porém isso é impossível do ponto de vista computacional. Testes realizados com $t_{lim} = 1000N$ mostraram que não há mudança quantitativa significativa nos valores da energia crítica, E_c , encontrados.

Analisaremos os resultados por meio de gráficos de T_{max} como função da energia de excitação E , onde T_{max} representa o tempo necessário para que a energia

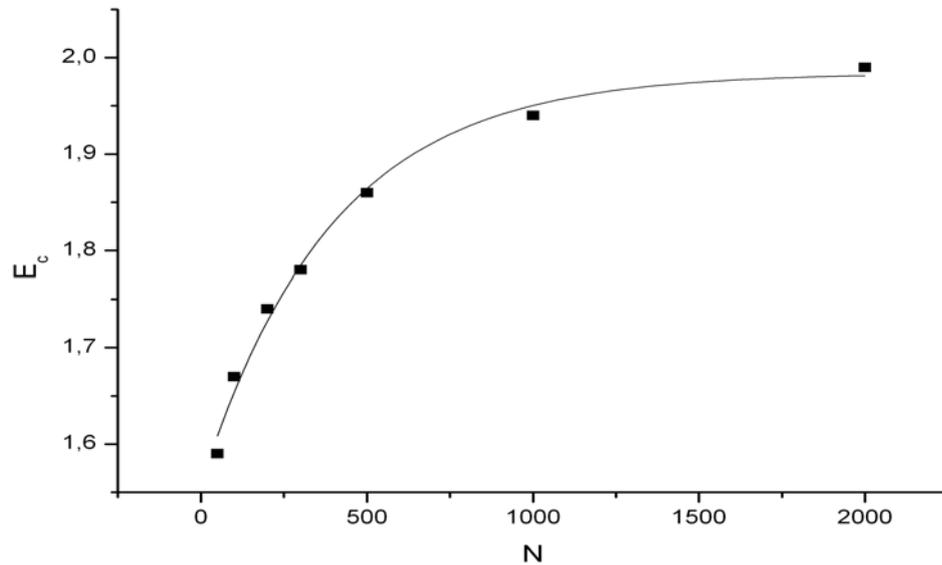


Figura 3.1: Gráfico da Energia Crítica para localização em função do tamanho da cadeia N . A linha apenas liga os pontos.

se distribua por um número de osciladores maior que o estabelecido por η_L , ou ser igual a t_{lim} , caso no qual dizemos que há localização de energia. Tipicamente, esse gráfico mostra uma inflexão abrupta, tipo sigmóide, com a energia. A energia crítica para localização é o valor da energia no ponto médio dessa inflexão.

O emparelhamento (A-T) forma duas ligações de hidrogênio na molécula de DNA, enquanto o emparelhamento (C-G) forma três, como já discutido na seção (1.1.2). Usando os parâmetros, usados no capítulo 2 obtemos $\mu_{A-T} = 2$ e $\mu_{C-G} = 3$. Insistimos que há, na literatura, uma grande variação nos valores destes parâmetros. Em função disto, avaliamos como varia o valor da energia

crítica E_c em função do parâmetro μ . Os resultados estão mostrados na figura (3.2). Vê-se que o crescimento do parâmetro μ torna a energia crítica para localização insensível ao valor do parâmetro.

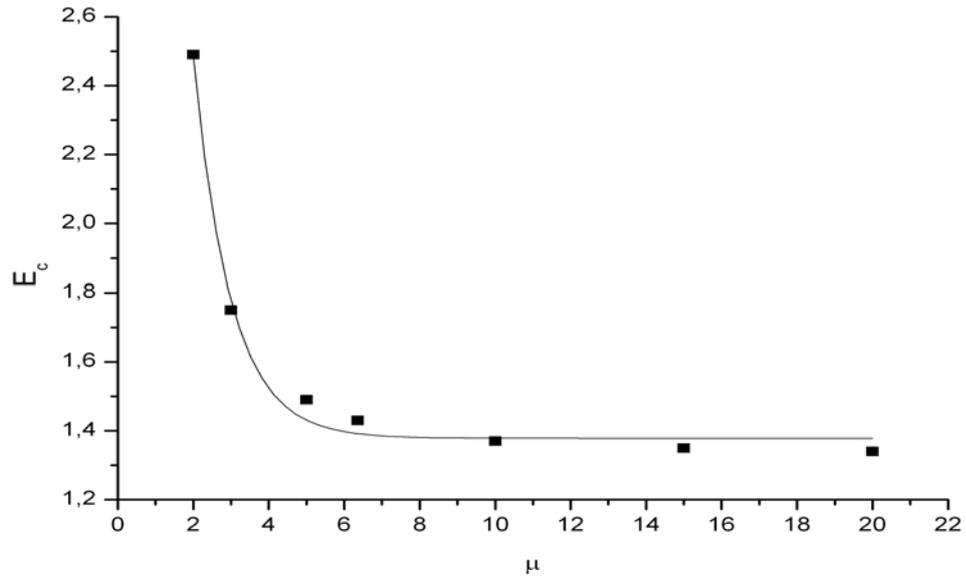


Figura 3.2: Gráfico da Energia Crítica para localização em função do valor de μ . A linha é a interpolação dos pontos usando uma função exponencial com decaimento de primeira ordem.

Outros testes foram realizados para discutir a influência da escolha de η_L na determinação do limiar de energia para a localização, para uma cadeia homogênea com $\mu_{(A-T)}$ e $\mu_{(C-G)}$. Na fig. (3.3) estão mostrados os resultados de T_{max} como função da energia fornecida à cadeia para vários valores de η_L .

No caso da figura (3.3(a)), $\mu_{(A-T)}$ o critério de localização ($n_{osc} < 0.05 \equiv$

η_L para $t \geq t_{lim}$) não foi satisfeito para os valores de energia inicial utilizados. Isso significa que, para esse valor do parâmetro $\mu_{(A-T)}$, a energia nunca permanece em menos que dez osciladores ($\eta_L = 0.05$) por um tempo suficientemente longo. O detalhamento da dinâmica da cadeia para esse valor de η_L pode ser obtido da figura (3.4), na qual a energia E_j de cada oscilador para um conjunto de instantes de simulação, nas mesmas condições da figura (3.3 (a)), está mostrado como função da posição do oscilador na cadeia. Em outras palavras a figura apresenta um conjunto de "fotografias" da cadeia, observada sob a ótica da energia. Vê-se claramente que para o conjunto de instantes mostrado, a energia se distribui por um número de osciladores maior do que dez, o número limite correspondente a $\eta_L = 0.05$ para localização.

Nas figuras (3.3) de (b) até (f), para $\mu_{A-T} = 2$, estão mostrados os dados de T_{max} como função da energia de excitação. Evidencia-se o limiar de energia para valores de η_L entre 0.1 e 0.30, por exemplo, para $\eta_L = 0.10$, o limiar de energia é aproximadamente, 2.49, ao passo que para $\eta_L = 0.30$, o limiar é de aproximadamente 1.63. Pode-se observar na medida que η_L aumenta, a energia necessária para localização diminui. Os dados obtidos destas figuras estão resumidos na tabela (3.1).

Os resultados do estudo do limiar de energia em função de η_L para μ_{C-G} estão, também, mostrados na tabela (3.1). Dois aspectos merecem destaque com relação ao caso $\mu_{(A-T)}$. Primeiro, encontra-se um limiar de energia mesmo para $\eta_L = 0.05$ com $E_c = 2, 31$, uma energia crítica menor que a energia necessária para

| η_L | N_{max} | $\mu_{(A-T)} = 2$ | $\mu_{(C-G)} = 3$ |
|----------|-----------|-------------------|-------------------|
| 0.05 | 10 | - | 2.31 ± 0.01 |
| 0.10 | 20 | 2.49 ± 0.01 | 1.75 ± 0.01 |
| 0.15 | 30 | 2.07 ± 0.01 | 1.52 ± 0.01 |
| 0.20 | 40 | 1.85 ± 0.01 | 1.42 ± 0.01 |
| 0.25 | 50 | 1.73 ± 0.01 | 1.36 ± 0.01 |
| 0.30 | 60 | 1.63 ± 0.01 | 1.32 ± 0.01 |

Tabela 3.1: Energia Crítica (E_c) necessária para que haja localização. N_{max} é o número de osciladores que tem energia maior do que E/N .

uma cadeia homogênea do tipo (A-T) obter localização com $\eta_L = 0.10$ ($N_{max} = 20$). Além disso, as energias críticas para os demais valores são menores. Para $\eta_L = 0.10$, por exemplo, o limiar de energia cai de aproximadamente 2.49 para $\mu_{(A-T)}$, para aproximadamente 1.74 para μ_{C-G} . Para ambos os casos pode-se observar que a energia crítica para valores intermediários de η_L entre 0.15 e 0.30 diminui lentamente. Este fato era esperado pois aumentar η_L implica que a energia pode se distribuir por um número maior de osciladores. Como consequência, com uma energia de excitação menor obtém-se a localização segundo os critérios estabelecidos. Por exemplo, para uma cadeia homogênea do tipo (C-G) com $\eta_L = 0.05$ tem-se $E_c = 2,31$, enquanto, para $\eta_L = 0.10$ tem-se $E_c = 1,74$. Isso significa que permitir que a localização passe de dez osciladores para vinte provoca uma diminuição de 32% na energia de localização. Enquanto que para $\eta_L = 0.15$

tem-se $E_c = 1,55$ e conseqüentemente uma diminuição de 12%, na energia crítica para localização em comparação com $\eta_L = 0.10$.

Estes resultados indicam que a não linearidade é, nesse caso de cadeias homogêneas, fundamental para localização da energia. Assim, quanto maior for a importância relativa desse termo, através do parâmetro μ , mais facilmente se manifesta o fenômeno da localização, embora isso não ocorra pra qualquer energia, como vimos na figura (3.2) e tornaremos a discutir no capítulo 5.

Outro aspecto interessante, é que a escolha de N_{max} permite caracterizar bem a localização (para casos particulares isso foi verificado para cadeias com $N = 500$, $N = 1000$ e $N = 2000$), além de ter uma correlação com o tamanho da bolha de transcrição do DNA, na qual ocorre o rompimento das ligações de hidrogênio numa pequena região de 15 a 20 pares de bases [24]. Dito de outra forma, embora hajam divergências quanto aos valores dos parâmetros k , D e a , que caracterizam o modelo original de Peyrard Bishop, ou no valor de μ nas equações escritas nesse trabalho, o modelo permite obter localização de energia, que pode implicar uma desnaturação local, em regiões com um número de osciladores compatível com os dados disponíveis.

Uma idéia da dinâmica da cadeia pode ser vista no conjunto de instantâneos mostrados na figura (3.5). No instante $t = 10.5$ a energia está concentrada nos osciladores em torno do pulso inicial (102). Está começando a se formar uma perturbação com baixa energia. Esta perturbação pode vista formada no instante $t = 60.5$. Deve-se notar a simetria das energias em relação ao oscilador

(102). No instante $t = 280.5$ vê-se que as perturbações de baixa energia estão se aproximando dos extremos da cadeia. O último instante mostra que a energia permanece localizada no pulso central e que as energias permanecem distribuídas de forma simétrica. Do conjunto de instantes vê-se que o pulso central, que concentra essencialmente toda a energia, permanece parado. Para uma cadeia homogênea com $\mu_{A-T} = 2$ qualitativamente esse quadro se repete.

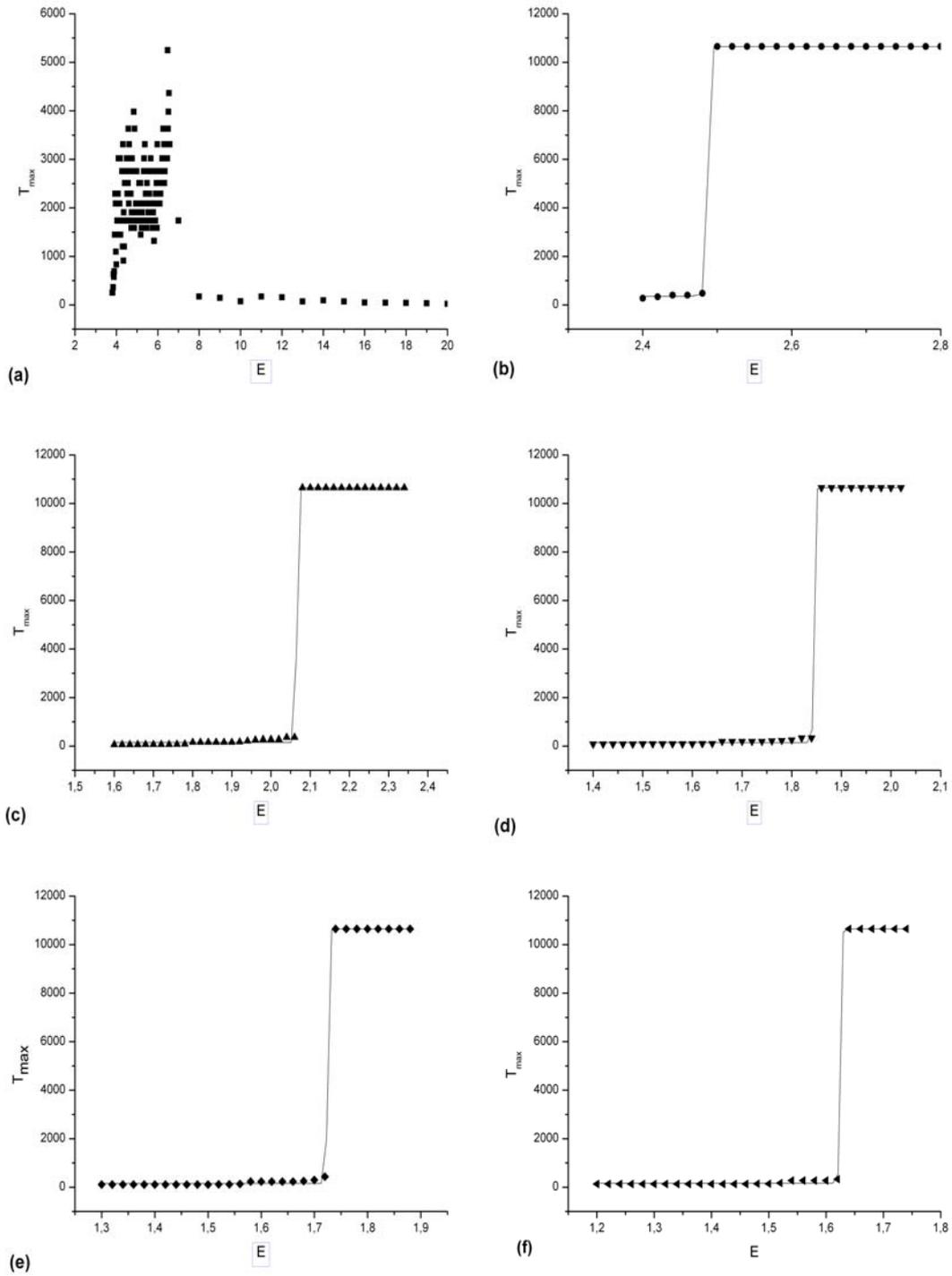


Figura 3.3: T_{max} como função da energia de excitação E com $N=200$, as unidades são arbitrárias com $\mu_{A-T} = 2$, (a) $\eta_L = 0.05$, (b) $\eta_L = 0.10$, (c) $\eta_L = 0.15$, (d) $\eta_L = 0.20$, (e) $\eta_L = 0.25$ e (f) $\eta_L = 0.30$ a linha é a interpolação dos pontos.

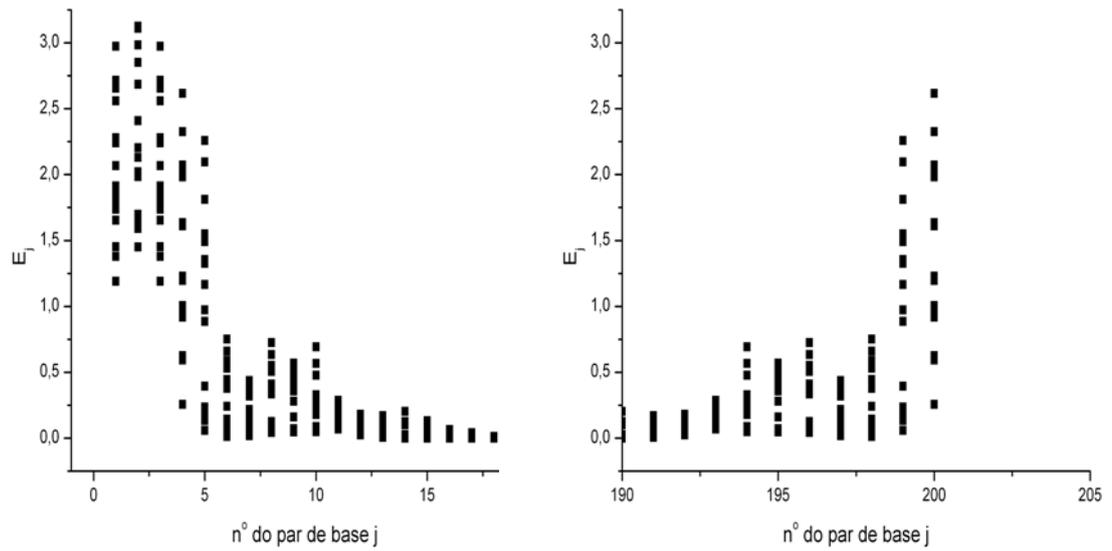


Figura 3.4: Energia do j -ésimo par de base contra o n^o do par de base. Para vários tempos de simulação

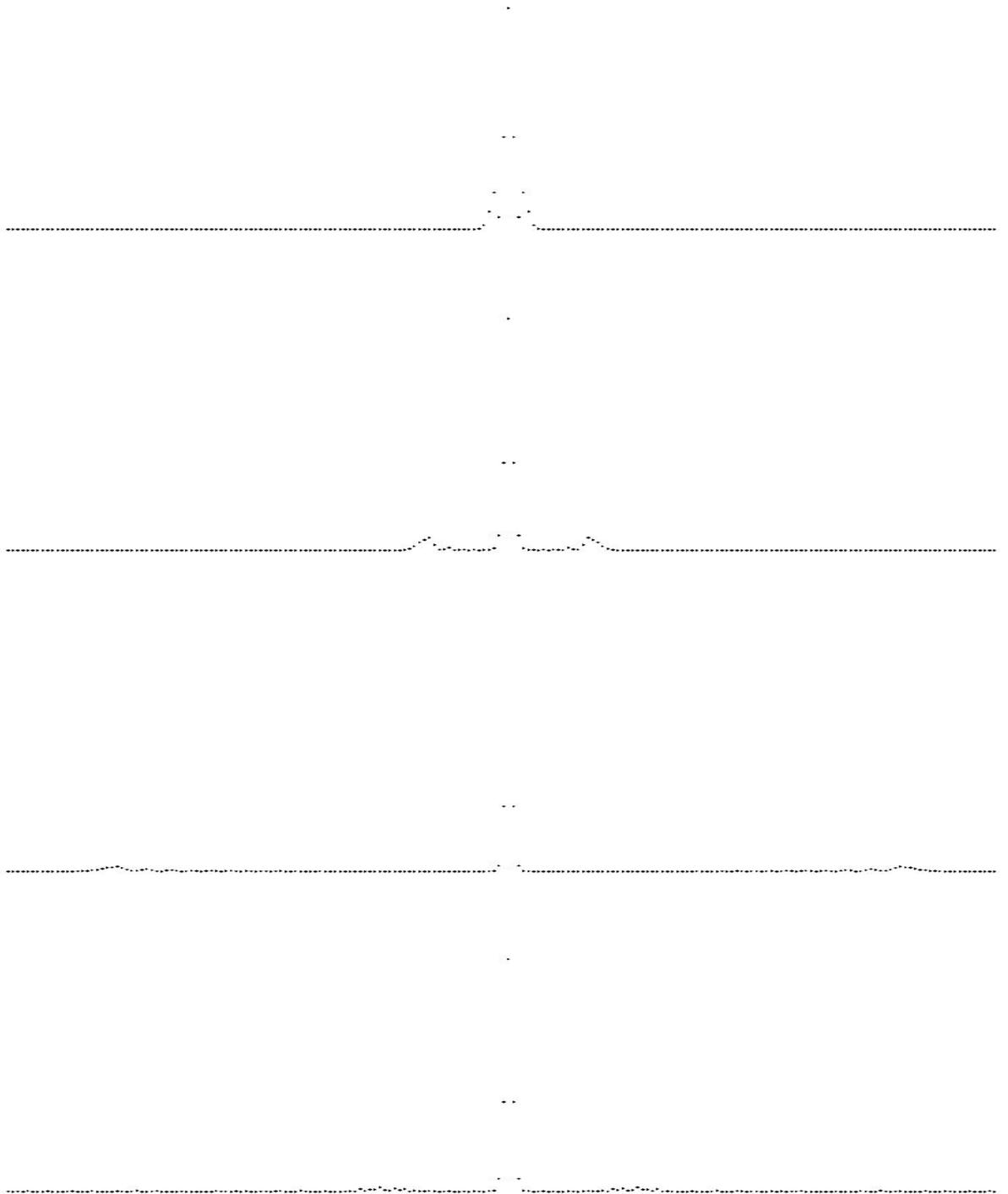


Figura 3.5: Instantâneos da distribuição de energia por oscilador na vertical compreende o intervalo $[-0.2, 2.5]$ cadeia homogênea com $N = 200$, $\mu_{C-G} = 3$, excitamos os osciladores 101, 102 e 103, energia total de excitação $E = 3.8$. Os instantes mostrados são de cima para baixo $t=10.5, 60.5, 280.5$ e 2060 .

Capítulo 4

Cadeias não-homogêneas

4.1 Preliminares

Nesta etapa, estamos interessados em estudar qual o efeito que não homogeneidades podem causar no limiar de energia para localização no modelo de Peyrard Bishop. As considerações com relação as simulações numéricas continuam válidas e, as mudanças no sistema serão discutidas abaixo. Coerente com os resultados do capítulo anterior, nas discussões abaixo, consideramos a rede com duzentos osciladores, $N = 200$, com condições periódicas de contorno e $\eta_L = 0.10$, ou seja, $N_{max} = 20$ osciladores.

A não homogeneidade será tratada na forma de blocos com o valor do parâmetro μ diferentes. De forma geral, a notação $\mu_{A-T}(k : l)$ ou $\mu_{(C-G)}(k : l)$ significa que do k -ésimo oscilador até o l -ésimo, o valor de μ é igual a dois ou três, respectivamente. Para cadeias com dois blocos, somente a notação se reduz a um segmento.

O outro segmento tem o valor de μ "complementar".

Como já mencionado anteriormente, teremos dois valores para μ dependendo do emparelhamento. Caso seja um emparelhamento (A-T) $\mu_{(A-T)} = 2$ e se for (C-G) $\mu_{(C-G)} = 3$. Trechos compostos com pares (A-T) serão denominados "macio" ($\mu_{(A-T)} = 2$) no que segue e trechos compostos de pares (C-G), de "rígidos" ($\mu_{(C-G)} = 3$).

4.2 Resultados

Na figura (4.1(a)) estão os resultados de T_{max} como função da energia de excitação para uma fita do tipo $\mu_{(C-G)}(1 : 100)$, significando que temos 50% de bases (C-G) e 50% de bases (A-T). Na fig. (4.1(a)) o pulso foi colocado no início da cadeia, ou seja, na parte mais rígida. Mas devido as condições de contorno periódicas, isso significa que o pulso está na interface "macio/rígido". Na figura (4.1(b)), por outro lado, o pulso inicial foi colocado nos osciladores 49, 50 e 51, ou seja, no meio da sub-cadeia "rígida".

Excitando-se o pulso no meio da sub-cadeia mais rígida obtém-se a energia crítica de $E_c = 1,62$. Este valor é próximo do valor da energia crítica de localização para uma cadeia homogênea do tipo (C-G), que é de 1,74. Observa-se da dinâmica, que o pulso se desloca ao longo da sub-cadeia, na qual foi colocado, refletindo nas interfaces. O comportamento da energia crítica é muito próximo aquele de uma cadeia homogênea com um número efetivo de osciladores menor

do que N , o que implica num η_L efetivo menor do que 0.1, ou num $N_{max} > 20$. Qualquer dessas mudanças tem como efeito diminuir a energia para localização.

Quando o pulso é excitado na interface entre dois blocos (isso é feito sempre de forma que o pulso inicial esteja em um único bloco), há uma mudança significativa no valor de E_c , ou seja, ele cai para $E_c = 0.72$, que é um valor menor que a metade do valor encontrado no caso anterior. A dinâmica revela, nesse caso, que o pulso também se move num único bloco, porém com uma velocidade maior. A variação no valor de E_c só pode ser atribuído à proximidade da interface no instante inicial, mas a razão para isso não é clara.

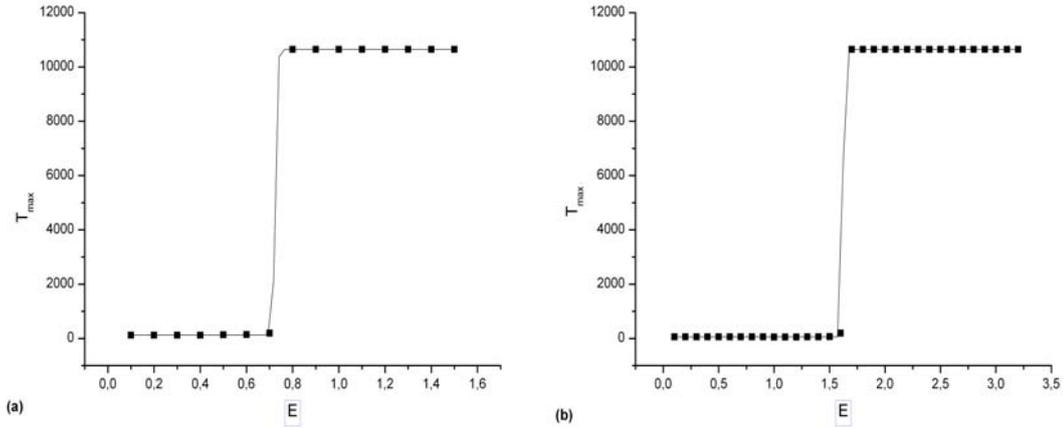


Figura 4.1: T_{max} como função da energia de excitação E com $N=200$, as unidades são arbitrárias, (a) $\mu_{C-G}(1 : 100) = 3$ com o pulso no início da cadeia, osciladores 1, 2 e 3, (b) mesmo caso da fig. (a) exceto pelo fato do pulso inicial estar nos osciladores 49, 50 e 51, a linha é a interpolação dos pontos.

Estes resultados sugerem que a localização dependa do tamanho das subcadeias, e da posição onde o pulso inicial é colocado. Então, enriqueceu-se o

sistema com várias interfaces "macio/rígido", e procurou-se um tamanho mínimo de uma porção para localização.

| Blocos | | Energia crítica por local de excitação | | | |
|--------|----------------------|--|-----------|------------|------------|
| Número | Tamanho por bloco | interface | interface | meio bloco | meio bloco |
| | | rígida | macia | rígido | macio |
| 1 | 200 | - | - | 1.75 | 2.49 |
| 2 | 100 | 0.72 | 0.69 | 1.62 | 2.29 |
| 4 | 50 | 0.68 | 0.66 | 1.58 | 2.04 |
| 5 | 40 | 0.60 | 0.63 | 1.59 | 2.02 |
| 7 | 30 | 0.53 | 0.53 | 1.51 | 1.91 |

Tabela 4.1: Energia crítica de localização em função do tamanho dos blocos e local da excitação inicial.

Na figura (4.2 (a) à (d)), faz-se os mesmos testes dos gráficos anteriores, mas para uma cadeia com quatro blocos, dois do tipo (C-G) e dois do tipo (A-T), distribuídos ao longo da cadeia da seguinte forma: $\mu_{C-G}(1 : 50)$, $\mu_{A-T}(51 : 100)$, $\mu_{C-G}(101 : 150)$ e $\mu_{A-T}(151 : 200)$. Para os dados da figura (4.2 (a)), o pulso inicial foi excitado nos osciladores 1, 2 e 3, na parte mais rígida da sub-cadeia. Para os dados da figura (4.2 (c)), o pulso foi excitado nos osciladores 51, 52 e 53, na parte mais macia da sub-cadeia. Os dois casos têm um aspecto comum, ambos os pulsos iniciais estão na interface "macio/rígido" e o resultado mostra que eles têm, praticamente, a mesma energia crítica de localização. Para o primeiro

caso (4.2 (a)), $E_c = 0.68$, enquanto para (4.2 (c)) tem-se $E_c = 0.66$. No caso da figura (4.2 (b)) o pulso inicial foi introduzido nos osciladores 24, 25 e 26, ou seja, no meio da sub-cadeia rígida e, a energia crítica encontrada foi $E_c = 1,58$, que é uma energia de localização próxima à obtida para uma cadeia homogênea do tipo (C-G) $E_c = 1,74$. Na figura (4.2 (d)) o pulso inicial foi introduzido nos osciladores 74, 75 e 76, no meio da sub-cadeia macia e a energia crítica encontrada foi $E_c = 2,04$, que é uma energia de localização não muito distinta daquela obtida para uma cadeia homogênea do tipo (A-T) $E_c = 2,48$. Os resultados dos testes com outras sub-divisões em blocos estão compilados na tabela (4.1). Observa-se que os valores de E_c são pouco sensíveis ao fato da excitação estar na parte macia ou rígida da interface. Há, contudo, um decréscimo suave no valor de E_c com a diminuição do tamanho do bloco. Essa mesma tendência é observada quando a excitação inicial é introduzida no meio de cada bloco, "longe" da interface, embora o valor de E_c dependa se o bloco excitado é macio ou rígido.

Assim como discutido para o caso de dois blocos, a dinâmica revela que a energia introduzida no sistema permanece no bloco de excitação. Isso significa que para blocos menores do que 20 osciladores, a localização satisfaz os critérios propostos para qualquer energia de excitação.

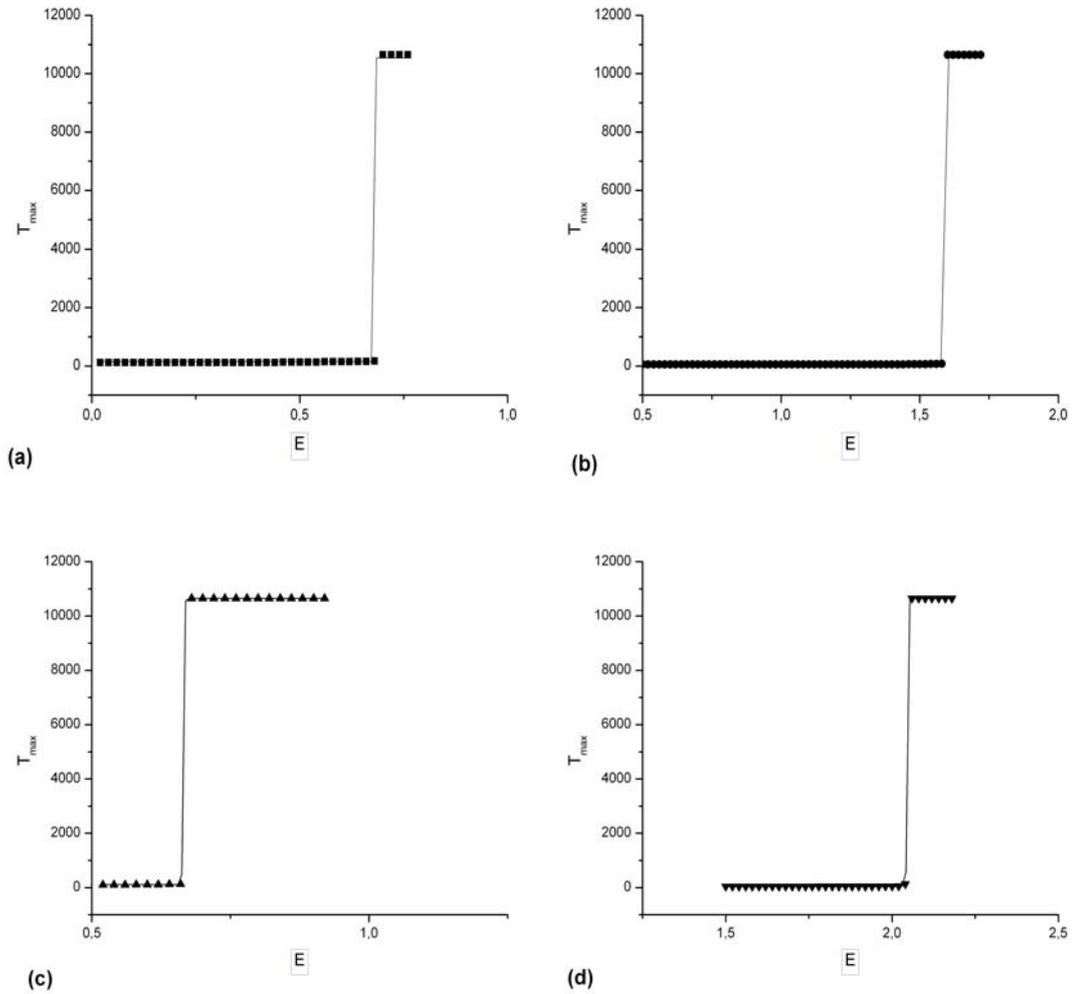


Figura 4.2: T_{max} como função da energia de excitação E com $N=200$, as unidades são arbitrárias, tem-se uma cadeia do tipo $\mu_{C-G}(1 : 50)$, $\mu_{A-T}(51 : 100)$, $\mu_{C-G}(101 : 150)$ e $\mu_{A-T}(151 : 200)$. Simula-se a mesma cadeia para quatro regiões de excitação inicial distintas: (a) pulso inicial nos osciladores 1, 2 e 3, (b) pulso inicial nos osciladores 24, 25 e 26, (c) pulso inicial nos osciladores 51, 52 e 53, (d) pulso inicial nos osciladores 74, 75 e 76. A linha é a interpolação dos pontos.

4.3 Estudo da dinâmica da cadeia de osciladores não-homogêneos

A idéia da dinâmica da cadeia pode ser vista no conjunto de instantâneos mostrados na figura (4.3). Da primeira figura ($t=10$), vê-se que o pulso inicial foi modificado com uma parte pequena da energia tendo passado aos vizinhos mais próximos. No tempo ($t=60$), percebe-se que a energia está concentrada no pulso inicial, distribuída simetricamente em relação ao oscilador central (50) e que um pulso de baixa energia se formou e está se propagando para as extremidades do bloco excitado. No tempo $t = 330$ os pequenos pulsos já refletiram nas interfaces e voltaram para o centro do bloco, onde está localizada a energia. Porém, perde-se a simetria e isso fará com que o pulso central se desloque pela cadeia, como se pode ver no último flash ($t=2680$), muito lentamente. Assim a introdução de interfaces "macio/rígido" produz o movimento do pulso que contém a energia, além de diminuir a energia crítica necessária para localização de energia, como visto acima.

Diferente da figura (4.3), a excitação, nos instantâneos mostrados na figura (4.4), está na interface do bloco (101, 102 e 103). Aqui, no instante $t = 20$, parte da energia cedida pelo pulso inicial já foi refletida na interface (100/101) e coloca o pulso central em movimento com velocidade significativamente maior do que no caso anterior analisado. Em $t=540$ o pulso central já sofreu uma reflexão na interface (200/1). O último instantâneo ($t=2750$) só confirma que a energia

continua localizada e que, apesar da excitação na interface, os osciladores de 1 a 100 pouco sentem (do ponto de vista energético) a presença da outra metade do sistema.

Na figura (4.5) mostramos o deslocamento da posição de equilíbrio ao invés da energia. Note que os osciladores da cadeia rígida (pulso inicial está na maioria) são colocados, como esperado, em movimento também, porém, com deslocamentos pequenos, ou seja, com energia muito pequena relativamente aquela que permanece no pulso central.

O conjunto de instantâneos da figura (4.6), mostram que qualitativamente o comportamento é o mesmo dos discutidos nos casos anteriores. A energia permanece localizada num pulso contendo de 7 a 9 osciladores, que se desloca pela rede restrito ao bloco em que foi excitado inicialmente. Vê-se que apenas parte muito pequena da energia atravessa a interface entre dois blocos.



Figura 4.3: Instantâneos da distribuição de energia por oscilador. O intervalo da escala vertical (energia por oscilador) é $[-0.1, 2.5]$. Cadeia homogênea com dois blocos, $\mu_{C-G}(1 : 100) = 3$ e $\mu_{A-T}(101 : 200) = 2$. A energia de excitação é 3.0 e os osciladores excitados juntamente com as condições iniciais utilizadas foram: $\xi_{49} = \xi_{51} = -0.1483$, $\xi_{50} = 0$, $\dot{\xi}_{49} = \dot{\xi}_{51} = 0.866$ e $\dot{\xi}_{50} = 1.988$. os instantes mostrados de cima para baixo são: $t=10.5, 60, 330$ e 2680 .



Figura 4.4: Instantâneos da distribuição de energia por oscilador. O intervalo da escala vertical (energia por oscilador) é $[-0.1, 2.5]$. Cadeia com dois blocos, $\mu_{C-G}(1 : 100) = 3$ e $\mu_{A-T}(101 : 200) = 2$. A energia de excitação é 3.0 e os osciladores excitados juntamente com as condições iniciais utilizadas foram: $\xi_{101} = \xi_{103} = -0.2844$, $\xi_{102} = 0$, $\dot{\xi}_{101} = \dot{\xi}_{103} = 1.299$ e $\dot{\xi}_{102} = 1.198$. Os instantes mostrados de cima para baixo são: $t=0.5, 20, 540$ e 2750 .

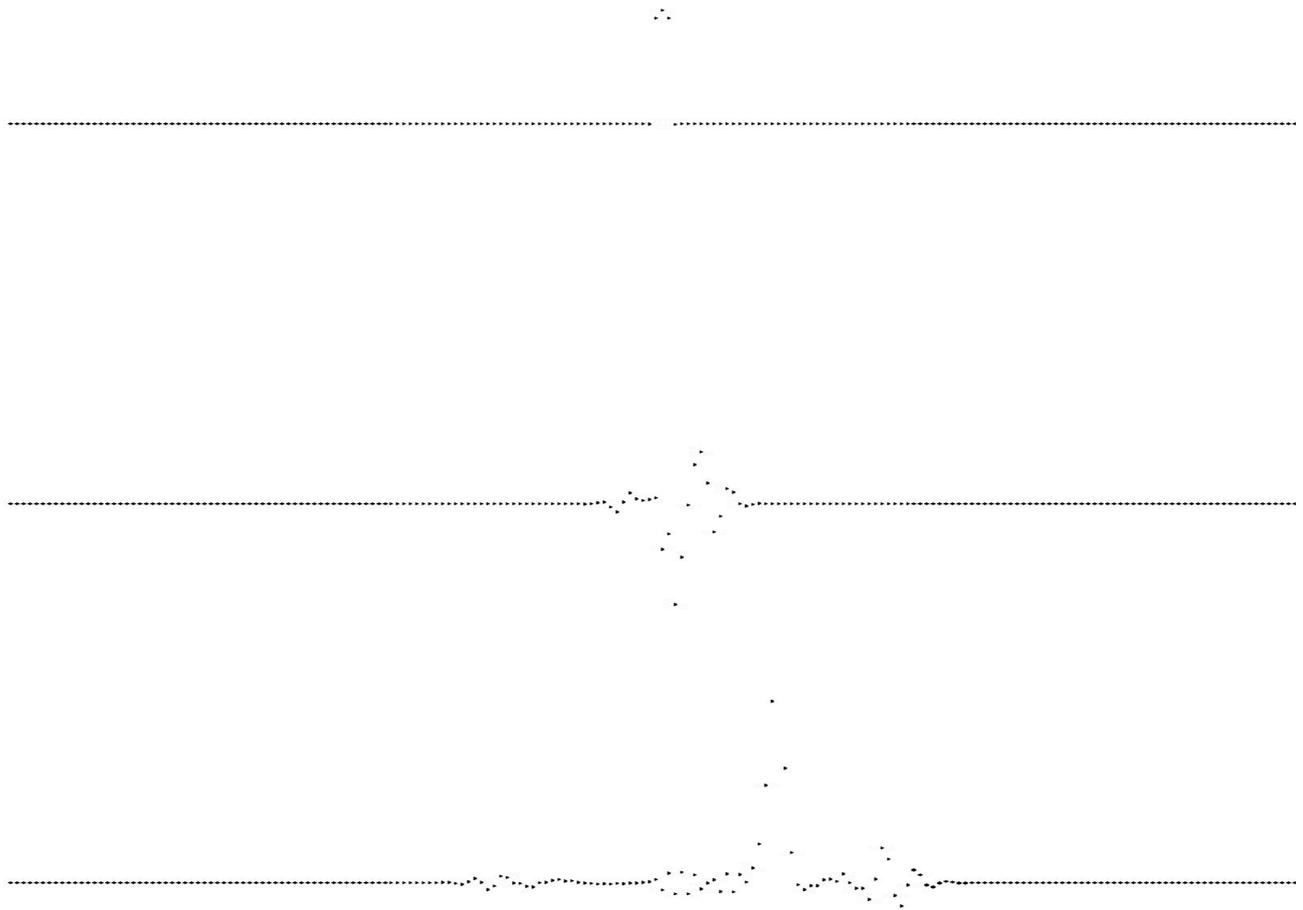


Figura 4.5: Instantâneos do deslocamento ξ_i compreende o intervalo da escala vertical $[-0.8, 0.8]$. As condições são as mesmas da figura (4.3). Os instantes mostrados de cima para baixo são: $t=0.5$, 20 e 90.

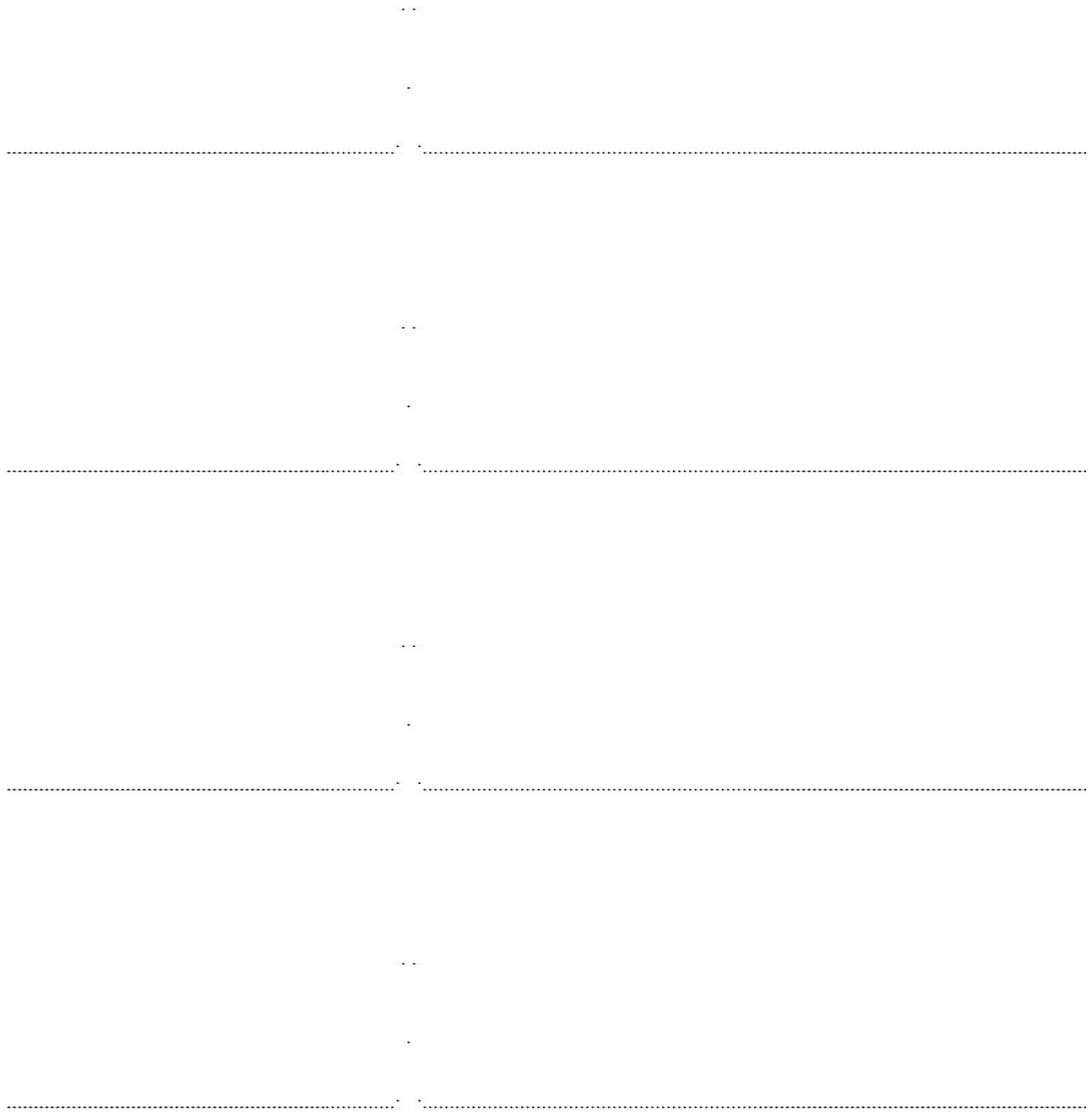


Figura 4.6: Instantâneos da distribuição de energia por oscilador. O intervalo da escala vertical (energia por oscilador) é $[0.1, 2.5]$. Cadeia com quatro blocos, $\mu_{C-G}(1 : 50) = 3$, $\mu_{A-T}(51 : 100) = 2$, $\mu_{C-G}(101 : 150) = 3$, $\mu_{A-T}(151 : 200) = 2$. A energia de excitação é 3.0 e os osciladores excitados juntamente com as condições iniciais utilizadas foram: $\xi_{74} = \xi_{76} = -0.2844$, $\xi_{75} = 0$, $\dot{\xi}_{74} = \dot{\xi}_{76} = 1.299$ e $\dot{\xi}_{75} = 1.1982$. Os instantes mostrados de cima para baixo são: $t=10, 40, 140$ e 2100 .

Capítulo 5

Deslocalização de energia

5.1 Preliminares

Verificou-se qualitativamente que, diferentemente do esperado, o sistema estudado ao invés de apresentar duas regiões distintas de energia, apresenta três regiões para valores de energia crítica de localização. A primeira região inicialmente é uma região de baixa energia onde não existe localização, em seguida tem-se uma região na qual ocorre a localização e posteriormente uma região onde esta localização é perdida. Isso foi verificado em cadeias homogêneas e não homogêneas e será discutido abaixo.

5.2 Cadeias homogêneas

Na figura (5.1) apresentamos o gráfico de T_{max} como função da energia de excitação, E , para uma cadeia homogênea do tipo (A-T). Deste gráfico, observar-se que para $E = 4.75$ o critério estabelecido para um estado localizado não é atendido. Dessa forma, para valores da energia de excitação, $E \geq 4.67 \pm 0.13$, o sistema perde a localização. Destaca-se também a região $5 < E_c < 17.5$, onde o tempo para energia espalhar pela rede é maior que N . Nessa região acredita-se que o potencial de Morse, para esta faixa de energia, ainda influencia significativamente, o sistema. A partir de $E_c > 20$ não se encontram valores de energia que satisfaçam os critérios para localização e, o curioso é que o tempo para energia se espalhar por mais de vinte osciladores é da ordem de 50 unidades de tempo. Para este caso testaram-se valores de energia até 120 unidades de energia.

O potencial de Morse para grandes amplitudes torna-se constante como discutido na seção (2.1) e, pode ser considerado como uma constante na hamiltoniana (2.8). No exemplo anterior para valores de energia acima de 20, e com a hipótese de que para estes valores de energia o sistema oscila com grandes amplitudes, tem-se uma cadeia de osciladores harmônica e homogênea, justificando a perda de localização.

Para verificar esta hipótese, foi feito a Transformada de Fourier do sistema. O resultado é apresentado na forma de gráficos do módulo da transformada complexa de Fourier $F(\omega)$ da coordenada do oscilador central do pulso inicial em

função da frequência ω . No conjunto de figuras (5.2(a)-(d)), mostramos os resultados obtidos para $\mu = \mu_{A-T} = 2$. A banda dispersiva é dada por $\mu \leq \omega \leq \sqrt{4 + \mu^2}$ ou, nesse caso, $2 \leq \omega \leq 2.83$. Na figura (5.2(a)) a energia de excitação é $E = 1.0$, menor que a energia crítica para localização ($E_c = 2.48$), como discutido no capítulo 3. As frequências com intensidade significativa estão todas na banda dispersiva. Na figura (5.2(b) e (c)), para $E = 2.8$ e $E = 3.8$, respectivamente, neste dois casos a energia está na faixa onde existe localização e, este fato é refletido nos resultados observados nos gráficos, pois, para as duas figuras observa-se dois picos de frequência que estão fora da banda de dispersão. Tanto na figura (b) quanto na (c) os picos estão próximos a $\omega = 1.8$ e $\omega = 3.75$. Na figura (5.2(d)) a energia de excitação foi $E = 27.3$, valor de energia para o qual não se observa localização e o tempo necessário para que a energia se espalhe pela rede é menor que N . A faixa com valores significativos de frequências estão entre $0 \leq \omega \leq 2.0$. Para uma cadeia puramente harmônica, as frequências dos modos normais de vibração são:

$$\omega(l) = \sqrt{4 \sin^2(l/2)}, \quad (5.1)$$

onde $l = (j\pi/N + 1)$ com $j = 1, \dots, N$.

A perda de localização também foi observada para cadeias não homogêneas, mas com os testes realizados até o momento, não é possível caracterizá-la. Na figura (5.3 (a) e (b)) mostramos os gráficos de T_{max} como função da energia de excitação para uma cadeia composta de dois blocos com $\mu_{C-G}(1 : 100)$ e $\mu_{A-T}(101 : 200)$. Na figura (5.3(a)) a excitação inicial foi colocada na interface

"macio/rígido", nos osciladores 1, 2 e 3. O valor da energia para o qual se perde a localização é 3.85, valor próximo ao encontrado para a cadeia homogênea com $\mu_{A-T} = 2$. Na figura (5.3(b)), o pulso de excitação foi dado no meio da subcadeia rígida, nos osciladores 49, 50 e 51. a energia de deslocalização obtida é 17.8, praticamente a mesma de uma cadeia homogênea com $\mu_{C-G} = 3$. Estes resultados mostram que repete-se para a energia crítica de deslocalização o comportamento discutido no capítulo 4 para energia crítica de localização, ou seja, para excitação no meio do bloco a energia pouco sente a existência de interfaces. Para o pulso inicial da interface a energia crítica de localização ou deslocalização é muito afetada, com um resultado que é menor do que o observado para uma cadeia homogênea macia.

Na figura (5.4) mostramos, para uma única localização do pulso inicial, a curva de $T_{max} \times E$, para uma cadeia com 4 blocos ($\mu_{C-G}(1 : 50)$, $\mu_{A-T}(51 : 100)$, $\mu_{C-G}(101 : 150)$ e $\mu_{A-T}(151 : 200)$). A excitação foi colocada nos osciladores 1, 2 e 3. A curva mostrada evidência a ocorrência da perda da localização por energias de excitação maiores que $E \sim 4.23$.

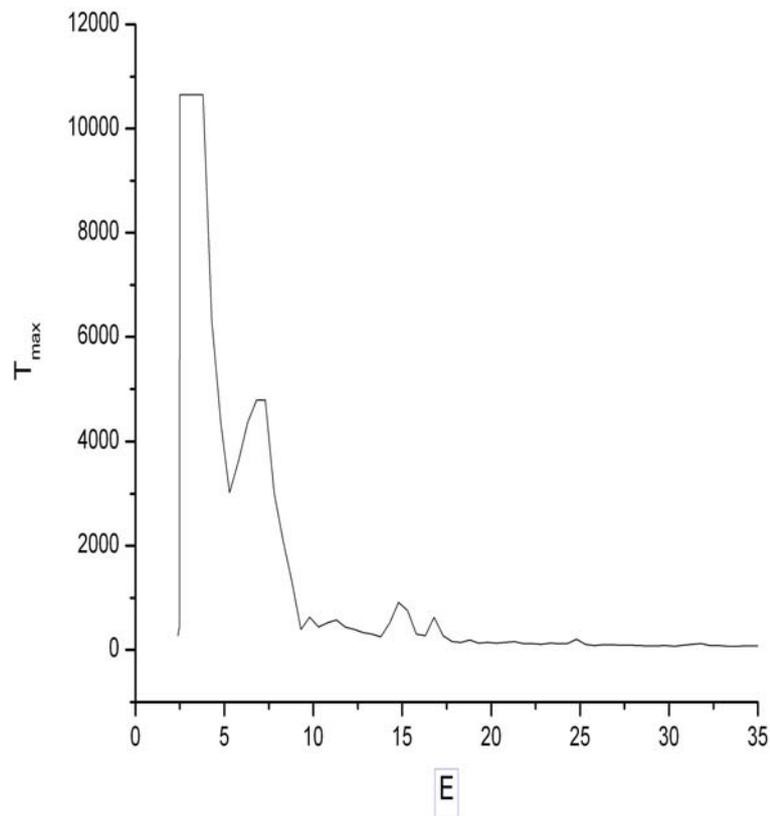


Figura 5.1: T_{max} como função da energia com $N=200$, as unidades são arbitrárias, para uma cadeia homogênea do tipo $\mu_{A-T}(1 : 200)$, o pulso inicial foi dado nos osciladores 1, 2 e 3.

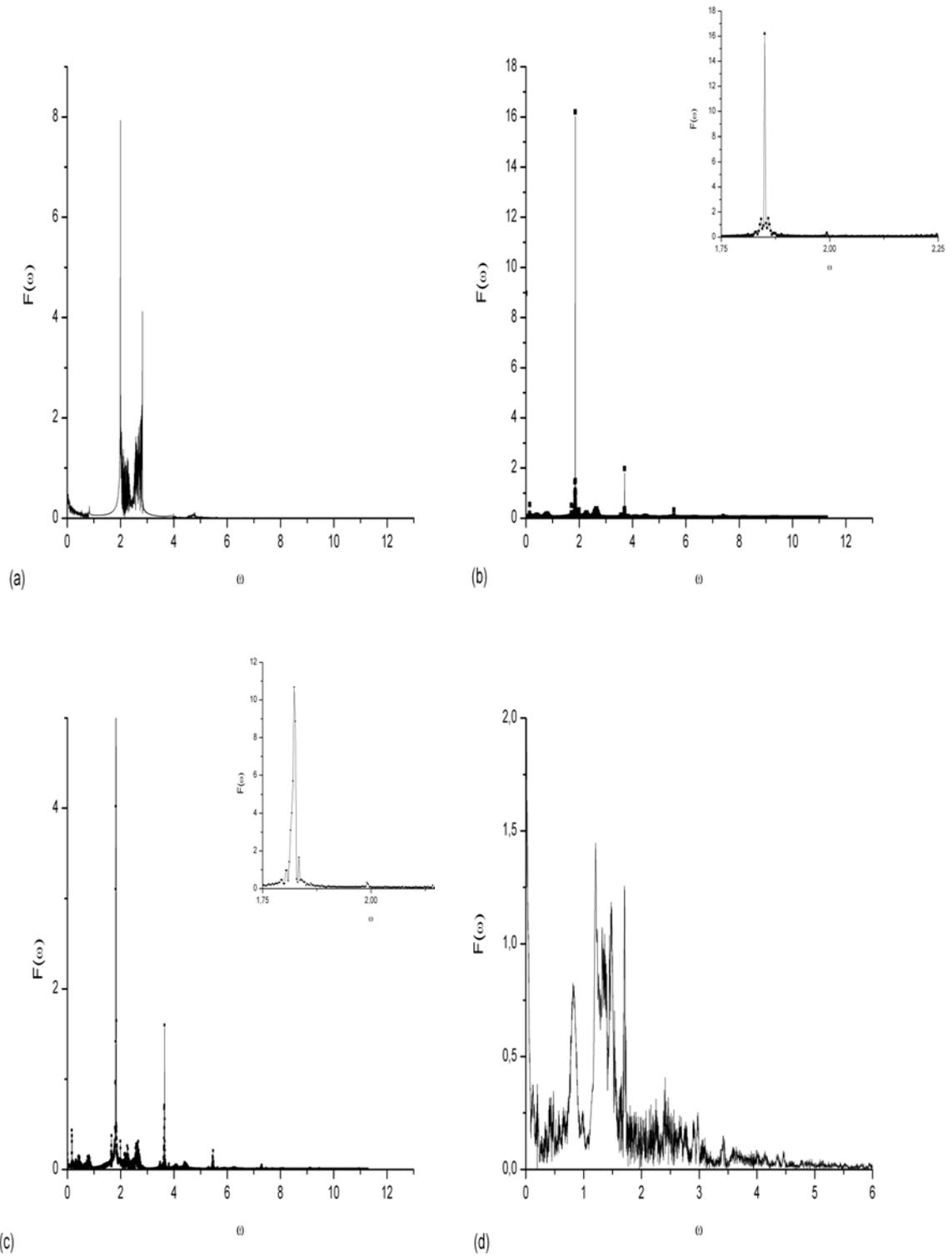


Figura 5.2: Transformada de fourier com o pulso inicial dado nos osciladores 1, 2 e 3. A cadeia é homogênea com $\mu = 2$ para (a) $E=1.0$, (b) $E=2.8$, (c) $E=3.8$ e (d) $E=27.3$. É plotado o módulo $F(\omega)$ da transformada complexa de Fourier.

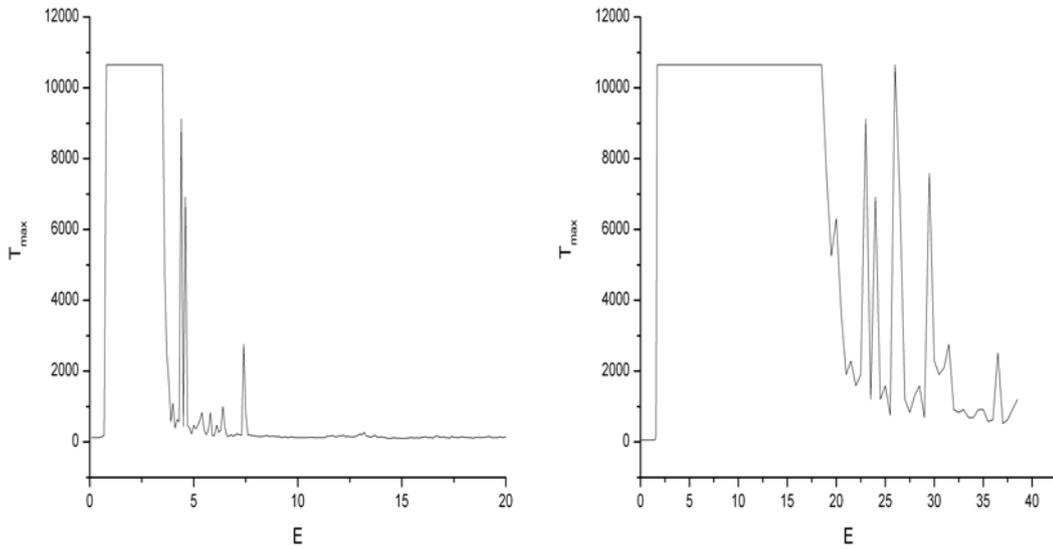


Figura 5.3: T_{max} como função da energia de excitação E com $N=200$, as unidades são arbitrárias, (a) $\mu_{C-G}(1 : 100) = 3$ com o pulso no início da cadeia, osciladores 1, 2 e 3, (b) mesmo caso da fig. (a) exceto pelo fato do pulso inicial estar nos osciladores 49, 50 e 51.

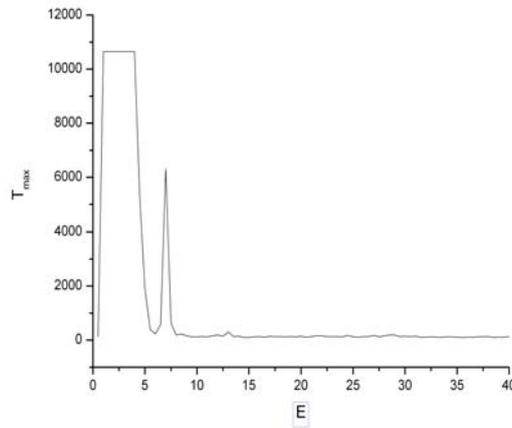


Figura 5.4: T_{max} como função da energia de excitação E com $N=200$, as unidades são arbitrárias, para uma cadeia do tipo $\mu_{C-G}(1 : 50)$, $\mu_{A-T}(51 : 100)$, $\mu_{C-G}(101 : 150)$, $\mu_{A-T}(151 : 200)$, pulso inicial nos osciladores 1, 2 e 3.

Capítulo 6

Conclusões

O objetivo deste trabalho era estudar os efeitos das não homogeneidades em cadeias de osciladores tratado no modelo de Peyrard-Bishop. Como proposto, essas não homogeneidades seriam introduzidas por blocos.

Começamos analisando o problema de cadeias homogêneas, discutindo o comportamento do limiar de energia, a energia crítica E_c , para a localização em função de: 1) número de osciladores da rede; 2) do parâmetro η_L que estabelece um dos critérios para localização e 3) do parâmetro μ que é uma medida dos acoplamentos dos osciladores na rede. Destas análises concluimos que $N = 200$ é um tamanho de rede suficiente para discutir o problema de localização, que $\eta_L = 0.1$ ($N_{max} = 20$, no caso de cadeia com 200 osciladores) é um bom valor, porque pode-se correlacionar a localização de energia com as bolhas de transcrição no DNA reportadas na literatura e que os parâmetros $\mu_{A-T} = 2$ e $\mu_{C-G} = 3$, obtidos ($\mu = \frac{4Da^2}{k}$) de valores encontrados na literatura, apesar das ressalvas feitas no

corpo do trabalho, são suficientes para caracterizar a localização por meio de E_c .

Para $\mu > 15$ o valor de E_c é, praticamente, constante com μ .

Dos resultados da dinâmica da cadeia homogênea, vimos que a energia permanece concentrada em torno do pulso inicial, com uma pequena perturbação, com baixa energia, que se move ao longo da cadeia e se espalha pela rede simetricamente em relação ao pulso central que permanece estático.

As análises das cadeias não homogêneas a blocos foram concentradas no número de osciladores por bloco e na posição de excitação do pulso inicial. Os resultados obtidos mostram que a energia crítica é dependente da existência das interfaces e, caso particular da excitação próxima à interface, essa dependência é muito forte, independente do tamanho dos blocos. Da análise da dinâmica mostramos que a energia introduzida no pulso inicial permanece, na sua quase totalidade, sempre no mesmo bloco embora o pulso de localização se desloque, ao contrário do caso homogêneo, de um lado para o outro no bloco, refletindo nas interfaces. A razão para esse "aprisionamento" do pulso no bloco não nos é clara.

Por fim, mostramos que se a energia de excitação cresce além de um certo limite, uma "energia de deslocalização", a localização é perdida. Tal efeito pode ser relacionado ao aspecto não confinante do potencial de Morse para valores positivos (à direita do mínimo) do deslocamento. Nesse caso, a cadeia ou um conjunto de osciladores, deixariam de sentir a não linearidade e espalhariam a energia para toda a rede. Um estudo mais aprofundado, contudo, se faz necessário.

Em conclusão, os resultados apresentados indicam que a presença das in-

terfaces, a não homogeneidade, facilita a localização e também a deslocalização (veja valores de E_c e da energia de deslocalização), implicando num intervalo de energias em que é possível obter e manter a localização. Além disso, limitar N_{max} em 20 osciladores mostra que o modelo pode ser correlacionado com dados experimentais. Das dinâmicas, vê-se que o número médio de osciladores do pulso central é de 7 a 9, praticamente do tamanho do TATAbox do DNA. O modelo é, dessa forma, um bom candidato para a descrição e compreensão da dinâmica do DNA.

Apêndice A

Transformação da lagrangiana L_{PB} em termos dos modos Normais

A Lagrangiana para o modelo da macromolécula DNA, proposto por Peyrard-Bishop (PB) é escrita como

$$L_{PB} = \sum_{j=1}^N \frac{m}{2} (\dot{u}_j^2 + \dot{v}_j^2) - \frac{k}{2} (u_{j+1} - u_j)^2 - \frac{k}{2} (v_{j+1} - v_j)^2 - D [\exp a(v_j - u_j) - 1]^2, \quad (\text{A.1})$$

Podemos escrever a Lagrangiana (A.1) em termos dos modos normais, definidos

$$x_j = \frac{(u_j + v_j)}{\sqrt{2}} \quad y_j = \frac{(u_j - v_j)}{\sqrt{2}}. \quad (\text{A.2})$$

Logo,

$$u_j = \frac{\sqrt{2}}{2} (x_j + y_j), \quad v_j = \frac{\sqrt{2}}{2} (x_j - y_j) \quad \text{e} \quad (v_j - u_j) = -\sqrt{2} y_j \quad (\text{A.3})$$

e

$$\dot{u}_j = \frac{\sqrt{2}}{2}(\dot{x}_j + \dot{y}_j) \quad \dot{v}_j = \frac{\sqrt{2}}{2}(\dot{x}_j - \dot{y}_j). \quad (\text{A.4})$$

Calculando;

$$\dot{u}_j^2 + \dot{v}_j^2 = \dot{x}_j^2 + \dot{y}_j^2, \quad (\text{A.5})$$

$$(u_{j+1} - u_j)^2 = \frac{1}{2}[x_{j+1}^2 + y_{j+1}^2 - 2(x_{j+1}x_j + y_{j+1}y_j) + x_j^2 + y_j^2], \quad (\text{A.6})$$

$$(v_{j+1} - v_j)^2 = \frac{1}{2}[x_{j+1}^2 + y_{j+1}^2 - 2(x_{j+1}x_j + y_{j+1}y_j) + x_j^2 + y_j^2]. \quad (\text{A.7})$$

Somando (A.6) com (A.7) obtemos:

$$\begin{aligned} (u_{j+1} - u_j)^2 + (v_{j+1} - v_j)^2 &= x_{j+1}^2 + y_{j+1}^2 - 2(x_{j+1}x_j + y_{j+1}y_j) + x_j^2 + y_j^2, \\ &= (x_{j+1} - x_j)^2 + (y_{j+1} - y_j)^2. \end{aligned} \quad (\text{A.8})$$

Assim temos

$$L_{PB} = \sum_{j=1}^N \frac{m}{2}(\dot{x}_j^2 + \dot{y}_j^2) - \frac{k}{2}(x_{j+1} - x_j)^2 - \frac{k}{2}(y_{j+1} - y_j)^2 - D[\exp(-\sqrt{2}ay_j) - 1]^2, \quad (\text{A.9})$$

$$L_X = \sum_{j=1}^N \frac{m}{2}\dot{x}_j^2 - \frac{k}{2}(x_{j+1} - x_j)^2, \quad (\text{A.10})$$

que depende apenas da coordenada x , e L_Y , que depende somente da coordenada y , como

$$L_Y = \sum_{j=1}^N \frac{m}{2}\dot{y}_j^2 - \frac{k}{2}(y_{j+1} - y_j)^2 - D[\exp(-\sqrt{2}ay_j) - 1]^2. \quad (\text{A.11})$$

Apêndice B

Adimensionamento da Lagrangiana

L_Y

Admitindo a Lagrangiana L_Y podemos sugerir uma mudança de variável nas coordenadas $\xi_j = \sqrt{2}ay_j$ e $d\xi_j = \sqrt{2}ady_j$. Uma mudança no tempo do tipo $\tau \equiv \sqrt{\frac{k}{m}}t$ isto implica que $d\tau \equiv \sqrt{\frac{k}{m}}dt$, e com estas mudanças:

$$\dot{y}_j = \frac{dy_j}{dt} = \frac{1}{\sqrt{2}a} \sqrt{\frac{k}{m}} \frac{d\xi_j}{d\tau}. \quad (\text{B.1})$$

O ponto denota, agora, a derivada com relação a τ . Reescrevendo L_Y com este adimensionamento temos:

$$L_Y = \sum_{j=1}^N \frac{m}{2} \left(\frac{1}{\sqrt{2}a} \sqrt{\frac{k}{m}} \dot{\xi}_j \right)^2 - \frac{k}{2} \left(\frac{1}{\sqrt{2}a} (\xi_{j+1} - \xi_j) \right)^2 - D \left[\exp\left(-\sqrt{2}a \frac{\xi_j}{\sqrt{2}a}\right) - 1 \right]^2. \quad (\text{B.2})$$

Logo,

$$L_Y = \frac{k}{2a^2} \sum_{j=1}^N \frac{1}{2} \dot{\xi}_j^2 - \frac{1}{2} (\xi_{j+1} - \xi_j)^2 - \frac{1}{2} [\exp(-\xi_j) - 1]^2 \quad \text{onde} \quad \mu^2 = \frac{4Da^2}{k}. \quad (\text{B.3})$$

Assim temos,

$$L = \sum_{j=1}^N \frac{1}{2} \dot{\xi}_j^2 - \frac{1}{2} (\xi_{j+1} - \xi_j)^2 - \frac{\mu^2}{2} [\exp(-\xi_j) - 1]^2 \quad (\text{B.4})$$

e

$$L_Y \equiv \frac{kL}{2a^2} \quad (\text{B.5})$$

Apêndice C

Espectro de frequência dos modos normais

As equações de movimento para o sistema descrito pela lagrangiana (2.6) são:

$$\ddot{\xi}_j - \xi_{j+1} + 2\xi_j - \xi_{j-1} + \mu^2[\exp(-\xi_j) - 1] \exp(-\xi_j) = 0 \quad i = 0, 1, 2, \dots, N+1. \quad (\text{C.1})$$

Para pequenas amplitudes, faz-se uma expansão em série de Taylor da parte não linear e considerando apenas os termos de segunda ordem assim é possível obter uma equação linearizada para o sistema na seguinte forma:

$$\ddot{\xi}_j - \xi_{j+1} + (2 + \mu^2)\xi_j - \xi_{j-1} = 0 \quad i = 0, 1, 2, \dots, N+1. \quad (\text{C.2})$$

Considerando por simplicidade que quando o sistema está em repouso o comprimento livre de cada mola seja a , ou seja a j -ésima partícula está à distância ja de um sistema fixo à parede rígida em $x_0 = 0$, e considerando uma dependência

harmônica no tempo para cada $\xi_j(t)$, ou seja

$$\xi_j(t) = \xi_j \cos \omega t, \quad (\text{C.3})$$

a equação para a j -ésima partícula se escreve

$$-\xi_j \omega^2 = \xi_{j-1} - (2 + \mu^2) \xi_j + \xi_{j+1}. \quad (\text{C.4})$$

Vamos assumir que ξ_j depende senoidalmente da distância ja :

$$\xi_j = C \sin qja, \quad (\text{C.5})$$

onde C é uma constante a ser determinada. Se introduzirmos esta condição na equação C.4, teremos:

$$-\omega^2 C \sin qja = C \sin q(j-1)a - (2 + \mu^2) C \sin qja + C \sin q(j+1)a \quad (\text{C.6})$$

ou, simplificando

$$\omega^2 = [(2 + \mu^2) - 2 \cos qa]. \quad (\text{C.7})$$

Para calcular o valor de q , faz-se uso das condições de contorno que nos dizem que $\xi_0 = \xi_{n+1} \equiv 0$ então:

$$\xi_0 = C \sin q0a = 0; \quad \xi_{n+1} = C \sin q(n+1)a = 0, \quad \text{então} \quad (\text{C.8})$$

$$q_l = q = \frac{l\pi}{(n+1)a} \quad l = 1, 2, \dots, n \quad (\text{C.9})$$

é possível escrever $\cos q_l a$ como sendo $\cos \left(\frac{q_l a}{2} + \frac{q_l a}{2} \right) = \cos^2 \frac{q_l a}{2} - \sin^2 \frac{q_l a}{2}$ então $\cos q_l a = 1 - 2 \sin^2 \frac{q_l a}{2}$, com isso

$$\omega_l = \sqrt{\mu^2 + 4 \sin^2 \left(\frac{k}{2} \right)}, \quad (\text{C.10})$$

onde $k = \frac{l\pi}{(n+1)}$.

Referências Bibliográficas

- [1] WATSON, J. D., CRICK, F. H. C. A structure for deoxyribose nucleic acid. *Nature*, vol.171, p.737-38, 1953.
- [2] PEYRARD, M., BISHOP, A. R. Statistical mechanics of a nonlinear model for DNA denaturation. *Phys. Rev. Lett.*, vol.62, p.2755-58, 1989.
- [3] PEYRARD, M. Nonlinear dynamics and statistical physics of DNA. *Nonlinearity*, vol.17, p.R1-R40, 2004.
- [4] YAKUSHEVICH, L.V. **Nonlinear Physics of DNA**. New York: John Wiley & Sons, 1998.
- [5] DRIGO, E., RUGGIERO, J.R. Parameters describing H-bond in DNA. *Phys. Rev. A*, vol.44, p.8435-8436, 1991.
- [6] DRIGO, E., RUGGIERO, J.R. H-bond simulation in DNA using a harmonic oscillator isospectral potential. *Phys. Rev. E*, vol.56, p.4486-4488, 1997.
- [7] PILZ, A.P.S. **Modelo estatístico para a molécula de DNA**. São José do Rio Preto: Tese Doutorado, 2003.
- [8] FERREIRA S.E. **Modelo dinâmico para a molécula de DNA**. São José do Rio Preto: Dissertação de Mestrado, 2003.

- [9] LUCA, J.De L.,DRIGO, E.F., PONNO, A. AND RUGGIERO, J.R. Energy localization in the Peyrard-Bishop DNA model. *Phys. Rev. E*, vol.70, p.026213-1-026213-9, 2004.
- [10] ZDRAVKOVIC, S. AND STARIC, M. V. The impact of viscosity on the DNA dynamics. *Physica Scripta*, vol.64, p.612-19, 2001.
- [11] DAUXOIS, T. Dynamics of breathers modes in a nonlinear "helicoidal" model of DNA. *Phys. Lett. A*, vol.159, p.390-95, 1991.
- [12] KUDRISTSKAYA, Z. G., DANILOV, V. I. Quantum mechanics study of bases interactions in various associates in atomic dipole approximation. *Theor. Biol.*, vol.59, p.303-18, 1976.
- [13] FERMI, E., PASTA J. and ULAM S., in *Collected Pappers of E. FERMI. University of Chicago*, 1965.
- [14] DAUXOIS T., PEYRARD M., RUFFO S., The Fermi-Pasta-Ulam 'numerical experiment': history and pedagogical perspectives. *European Journal of Physics*, vol.26, p.S3-S11, 2005.
- [15] METROPOLIS, N., The age of computing - A personal Memoir. *Daedalus*, vol.121, p.119-130, 1992.
- [16] TSITOURAS, CH. A tenth order symplectic Runge-Kutta-Nystrom method. *Celest. Mech. Din. Astron.*, vol.74, p.223-230, 1999.
- [17] CHANNELL, P.J. AND SCOVEL J.C. Symplectic integration of Hamiltonian systems. *Nonlinearity*, vol.3, p.231-259, 1990.

-
- [18] SANZ-SERNA, J.M. AND CALVO M.P. **Numerical Hamiltonian Problems**. London: Chapman and Hall, 1994.
- [19] MIRNOV, V.V., LICHTENBERG A.J. AND GUCLU H. Chaotic breather formation, coalescence, and evolution to energy equipartition in an oscillatory chain. *Physica D*, vol.157, p.251-282, 2001.
- [20] SHANNON, E.C., Mathematical theory of communication. *Mobile Computing and Communications Review*, vol.05, p.3-55, 2001.
- [21] CAMPA, A. Bubble propagation in helicoidal molecular chain. *Phys. Rev. E*, vol.63, p.021901-1-021901-10, 2001.
- [22] Anderson P. W., Absence of diffusion in certain random lattices. *Physical Review*, vol.109(5), p.1492-1505, 1958.
- [23] OKUNBOR, D. SKEEL R.D. Explicit canonical methods for hamiltonian systems. *Mathematics of Computation*, vol.59, p.439-455, 1992.
- [24] DREW, B. **Understanding DNA**. New York Oxford University Press, 1997.