ISAS - INTERNATIONAL SCHOOL FOR ADVANCED STUDIES

Dynamics of homopolymer chain models

Thesis submitted for the degree of

"Magister Philosophiæ"

CANDIDATE

Cecilia Clementi

SUPERVISOR

Prof. Amos Maritan

October 1996

SISSA - SCUOLA INTERNAZIONALE SUPERIORE DI STUDI AVANZATI

TRIESTE
Strada Costiera 11

TRIESTE

SISSA ISAS
SCUOLA INTERNAZIONALE SUPERIORE DI STUDI AVANZATI
INTERNATIONAL SCHOOL FOR ADVANCED STUDIES

Dynamics of homopolymer chain models

Thesis submitted for the degree of

"Magister Philosophiæ"

CANDIDATE

Cecilia Clementi

SUPERVISOR

Prof. Amos Maritan

October 1996

Table of Contents

1 Introduction

1.1 Strategic motivation

A protein is an heteropolymer chain, comprised of the 20 naturally occurring aminoacids, which can perform biological functions. It is then likely that existing sequences have been selected through natural evolution. A protein is usually modeled as a self-avoiding chain of tens to thousands of monomer units. The monomers belong to the natural set of amino acids.

A class of proteins of major interest is that of globular proteins, existing in typical living cell conditions, i.e. in aqueous solvents, near neutral pH and at 20°-40° C. Under these physiological conditions proteins assume a unique three dimensional globular conformation. In such a folded state (the "native state") they are biologically active [1, 2]. The structural information is encoded in a yet unknown way in the linear amino acid sequence.

The protein folding problem (P.F.P.) is that of predicting this compact three dimensional structure from the knowledge of the amino acid sequence.

The native structure must fulfil two requirements, it must be thermodynamically stable

and it must be kinetically accessible [3, 4, 5].

If we consider the enormously large number of conformations in which the protein could find itself, the protein folding is a very fast process. Would a protein find its native structure after a systematic testing of all the allowed configurations. the folding time should be measured in cosmological time units.

Nature selects those aminoacid sequences that render fast the dynamic evolution toward equilibrium. It is well known that choosing at random the sequence of aminoacids, with probability close to 1, a system is found with a huge degeneracy of the energy minima, and hence with a glassy dynamic behavior that prevents the system from finding a unique stable equilibrium and from quickly finding it. In fact, the dynamics of the random heteropolymer is characterized by frustration, arising by antagonistic interactions between aminoacids, and by the chain topology constraint, which produce a complex connectivity pattern between low energy states. As a result, the energy landscape of a random polypeptide sequence is rugged, with energy barriers of any height [6]. Shackanovic and coworkers [4, 7] have shown by Monte Carlo simulations, that the ground state could possibly be found only for special protein sequences.

At present, Metropolis MonteCarlo dynamics has been the principal tool for studying model proteins, which is equivalent to a statistical mechanical description.

Also molecular dynamics simulations has been widely used to try to simulate the folding of the proteins. Because the harmonic motions of bounded atoms have characteristic times of about 10^{-14}-10^{-13}, stable numerical integration requires femtosecond (10^{-15} second) time

steps. Supercomputers can currently simulate up to nanoseconds (10^{-9} second) of real time protein dynamics with such short time steps, but this scale doesn't approach the 10^{-1}-10^{-3} seconds typically required to fold real proteins. This notwithstanding, natural dynamics might convey informations complementary to MonteCarlo dynamics. In fact, recent developments (numerical and theoretical) in nonlinear Hamiltonian dynamics provide new insights in its rich phenomenology; it has been shown, for instance, that a very rapid increase of the time scales to reach thermodynamic equilibrium is a typical feature of many degrees-of-freedom Hamiltonian systems when energy is lowered below some threshold value [8, 9]. This is a general result [10, 11], related with the existence of qualitatively different chaotic regimes (weak and strong chaos), whose existence can be also analytically predicted on the basis of a new method to treat chaotic Hamiltonian dynamics [12, 13, 14]. In the weakly chaotic regime, a rich variety of dynamical behaviors exists (breathers, metastable coherent structures of solitary type), that keep a system far from thermal equilibrium for a long time. During these long-living non equilibrium states, in the natural (i.e. Hamiltonian) dynamics the decay of correlations is slow, which makes an important difference with respect to a MonteCarlo dynamics, i.e. a Markovian dynamics.

A basic question is how to characterize the dynamics of "good" sequences, i.e. those having a unique native state and short folding times. What kinds of dynamical indicators make possible to discriminate between a "good" and a "bad" sequence?

This work is intended to make a first step toward this problem by studying simplified models of homopolymers, in order to assess the actual interest and feasibility of the

Hamiltonian dynamics approach.

Such a study of conformational properties of chains in three-dimensional space has never been tackled from the dynamical point of view. A natural starting point along this line of research is suggested by the existence of several results [15, 16, 17, 18, 19, 20] obtained by means of equilibrium statistical mechanics - predicting the scaling behavior of conformational quantities, like the end-to-end distance of a polymer chain, with the number N of atoms in the chain according to the type of interactions (short-range only, short plus long-range, attractive and repulsive, etc) - that can be compared with the outcome of numeric simulations.

We can a-priori guess that the predictions of statistical computations will hold true after a relaxation time scale that in some physical conditions could become very long. The wildly fluctuating energy landscape of the glassy phase - in the statistical description with the canonical ensemble - has eventually a microcanonical counterpart in the very slow phase diffusion of the weakly chaotic regime. In this case a new and very interesting question arises: is it possible to find out how the threshold energy between weak and strong chaos moves passing from a "good" to a "wrong" sequence of "aminoacids" in some toy-model? Here one of the basic questions in the P.F.P. is rephrased in the language of natural dynamics, making a link with a well defined dynamical phenomenon (threshold between weak and strong chaos) for which recently developed theoretical tools are available.

2 Dynamical characterization of homopolymer chain models

2.1 Basic Model

We will assume, as is usually done in simple models, that a polymer is represented by a chain of N beads. A configuration of the chain is defined by the positions $\mathbf{r}_1, \ldots, \mathbf{r}_N$ of the beads in the three-dimensional continuous space. A simple choice for the interaction potential [21, 22, 23] is:

$$V_{ij} = \delta_{i,j+1} f(r_{i,j}) + (\frac{\sigma}{r_{ij}})^{12} - \eta_{i,j}(\frac{\sigma}{r_{ij}})^6, \tag{2.1}$$

where $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ are the interparticle distances.

The interaction energy of monomer i with all the other monomers is the sum $\sum_{j \neq i} V_{ij}$. The parameters entering Eq. (2.1) have to be adjusted to fit both the complex interactions between the various groups of aminoacids and the interactions with the solvent. We choose for the energy bond function the expression:

$$f(x) = a(x - d_0)^2 + b(x - d_0)^4, \tag{2.2}$$

that is the fourth-order expansion of a generic and symmetric function of the distance x. The parameter d_0 represents the equilibrium distance of the nearest neighbors along the chain. The Hamiltonian is defined as:

$$H = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2} + \sum_{i=1}^{N} \sum_{j>i} V_{i,j}. \tag{2.3}$$

The first term is the classical kinetic energy of the chain, where the $\mathbf{p}_i$'s are the canonically conjugated variables of the $\mathbf{r}_i$'s.

2.2 The first step: the "free chain"

We start by "switching off" the interparticle potential $V_{i,j}$, i.e. by setting all the constants $\eta_{i,j}$ and σ equal to zero. In so doing, we define the simplest non-trivial idealization of a flexible polymer chain: the Hamiltonian dynamical counterpart - in the three dimensional space - of the random walk model (RWM) on a periodic lattice. In the RWM, the polymer chain configurations are mapped into histories of a random walk on a lattice, i.e. successions of N steps on the lattice, with arbitrary end points. At each step, the next jump may proceed toward any of the nearest-neighbor sites, and the statistical weight for all these possibilities is the same. All the properties of the simple statistical model are easily computed and visualized [15].

Our dynamical model consists of a chain on N "beads" $\mathbf{r}_1, \ldots, \mathbf{r}_N$ separated by "springs"

along the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_{N-1}$ in the three dimensional continuous space ($\mathbf{a}_k = \mathbf{r}_{k+1} - \mathbf{r}_k$). At initial time we generate a random sequence of beads in the space, so the distance between two neighboring beads is $d_0 + \Delta d$. Then the system has the total energy $(N-1)[a(\Delta d)^2 + b(\Delta d)^4]$.

This kind of random configurations, generated at different N and at different energies, are used to initialize the numeric integration of Hamiltonian equations of motion

$$\dot{\mathbf{r}}_i = \frac{\partial H}{\partial \mathbf{p}_i}, \; \dot{\mathbf{p}}_i = -\frac{\partial H}{\partial \mathbf{r}_i}, \quad i = 1, \ldots, N \tag{2.4}$$

from which the dynamical evolution of the chain is obtained. As the initial configurations are chosen at random, we may expect a convergence of time to ensemble averages - of the relevant observables - only after some transient non-equilibrium regime. Eventually, equilibrium will be attained and time averages will only fluctuate around ensemble averages. The following questions naturally arise:

- how fast is the convergence of time to ensemble averages?
- is the relaxation time τ_R of the non equilibrium transient trivial or not as a function of energy density (energy per degree of freedom)?
- is it there any effect of the transition from weak to strong chaos on the relaxation time τ_R?

Let us briefly recall that the ergodic hypothesis at the grounds of statistical mechanics is not sufficient to ensure the finite-time convergence of time and ensemble averages, this finite time convergence requires a stronger condition: a phase space mixing dynamics. Hence it is

understood that a mixing dynamics is required to make useful the predictions of statistical mechanics. Dynamics is mixing if it is unstable with respect to variations, however small, of the initial conditions. With the exception of integrable systems, the generic situation of classical dynamical systems describing N particles interacting through physical potentials, is *instability* of the trajectories in the Lyapunov sense. Nowadays such an instability is called intrinsic stochasticity, or chaoticity, of the dynamics and is a consequence of nonlinearity of the equations of motion. Likewise any other kind of instability, dynamical instability brings about the exponential growth of an initial perturbation, in this case it is the distance between a reference trajectory and any other trajectory originating in its close vicinity that *locally* grows exponentially in time. Quantitatively, the degree of chaoticity of a dynamical system is characterized by the largest Lyapunov exponent λ_1 that – if positive – measures the mean instability rate of nearby trajectories averaged along a sufficiently long reference trajectory. The exponent λ_1 also measures the typical time scale of memory loss of the initial conditions.

Let us recall that if

$$\dot{x}^i = X^i(x^1 \dots x^N) \tag{2.5}$$

is a given dynamical system, and if we denote by

$$\dot{\xi}^i = \mathcal{J}^i_k[x(t)]\, \xi^k \tag{2.6}$$

the usual tangent dynamics equation, where $[\mathcal{J}^i_k]$ is the Jacobian matrix of $[X^i]$, then the

largest Lyapunov exponent λ_1 is defined by

$$\lambda_1 = \lim_{t\to\infty} \frac{1}{t} \ln \frac{||\xi(t)||}{||\xi(0)||} \tag{2.7}$$

and, by setting $\Lambda[x(t),\xi(t)] = \xi^T \mathcal{J}[x(t)]\xi / \xi^T\xi \equiv \xi^T\dot{\xi}/\xi^T\xi = \frac{1}{2}\frac{d}{dt}\ln(\xi^T\xi)$, this can be formally expressed as a time average

$$\lambda_1 = \lim_{t\to\infty} \frac{1}{2t} \int_0^t d\tau\, \Lambda[x(\tau),\xi(\tau)] \quad . \tag{2.8}$$

In recent papers [8, 9, 10, 11] it has been shown that different regimes of mixing exist, fast and slow, characterized by a sudden change of the energy density dependence of τ_R and tightly related to a cross-over between different scaling laws with the energy density ϵ of the largest Lyapunov exponent $\lambda_1(\epsilon)$ defining the transition from weak to strong chaos.

For the "free chain" model of homopolymer, we can analytically compute the ϵ-dependence of the largest Lyapunov exponent, in the large N limit, by taking advantage of a recently developed method [24, 25].

The theoretical prediction of $\lambda_1(\epsilon)$, together with other statistical averages that are analytically computable for the "free chain" model, are then compared with the time averages of the same quantities computed along the dynamical trajectories. Moreover, in the dynamical simulations, we also measure the relaxation time τ_R needed by an average to converge to its expected value.

The dynamical simulations of the system are performed for a finite number N of degrees of freedom; therefore it is also interesting to compare the analytic results, obtained in the limit $N \to \infty$, with the numeric results obtained at different finite N values.

2.2.1 Calculations and results

2.2.1.1 Statistical averages

Statistical averages are computed in the canonical ensemble. The partition function for the system in the d-dimensional space

$$Z = \int \prod_{i=1}^{N} d\mathbf{p}_i \int \prod_{i=1}^{N} d\mathbf{r}_i \exp\left[-\beta H(\mathbf{p},\mathbf{r})\right], \tag{2.9}$$

is then written with the Hamiltonian

$$H = \sum_{i=1}^{N-1} [\frac{\mathbf{p}_i^2}{2} + \frac{a}{2}(|\mathbf{r}_{i+1} - \mathbf{r}_i| - d_0)^2 + \frac{b}{4}(|\mathbf{r}_{i+1} - \mathbf{r}_i| - d_0)^4]. \tag{2.10}$$

Integrating over the $\mathbf{p}$ variables and setting $\mathbf{w}_i = (\mathbf{r}_{i+1} - \mathbf{r}_i)$, $w = |\mathbf{w}|$, we obtain

$$Z = (\frac{4\pi}{\beta})^{dN} (\Omega_d)^N \left[\int_0^{+\infty} w^{d-1} dw \exp(-\frac{\beta a}{2}(w - d_0)^2 - \frac{\beta b}{4}(w - d_0)^4)\right]^N. \tag{2.11}$$

The square of the distance from one end to the other can be expressed in terms of the w variables as follows

$$\begin{aligned} R^2 &= |\mathbf{r}_N - \mathbf{r}_1|^2 = [\sum_{i=1}^{N-1} (\mathbf{r}_{i+1} - \mathbf{r}_i)] \cdot [\sum_{j=1}^{N-1} (\mathbf{r}_{j+1} - \mathbf{r}_j)] \\ &= \sum_{i,j=1}^{N-1} (\mathbf{r}_{i+1} - \mathbf{r}_i) \cdot (\mathbf{r}_{j+1} - \mathbf{r}_j) = \sum_{i,j=1}^{N-1} \mathbf{w}_i \cdot \mathbf{w}_j \end{aligned} \tag{2.12}$$

Hence for the mean value of the end-to-end distance we find

$$\begin{aligned} \langle R^2 \rangle &= \frac{\int \prod_{i=1}^{N-1} d\mathbf{w}_i (\sum_{j,k=1}^{N-1} \mathbf{w}_j \cdot \mathbf{w}_k) \exp -\beta[\sum_{i=1}^{N} \frac{a}{2}(w_i - d_0)^2 + \frac{b}{4}(w_i - d_0)^4]}{\int \prod_{i=1}^{N-1} d\mathbf{w}_i \exp -\beta[\sum_{i=1}^{N} \frac{a}{2}(w_i - d_0)^2 + \frac{b}{4}(w_i - d_0)^4]} \\ &= (N-1)\frac{\int dw w^2 \exp[-\beta\frac{a}{2}(w - d_0)^2 - \beta\frac{b}{4}(w - d_0)^4]}{\int dw \exp[-\beta\frac{a}{2}(w - d_0)^2 - \beta\frac{b}{4}(w - d_0)^4]} \\ &= C(\beta)(N-1). \end{aligned} \tag{2.13}$$

Thus, with the canonical measure, in the limit $N >> 1$ we find the same power law $\langle R^2 \rangle \sim N$ of the random walk model on the lattice. The proportionality constant $C(\beta)$ is now dependent on the temperature, and represents the "mean bond distance" in the chain (that is the lattice constant in the discrete model). It is known [26] that, in the limit $N \to \infty$, we can obtain the microcanonical average $\langle f \rangle_\mu$ of any observable function $f(\mathbf{r})$ in the following parametric form

$$\langle f \rangle_\mu(\epsilon) \to \begin{cases} \langle f \rangle_\mu(\beta) = \langle f \rangle^G(\beta) \\ \\ \epsilon(\beta) = \dfrac{d}{2\beta} - \dfrac{1}{N}\dfrac{\partial}{\partial\beta}[\log Z_C(\beta)] \; . \end{cases} \tag{2.14}$$

where $\langle f \rangle^G$ is the canonical average of the observable f. At finite N it is $\langle f \rangle_\mu(\beta) = \langle f \rangle^G(\beta) + \mathcal{O}(\frac{1}{N})$.

Then, we can expect that, for sufficiently large N and after a suitable transient, the time averages of the end to end distance $\overline{R^2}$ of the chain, numerically computed along the dynamical trajectories of the system, will agree with the scaling law $\overline{R^2} \sim N$, at any fixed value of the energy density $\epsilon = \frac{E}{N}$.

The gyration radius of an homopolymer chain is defined by

$$R_g^2 = \frac{1}{N}\sum_{i=1}^{N} |(\mathbf{r}_i - \mathbf{r}_{cm})|^2 \tag{2.15}$$

where $\mathbf{r}_{cm}$ is the center of mass of the system. The mean square value of this quantity is always proportional to the mean square value of the end-to-end distance. In the numerical calculations it is better to work out the gyration radius rather than the end-to-end distance because the former is computed using the coordinates of all the particles while the latter

using only the coordinates of the ends. So, the numerical fluctuations around the mean value is smaller in the first case than in second one.

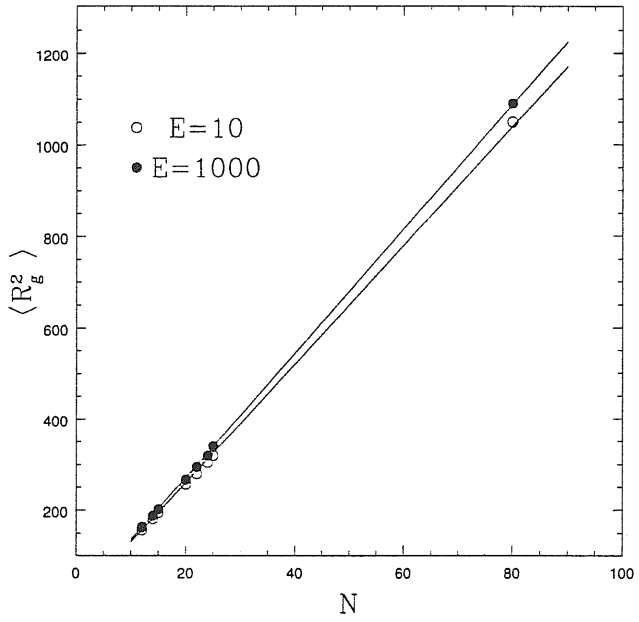

Figure 2.1: Check of the power law $R^2 \sim N$ for two different values of ϵ.

In Fig. 2.1 we show the comparison between analytic and numerical results for two different values of ϵ. The agreement between ensemble and time average is very good, even if the values of N, chosen in the interval $[12, 100]$ in our numeric simulations, are not very large.

All the numerical calculation have been carried out using an efficient algorithm described in appendix (§4).

2.2.1.2 The Lyapunov exponent and the relaxation time

Now we wonder if we can a-priori predict the ϵ - value at which a cross-over of $\lambda_1(\epsilon)$ occurs, and correspondingly a transition between slow and fast mixing will take place (i.e. a transition between slow and fast relaxation of time averages to ensemble averages). A

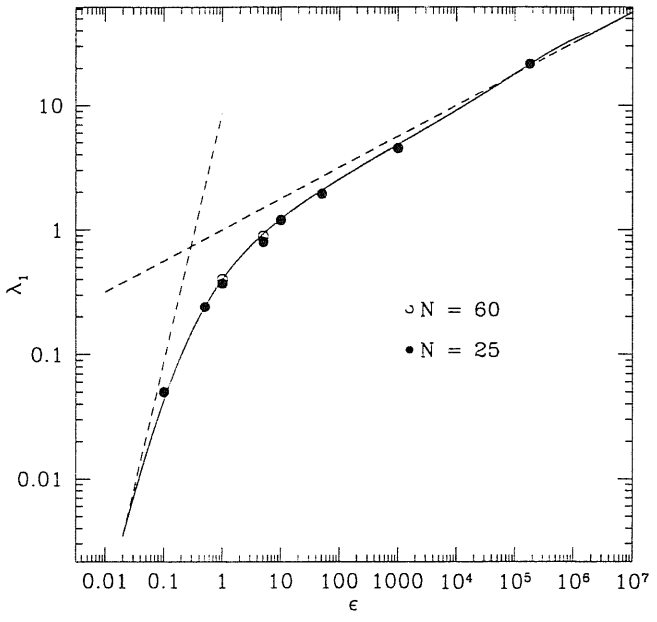

Figure 2.2: Largest Lyapunov exponent λ_1 vs. energy density ϵ: comparison between theoretical prediction (solid line) and numerical results (circles); dotted lines are references to asymptotic power laws ϵ^2 and $\epsilon^{2/3}$, and their intersection defines the SST

recently developed method [24, 25] makes possible to cope with the problem of analytically computing the largest Lyapunov exponent λ_1 for many degrees of freedom Hamiltonian systems as a function of the energy density ϵ. A brief explanation of this method is reported in appendix (§5).

We have made use of this method to compute the exponent λ for a "free chain" model in the large N limit. The results of the analityc computations are shown in Fig 2.2.

We have also numerically computed the "experimental" values of λ_1 by means of the standard algorithm [27] at $N = 25$ and $N = 60$ for different values of ϵ. The agreement between analytic and numeric results is strikingly good. Fig. 2.2 also shows the existence

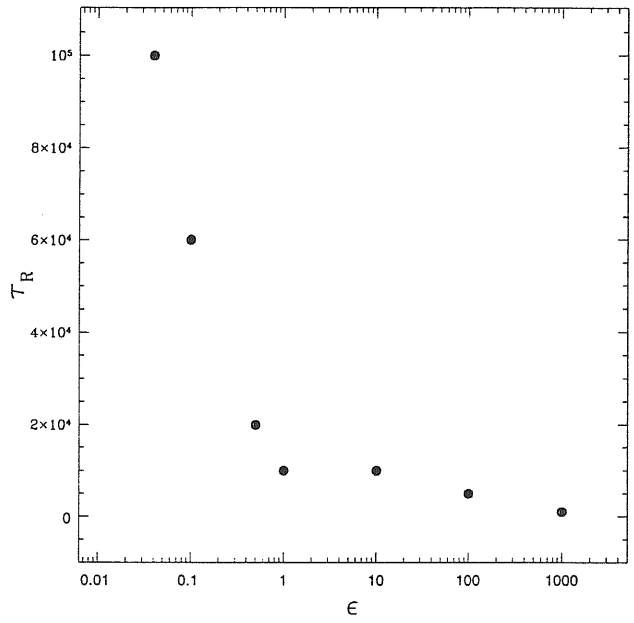

Figure 2.3: Relaxation time τ_R vs. energy density ϵ

of a cross-over energy density ϵ_c between two different scaling-laws of the exponent λ_1. Such a bimodality is characteristic of the transition between regimes of a qualitatively different chaoticity, i.e. of the existence of a strong stochasticity threshold (SST). At $\epsilon > \epsilon_c$ the dynamics possesses good properties of randomness (of course if it is observed with a suitable temporal coarse graining). On the contrary, at $\epsilon < \epsilon_c$ phase-space diffusion is much slower because of a major change of the topology of phase trajectories whose chaoticity gets definitely weaker. Usually, in this weakly chaotic regime, the behavior of a nonlinear, nonintegrable Hamiltonian system displays very long mixing times. We have estimated these mixing times through the numerical relaxation times $\tau_R(\epsilon)$ of the time averages of the mean square radius $\overline{R^2}$ to its theoretically predicted values at different energy densities.

In Fig. 2.3 we show that there is a rather sudden change of the behavior of $\tau_R(\epsilon)$ at

the same value ϵ_c at which the SST occurs. this is a very interesting result in asmuch the dynamics provides us with a new physical parameter - τ_R - which is dynamically generated and whose energy dependence is non-trivial. This time scale has to be eventually compared with other physical time scales of the system under consideration, thus a-priori enriching the physical phenomenology that can be described by the natural dynamics.

2.3 The second step: the "self-avoiding chain"

Let us now "switch on" the repulsive interaction among the particles by modifing the "free chain" Hamiltonian as follows

$$H = \sum_{i=1}^{N-1} [\frac{\mathbf{p}_i^2}{2} + \frac{a}{2}(|\mathbf{r}_{i+1} - \mathbf{r}_i| - d_0)^2 + \frac{b}{4}(|\mathbf{r}_{i+1} - \mathbf{r}_i| - d_0)^4] + \sum_{i>j}(\frac{\sigma}{|\mathbf{r}_i - \mathbf{r}_j|})^{12} \qquad (2.16)$$

so that we have defined a Hamiltonian counterpart of the self-avoiding walks (SAW) on a periodic lattice, these are the walks which can never self-intersect. Certain mathematical properties that are trivial for simple random walks become complex in the SAW case. Indeed the SAW is a rather accurate model for the real chain in a *good*[1] solvent. The end-to-end distance has a mean square average which now scales as:

$$\langle R^2 \rangle \sim N^{2\nu}. \qquad (2.17)$$

The existence of a non-trivial exponent ν is proved by the renormalization group calculations, but the exponent ν is not exactly known. For three dimensional systems, numerical

[1] In a good solvent the polymers are in the swollen phase.

and experimental results agree with a value of ν very close to 3/5. Long ago, Flory devised a simple and brilliant scheme for computing the exponent ν, which gives excellent values for all dimensionality [28]. He roughly estimated the free energy as a sum of a repulsive energy proportional to the square of the average monomer concentration

$$\frac{F_{rep}}{T} \sim \left(\frac{N}{R^d}\right)^2 R^d \tag{2.18}$$

and an entropy term computed for an ideal chain as

$$S \sim \frac{R^2}{N}. \tag{2.19}$$

It is easily found that this free energy has a minimum for a radius $R = R_F$ defined by the relation:

$$R_F^{d+2} \sim N^3. \tag{2.20}$$

This equation states that

$$\nu = \frac{3}{2+d}, \tag{2.21}$$

which is an amazingly good assessment; it gives the correct value for $d = 1$ and the values for $d = 2$ and $d = 3$ are within a percent of the most accurate numerical results.

2.3.1 Calculations and results

On a lattice, the effect of the repulsive interaction between the monomers is built in the self-avoidness condition. Two non-nearest-neighbors monomers along the chain must stay more distant, one from the other, than one lattice constant. If $\sigma = 0$, again we have a random walk. As soon as we switch-on the repulsion, we prevent two monomers from contact. The

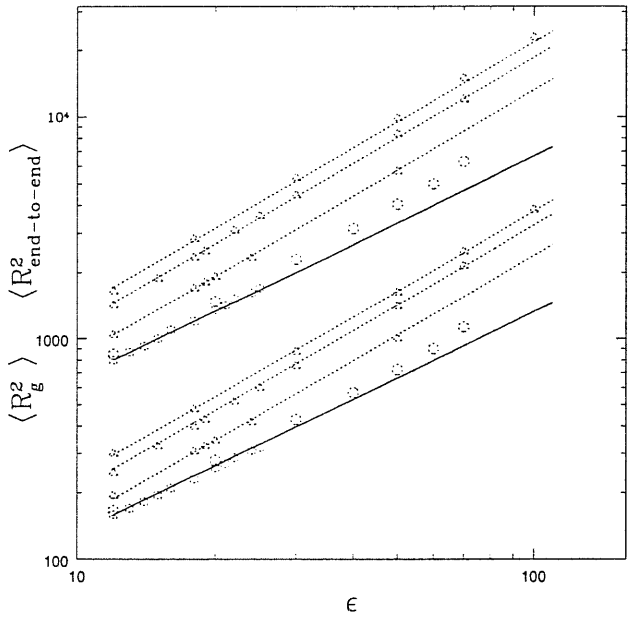

Figure 2.4: Behaviour of the gyration radius (lower curves) and the end to end distance (upper curves) for different values of the ratio σ/d_0; solid lines are references to power law $R^2 \sim N^{6/5}$, dotted lines to $R^2 \sim N$.

repulsive potential $(\frac{\sigma}{r})^{12}$ is a very steep potential and, roughly speaking, we can say that monomers i, j cannot get closer than a distance $\mathbf{r}_{i,j} \sim \sigma$. In the Hamiltonian (2.16) there is another scale of distance, that is fixed from the parameter d_0, the equilibrium distance of two nearest neighbors along the chain. When $\sigma << d_0$, a very large number N of monomers in the chain are needed in order to appreciate the change of the power law from $R^2 \sim N$, to $R^2 \sim N^{2\nu}$. We are interested in studying the dynamical behavior of the system in a finite range of N ($N \in [10, 100]$), where we expect the correct scaling with the exponent $\nu \simeq 3/5$ if the ratio σ/d_0 exceeds a critical value.

Fig. 2.4 shows the results for $\sigma/d_0 = \sqrt[12]{500}/10$, $\sqrt[12]{10^5}/10$, $\sqrt[12]{10^8}/10$, $\sqrt[12]{10^{10}}/10$, $\sqrt[12]{10^{11}}/10$

in this range of N.

- For $\sigma/d_0 = \sqrt[12]{500}/10$ we have the scaling $R^2 \sim N$, as in a random walk.
- For $\sigma/d_0 > \sqrt[12]{10^8}/10$ we have the correct scaling of the SAW.
- For $\sigma/d_0 = \sqrt[12]{10^5}/10$ we note a change of the power law at an intermediate value of N in the range considered.

We have used a Flory-like method, with its approximations, modifying it as to account for a dependence on the parameter σ/d_0, in order to fit different curves by means of a unique expression.

We guess a scaling function for the radius as follows:

$$R^2(\varepsilon, N) \sim N^{2\nu} f\left(\frac{N}{N(\varepsilon)}\right) \tag{2.22}$$

where f is an universal and unknown function. The parameter ε is a measure of the "efficiency" of the repulsion, then we put

$$\varepsilon \sim \left(\frac{\sigma}{d_0}\right)^{\phi}, \tag{2.23}$$

where ϕ is an exponent to be determined. $N(\varepsilon)$ is the value of N above which the exponent ν is that of the SAW, we assume that it obeys the following expression:

$$N(\varepsilon) \sim \varepsilon^{-\psi} \tag{2.24}$$

where we introduce another exponent ψ. The function f must be such as to reproduce always the SAW exponent ν in the limits $N \to \infty$ and finite ε, and the random walk

exponent (= 0.5) in the limit $\varepsilon \to 0$, i. e.:

$$\begin{aligned} f(x) &\to const. \quad \text{as} \quad x \to \infty \\ f(x) &\to x^{1-2\nu} \quad \text{as} \quad x \to 0 \end{aligned} \tag{2.25}$$

Then, like in the Flory's scheme, we proceed to a rough estimate of the entropy that enters the free energy expression, the repulsion energy must be now dependent on ε, so the following simple approximation is natural:

$$\frac{F_{rep}}{T} \sim \varepsilon \left(\frac{N}{R^d}\right)^2 R^d. \tag{2.26}$$

Hence we can write (for a generic dimension d):

$$R^2 \sim N^{2\nu} \varepsilon^{\frac{2}{2+d}}. \tag{2.27}$$

For each value of ε, there is a critical value $N(\varepsilon)$ at which a cross-over of the two different power laws occurs, then we can write

$$N(\varepsilon)^{\frac{1}{2}} \sim N(\varepsilon)^{\nu} \varepsilon^{\frac{1}{2+d}}, \tag{2.28}$$

and therefore:

$$N(\varepsilon) \sim \varepsilon^{-\frac{2}{4-d}} \tag{2.29}$$

that fixes the value of the parameter ψ to $\psi = 2/(4-d)$. This expression is reasonable because $d = 4$ is the upper critical dimension of the SAW. In our numerical simulations $d = 3$, and then $\psi = 2$.

In Fig. 2.5 we show that all the curves of the previous figure collapse an a unique curve if we take for the exponent ϕ a value of $\phi = 3$. So the guessed scaling function (2.22) is verified.

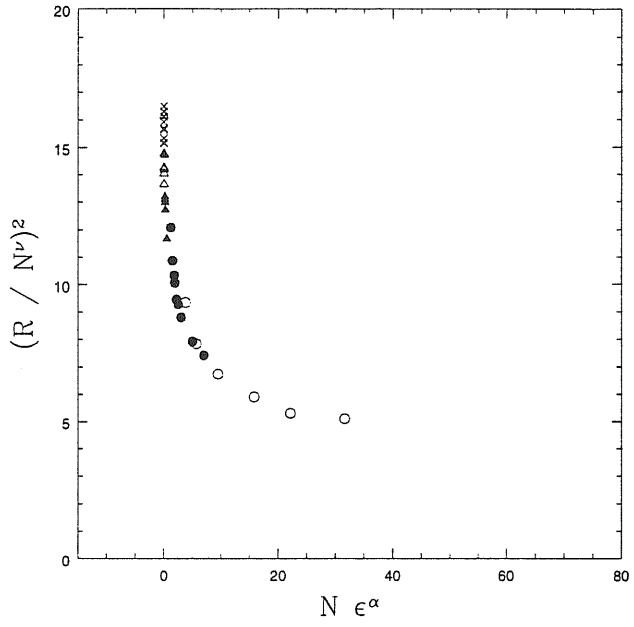

Figure 2.5: Collapse plot for the scaling law (2.22)

2.3.1.1 Dynamical behavior

We can now study the behavior of the Lyapunov exponent of the self-avoiding chain. We compute it numerically, by the standard algorithm [27].

In Fig. 2.6 the comparison among the numerical results, for the case $\sigma = \sqrt[12]{10^{11}}$, and the analytic ones, for the case of the random walk ($\sigma = 0$), is plotted. Notice that again for $\sigma = \sqrt[12]{10^{11}}$ the analytic curve computed at $\sigma = 0$ gives a good approximation of the behavior of the Lyapunov exponent $\lambda_1(\epsilon)$ as a function of the energy density ϵ. So, while the scaling law changes from $R^2 \sim N$ to $R^2 \sim N^{2\nu}$, in the same range of parameters, the dynamics seems to keep the features of the random walk case. Indeed, the introduction of the repulsive energy in the Hamiltonian has apparently no effect on the degree of instability of the dynamics in the three dimensional space. The repulsive potential $(\sim 1/r)^{12}$ grows

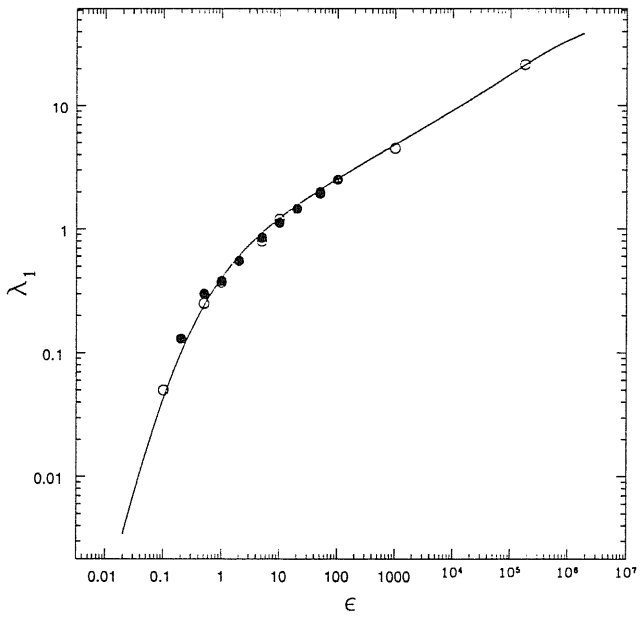

Figure 2.6: Comparison among the largest Lyapunov exponent for the RWM case (open circles) and the SAW case (full circles). The solid line is the analytic prediction for the RW case

up very steeply for $r < 1$. Thus the phase space can be viewed like the previous one (i. e. the random walk case) with some forbidden regions. If these forbidden regions of phase space are not too large (in a sense that should be made mathematically precise, what is now beyond the aims of the present work) we can think that some observables, that are sufficiently uniform in phase space, will not be very much affected. Now, in the strongly chaotic regime all the geometric properties - of configuration space and of phase space - that determine the value of Lyapunov exponent (see Appendix 2), are rather uniform: they do not change significantly when are measured in different regions of the ambient manifold. Therefore, being the average Ricci curvature and its fluctuations rather uniform, they are not very sensitive to the esclusion of some regions of phase space, consequently neither the Lyapunov exponent will be affected. We may expect that this reasoning will cease to be valid when the range of the repulsive potential will become closer to the interatomic distance.

2.4 The θ point of a linear polymer in two dimensions

The study of some dynamical properties of the free-chain and of the self-avoiding chain has positively answered to our initial question about the interest and the feasibility of the Hamiltonian dynamical approch, therefore we can now begin to study a more complex case, in which also the attractive forces are relevant. This is a more realistic choice for a potential between the amino-acids in a protein. The prevalence of the attractive forces or of the repulsive ones between the amino-acids depends on the range of the interaction, that

in turn depends on the pair of the involved aminoacids. Here we study the simplest model where the same type of monomer is put on all the sites of the chain. This corresponds to set the parameters $\eta_{i,j} = \eta = const.$, $\forall i, j$ in the expression (2.1). Hence we deal with an homopolymer chain with pair-wise interparticle interaction of Lennard-Jones type. We introduce the parameters $\gamma = \frac{\eta^2}{4}$ and $\lambda = \frac{\sigma}{\sqrt[6]{\eta}}$ so that we can write the Lennard-Jones potential in the standard form:

$$V(r) = 4\gamma \left[\left(\frac{\lambda}{r}\right)^{12} - \left(\frac{\lambda}{r}\right)^{6} \right]. \tag{2.30}$$

It is been known for a long time [28], that by varying the temperature an homopolymer chain presents two different phases according to the dominance of actractive or repulsive interaction energy (as it can be the with the Lennard-Jones potential). At high temperature, the attractive part of the potential is negligible and the chain is in the swollen phase: in these conditions the system closely behaves as in the case with only repulsive forces (SAW). So, the same power law of the self-avoidinge case, $R^2 \sim N^{2\nu}$, is found. As the temperature lowers, the attractive terms become more relevant and the chain becomes more and more compressed, then the radius of system scales as $R^2 \sim N^{2/d}$. The separation between the regimes of self-avoiding chain and of collapsed chain is marked by a θ temperature; at this temperature a phase transition occurs. Here we study the properties of this model on two-dimensional space. For $d = 2$ the Flory estimation of the exponent ν provides $\nu = 3/4$, whereas in the collapsed case we have $\nu = 1/2$.

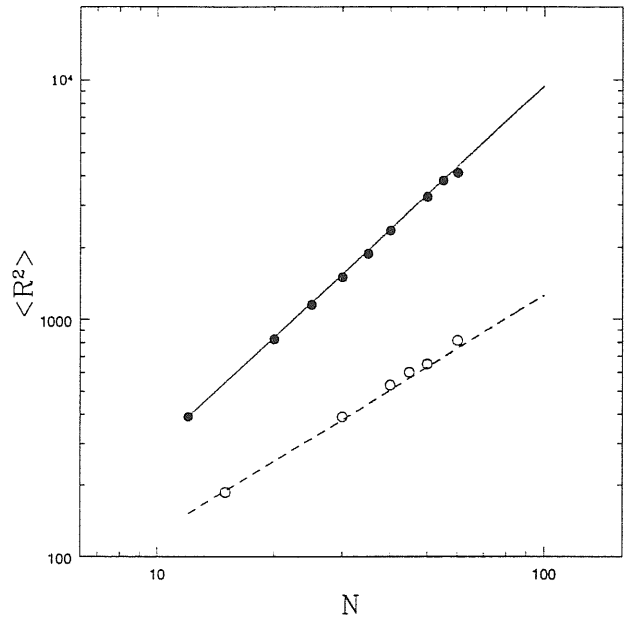

Figure 2.7: Check of the two different scaling laws above (full circles) and below (open circles) the theta point. Solid line is reference to $R^2 \sim N^{3/2}$, dotted line to $R^2 \sim N$.

2.4.1 Calculations and results

First of all, we have tried to obtain the correct power law by dynamical simulations performed in the range $N \in [10 - 100]$. We fix the parameters γ and λ so that at high temperature the system is in the case of the correct scaling in this range of N (§2.3.1). In the dynamical simulations we vary the energy density $\epsilon = E/N$, which is a function of the temperature [Eq.(2.14)], that in our microcanonical simulations acquires the meaning of a mean kinetic energy per degree of freedom.

In Fig. 2.7 we show that the two different scaling laws are found for values of the energy density $\epsilon = 10$ and $\epsilon = -2$ [2]. Hence at these values of ϵ the system is in two different phase.

[2]Here we use dimensionless quantities. Physical units can be easily introduced when we want to describe

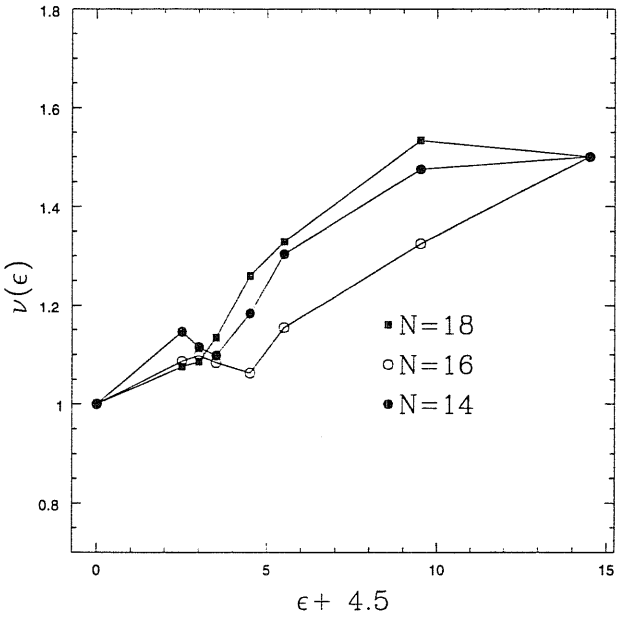

Figure 2.8: Plot of the behaviour of $\nu(\epsilon)$ for different values of N.

In order to have an extimation of the critical value ϵ_θ at which the transition occurs, we have performed a finite size scaling analysis. An immediate way of analyzing approximate results for R_N^2, at finite N and near a θ point, consists in computing effective exponents of the type:

$$2\nu_N = \frac{\ln(R_N^2/R_{N-2}^2)}{ln\left(\frac{N}{N-2}\right)}. \tag{2.31}$$

The curves at different N have a clear tendency to intersect each other in a restricted region, as showed in Fig. 2.8.

Abscissas and ordinates of these intersection points yield approximations to ϵ_θ and ν_θ, respectively [29].

---

a given real system for which masses, lengths and parameters of the potentials are assigned

2.4.1.1 The Lyapunov exponent and the dynamical characterisation of the θ transition

Recently it is been conjectured that there exists a non-trivial relationship between the dynamical properties (like those described by Lyapunov exponents) and the behaviour of the statistical observables when the system under consideration undergoes a second order phase transition [30, 31, 32, 33, 34, 35]. It is observed that the behaviour of the Lyapunov exponent is not smooth near the transition point: the dependence of λ_1 on ϵ changes rather abruptly at ϵ_c. A differential geometric approch to Hamiltonian [12, 13, 14] is useful in trying to explain this result. Indeed the curvature fluctuations of the Riemaniann manifold whose geodesics are the trajectories of the dynamical system undergoing a phase transition, exhibit a singular behaviour at the critical energy density. This singular behaviour is reproduced in abstract geometric models and suggests that the phase transition corresponds to a change in the topology of the ambient manifold. The relevance of the topological concepts for the theory of phase transitions has already been suggested in a more abstract context [36].

We have numerically computed the largest Lyapunov exponent $\lambda_1(\epsilon)$ of the system at varying the energy density ϵ. As shown in Fig. 2.9 we actually find a rather sharp change in the behaviour of the $\lambda_1(\epsilon)$ at an energy density that agrees with the critical value ϵ_c, previously extimated. The numerical calculations of the Lyapunov exponent have been performed at $N = 25$, but already seem to hint for the existence of the phase transition, which is properly defined in the thermodynamic limit. In order to check that these numerical

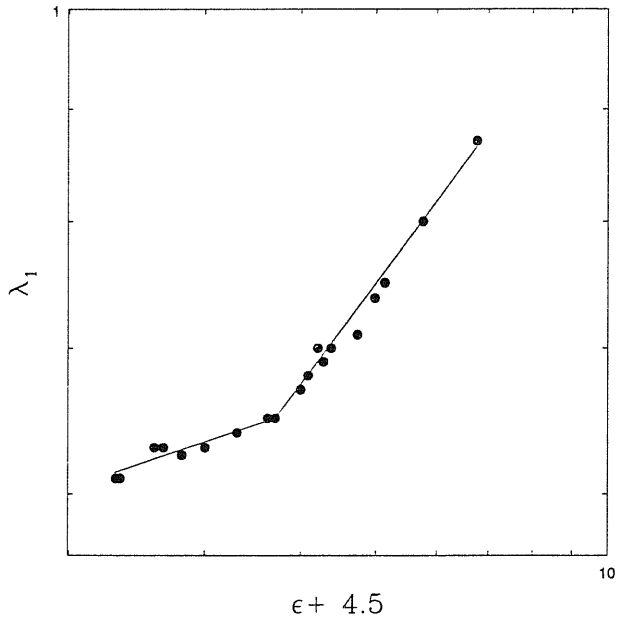

Figure 2.9: Cross-over between two different behaviors of the Lyapunov exponent at the theta point.

calculations really have some relevance to estimate the behaviour that some observables have in the limit $N \to \infty$, we have analytically computed the high temperature behaviour of $\lambda_1(\epsilon)$. Indeed, even if the statistical averages are not exactly computable, we can use the above explained method (§2.2.1.2) with the same approximation for the high temperature phase. When the attractive part of the potential is negligible, the chain is, in practice, a self-avoiding one. In the self-avoiding case, in the three dimensional space, we note that the behaviour of the Lyapunov exponent is not different from the random walk (see §2.3.1).

In the two dimensional space a similar result holds for the self-avoiding chain, but it is necessary a more careful estimate of the average quantities that are needed to apply the theoretical method of refs. [24, 25]. If, to obtain the mean value of same observable, we integrate over all the possible configurations by means of the canonical measure $\exp(-\beta V_0)$ of the free chain potential V_0 (Eq. 2.10), the integrals diverge. But it is very unlikely that two particles will happen to be closer one another than the length scale γ of Lennard-Jones interaction. Thus we can approximate the difference between the canonical measure for the free chain and that of the self-avoiding chain by the exclusion of the volume of the sphere of radius equal to γ around each particle. A similar procedure does not change the result in three dimensional space because, there, the integrand functions vanish in this region (see appendix §5).

In Fig. 2.10 we show the analytical result for the Lyapunov exponent obtained by the application of this method. The agreement with the numerical values is very good in the high temperature phase, at energy density greater than the critical value ϵ_c. By decreasing the

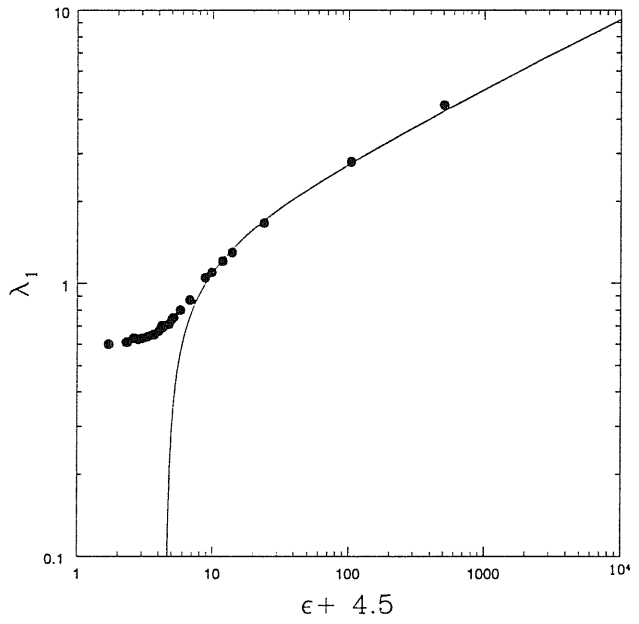

Figure 2.10: Comparison between the numerical values of the Lyapunov exponent (N=25) and the analytic prediction for the SAW case.

energy density, when the agreement of the numerical values with the theoretical prediction for the SAW case becomes poor, we can imagine that the attractive part of the Lennard-Jones potential ceases to be negligible. At even lower energy densities, the attractive energy becomes comparable with the repulsive one, and the theta transition shows up.

3 Conclusions and perspectives

The results here reported give a positive answer to the questions that have been formulated in the Introduction, concerning the feasibility and the interest of a dynamical approach to the statistical properties of homopolymer chain models.

Throughout the present work dynamics means natural (Hamiltonian) microscopic dynamics.

First, the dynamical approach has reproduced all the expected basic laws for spatial packing that are predicted for homopolymers on the basis of idealized models for the spatial conformations seen as random walks on a lattice. This is already a non trivial result because in our dynamical models the different monomer units are free to move continuously in space, moreover the change of conformation in time is driven by a non trivial differentiable dynamics. Thus it is interesting to find that, already at small number N of monomer units, the statistically predicted scaling laws are reproduced by natural dynamics. However the dynamics has something of its own to add to the existing knowledge. In fact, starting from a generic initial condition, only after some *relaxation* time τ_R the time averages converge to their statistically predicted values. Moreover, τ_R has a non trivial energy density

(ϵ) dependence. A sudden change of $\tau_R(\epsilon)$ is observed in correspondence of the Strong Stochasticity Threshold that generically exists for nonintegrable many-degrees-of-freedom Hamiltonian systems. In other words, the transition from different regimes of strenght of chaos, signaled by a cross-over in the ϵ-dependence of the largest Lyapunov exponent, brings about a qualitative change in the characteristic phase space mixing time.

It is also worth mentioning that the dynamical approach can convey other kinds of complementary physical informations, as is the case of the "self-avoiding chain" that compares to the SAW model on a lattice. In this latter model just by switching-on the condition of self-avoidedness induces a change in the scaling law with N of the end-to-end as well as of the gyration radius. At variance, in the Hamiltonian model the change of this scaling law depends on the relative magnitude of the effective range of the repulsive potential and of the interparticle distance, which is physically sound.

Another interesting result is that of a clear dynamical signature of the "theta"-transition between filamentary and globular conformations that already shows up at rather small N, a situation very far from the $N \to \infty$ limit which is implicit in the statistical mechanical approach. Here again dynamics gives complementary informations of physical relevance, as real polymers may exist with few tens of monomer units.

Finally, in view of attacking the study of the dynamics of heteropolymer chains and of proteins, the above mentioned transition between different regimes of chaos, and thus of phase space diffusion efficiency, suggests that phenomena like the glassy transition of heteropolymers might well have a dynamical counterpart whose study could benefit also

of new theoretical tools to study the transition between slow and fast phase space mixing. Hence the approach started in this work appears potentially very interesting to cope with the problem of dynamically recognizing the difference between "good" aminoacid sequences of real proteins and "bad" sequences of random heteropolymers.

4 Appendix 1: Numerical computations

A faithful numerical representation of a Hamiltonian flow (i.e. fulfilling Liouville theorem, energy conservation, conservation of Poincaré invariant) is guaranteed only by symplectic integration schemes. A very efficient and precise symplectic algorithm has been recently proposed [37]. This integrator alternates two different choices of the function generating the canonical transformation that maps coordinates and momenta from any time t to a subsequent time $t + \Delta t$. For this reason the algorithm is called "bilateral".

Let us recall that the well known leap-frog scheme which – for Hamiltonians like $H(\mathbf{q}, \mathbf{p}) = \sum_{i=1}^{N} p_i^2/2 + V(\mathbf{q})$ – reads as

$$\begin{aligned} q_i(t + \Delta t) &= q_i(t) + \Delta t\, p_i(t) \\ p_i(t + \Delta t) &= p_i(t) - \Delta t \frac{\partial}{\partial q_i} V\left(\mathbf{q}(t + \Delta t)\right) , \end{aligned} \tag{4.1}$$

is a canonical transformation of variables generated by the function

$$F(\mathbf{Q}, \mathbf{p}, \Delta t) = -\mathbf{Q} \cdot \mathbf{p} + \Delta t H(\mathbf{Q}, \mathbf{p}) , \tag{4.2}$$

(small letters – $\mathbf{p}$, $\mathbf{q}$ – refer to time t while capital ones to the same variables at time $t + \Delta t$).

At infinitesimal Δt, the function (4.2) becomes exactly the generating function of the natural motion of the system in phase space [38]. There is some arbitrariness in the choice of (4.2) to construct the numerical integration scheme: in fact, a function Φ obtained from (4.2) by interchanging the role played by the coordinates with that of their conjugated momenta,

$$\Phi(\mathbf{q}, \mathbf{P}, \Delta t) = \mathbf{q} \cdot \mathbf{P} + \Delta t H(\mathbf{q}, \mathbf{P}) , \tag{4.3}$$

has the same meaning of F for a vanishing Δt; the transformation generated by (4.3) is in fact an alternative form for the leap-frog algorithm.

The difference between these two limits shows up only at finite Δt, and so it has practical consequences only when these canonical transformations are used in the form of numerical integration algorithms. Hence the simple but ingenious idea of compensating the errors of each partial scheme by alternating them in a "bilateral" algorithm [37].

Moreover, by applying this idea to the general scheme to generate higher order symplectic algorithms reported in ref. [39], a second order bilateral algorithm is worked out [37] in the following form

$$\begin{aligned}
\tilde{q}_i &= q_i(t) \\
\tilde{p}_i &= p_i(t) - \frac{1}{2}\Delta t \frac{\partial}{\partial \tilde{q}_i} V(\tilde{\mathbf{q}}) \\
q_i(t + \Delta t) &= \tilde{q}_i + \Delta t\, \tilde{p}_i \\
p_i(t + \Delta t) &= \tilde{p}_i - \frac{1}{2}\Delta t \frac{\partial}{\partial q_i} V\left(\mathbf{q}(t + \Delta t)\right) \\
\hat{p}_i &= p_i(t + \Delta t)
\end{aligned}$$

$$
\begin{aligned}
\hat{q}_i &= q_i(t+\Delta t) + \frac{1}{2}\Delta t\, \hat{p}_i \\
p_i(t+2\Delta t) &= \hat{p}_i - \Delta t \frac{\partial}{\partial \hat{q}_i} V(\hat{\mathbf{q}}) \\
q_i(t+2\Delta t) &= \hat{q}_i + \frac{1}{2}\Delta t\, p_i(t+2\Delta t)\,. \qquad (4.4)
\end{aligned}
$$

5 Appendix 2: Riemannian theory of Hamiltonian chaos and Lyapunov exponent

Here we give a sketchy summary of the concepts involved in a differential-geometric method to describe hamiltonian chaos and to predict theoretically the value of the largest Lyapunov exponent. Longtime ago the geometrical formulation of the dynamics of conservative systems [40] was first used by Krylov in his studies on the dynamical foundations of statistical mechanics [41] and subsequently became a standard tool to study abstract systems like Anosov flows in the framework of ergodic theory [42]. In recent papers [12, 13, 24, 25] the geometric approach has been extended in order to consider physical models like coupled nonlinear oscillators.

Let us briefly recall that the dynamics of N-degrees-of-freedom systems defined by a Lagrangian $\mathcal{L} = T - V$, in which the kinetic energy is quadratic in the velocities, $T = \frac{1}{2} a_{ij} \dot{r}^i \dot{r}^j$, can be rephrased in geometrical terms due to the fact that the natural motions are

the extrema of the Hamiltonian action functional $\mathcal{S}_H = \int \mathcal{L} dt$ or of the Maupertuis' action $\mathcal{S}_M = 2 \int T\, dt$. In fact the geodesics of a Riemannian manifold are themselves the extrema of the arc-length functional $\ell = \int \sqrt{g_{ij} dr^i dr^j}$, hence a suitable choice of the metric tensor allows for the identification of the arc length with either $\mathcal{S}_H$ or $\mathcal{S}_M$, and of the geodesics with the natural motions of the dynamical system. Starting from $\mathcal{S}_M$ we obtain the so-called Jacobi metric on the accessible configuration space, $(g_J)_{ij} = [E - V(\{r\})]\, a_{ij}$. A description of the extrema of Hamilton's action $\mathcal{S}_H$ as geodesics of a manifold can be obtained using Eisenhart's metric [43] on an enlarged configuration spacetime ($\{t \equiv r^0, r^1, \ldots, r^N\}$ plus one real coordinate r^{N+1}, related to the action), whose arc-length is

$$ds^2 = -2V(\mathbf{r})(dr^0)^2 + a_{ij} dr^i dr^j + 2 dr^0 dr^{N+1} \,. \tag{5.1}$$

The manifold has a Lorentzian structure and the dynamical trajectories are recovered as those geodesics satisfying the condition $ds^2 = C dt^2$, where C is a positive constant. In the geometrical framework, the stability of trajectories is therefore mapped to the stability of geodesics, hence it is completely determined by the curvature properties of the underlying manifold: the field J which measures the deviation between nearby geodesics obeys Jacobi equation [44]

$$\nabla^2_{\dot\gamma} J + R(\dot\gamma, J)\dot\gamma = 0 \,, \tag{5.2}$$

where $\nabla_{\dot\gamma}$ stands for the covariant derivative along the geodesic $\gamma(s)$, R is the Riemann curvature tensor, and $\dot\gamma$ is the velocity field along γ. J is commonly referred to as the Jacobi field. In the case of hyperbolic isotropic manifolds the curvature term in Eq. (5.2) can be rewritten as $R(\dot\gamma, J)\dot\gamma = KJ$ and the sectional curvature K is a *negative* constant, thus

equation (5.2) has exponentially growing solutions and the system is dynamically unstable. This is the origin of chaotic dynamics in Anosov flows [42]. As far as coupled nonlinear oscillators are concerned the geometric picture is completely different since because all the curvatures (Ricci, scalar, sectional) are mainly (and in some cases strictly) *positive*. Actually, negative curvatures are not necessary to make chaos, and exponentially growing solutions of the stability equation (5.2) can be obtained through *parametric resonance* even if no negative curvature is experienced by the geodesics [12, 13, 24, 25]. In the large N limit, and under the assumption that the manifold is nearly isotropic, this mechanism can be modeled by replacing Eq. (5.2) with an effective scalar Jacobi equation [24, 25] which reads

$$\frac{d^2\psi}{ds^2} + \kappa(s)\,\psi = 0\ , \tag{5.3}$$

where $\psi^2 \propto |J|^2$ and the "effective curvature" $\kappa(s)$ is a stochastic, δ-correlated gaussian process whose mean κ_0 and variance σ_κ are identified respectively with the average and the rms fluctuations of the Ricci curvatures at any given point $k_R = K_R/N$ (which is itself the average of the sectional curvature over the directions of J) along a geodesic

$$\kappa_0 = \frac{\langle K_R\rangle}{N}\ , \qquad \sigma_\kappa^2 = \frac{\langle (K_R - \langle K_R\rangle)^2\rangle}{N}\ . \tag{5.4}$$

Using Eisenhart's metric $K_R = \Delta V = \sum_{i=1}^{N} \partial^2 V/\partial r_i^2$. The exponential growth rate λ of the envelope of the solutions of Eq. (5.3), which in this picture is the natural estimate of the Lyapunov exponent, can be computed exactly:

$$\lambda = \frac{\Lambda}{2} - \frac{2\kappa_0}{3\Lambda}, \qquad \Lambda = \left(2\sigma_\kappa^2 \tau + \sqrt{\frac{64\kappa_0^3}{27} + 4\sigma_\kappa^4 \tau^2} \right)^{\frac{1}{3}} \tag{5.5}$$

where $\tau = (\pi\sqrt{\kappa_0})/(2\sqrt{\kappa_0(\kappa_0 + \sigma_\kappa)} + \pi\sigma_\kappa)$; in the limit $\sigma_\kappa/\kappa_0 \ll 1$ one finds $\lambda \propto \sigma_\kappa^2$. The details can be found in Refs. [24, 25].

5.1 Application of this model to the homopolymer "free-chain"

The quantities needed by the above model to work out a value of the largest Lyapunov exponent λ_1 are the average and the rms fluctuations of Ricci curvature, i. e. - using Eisenhart's metric - the mean and the fluctuation of the function:

$$K_R = \sum_{i=1}^{N} \frac{\partial^2 V}{\partial r_i^2}. \tag{5.6}$$

In the case of the system described by the Hamiltonian (§2.10) in the d-dimensional space, this expression reads

$$\begin{aligned} K_R &= \sum_{i=1}^{N} \sum_{\mu=1}^{d} \frac{\partial^2 V(\mathbf{r})}{\partial r_{i,\mu}^2} \\ &= 2adN - 2a(d-1) \sum_{i=1}^{N} \frac{d_0}{|\mathbf{r}_{i+1} - \mathbf{r}_i|} + 6b \sum_{i=1}^{N} (|\mathbf{r}_{i+1} - \mathbf{r}_i| - d_0)^2 \\ &+ 2b(d-1) \sum_{i=1}^{N} \frac{(|\mathbf{r}_{i+1} - \mathbf{r}_i| - d_0)^3}{|\mathbf{r}_{i+1} - \mathbf{r}_i|}. \end{aligned} \tag{5.7}$$

In order to compute the Gibbsian average, we slightly modify the canonical partition function as follows:

$$\tilde{Z}(\alpha) = \int \prod_{i=1}^{N} d\mathbf{p}_i \int \prod_{i=1}^{N} d\mathbf{r}_i \exp\left[-\beta H(\mathbf{p}, \mathbf{r}) + \alpha K_R\right], \tag{5.8}$$

so the problem of computing the mean values is now reduced to that of working out

$$\langle \frac{K_R}{N} \rangle = \frac{1}{N} \left[\frac{\partial}{\partial \alpha} \tilde{Z}(\alpha) \right]_{\alpha=0} \tag{5.9}$$

$$\frac{\langle (K_R - \langle K_R \rangle)^2 \rangle}{N} = \frac{1}{N} \left[\frac{\partial^2}{\partial \alpha^2} \tilde{Z}(\alpha) \right]_{\alpha=0}. \tag{5.10}$$

Setting $\mathbf{w}_i = (\mathbf{r}_{i+1} - \mathbf{r}_i)$, integrating over the $\mathbf{p}$ we obtain:

$$\tilde{Z}(\alpha) = [\tilde{z}(\alpha)]^N, \tag{5.11}$$

where

$$\tilde{z}(\alpha) = \left(\frac{4\pi}{\beta} \right)^{\frac{d}{2}} \Omega_d \int_0^{+\infty} w^{d-1} dw \exp(-\frac{\beta a}{2}(w - d_0)^2 - \frac{\beta b}{4}(w - d_0)^4 + \alpha k(w)), \tag{5.12}$$

with

$$k(w) = 2ad - 2a(d-1)\frac{d_0}{w} + 6b(w - d_0)^2 + 2b(d-1)\frac{(w - d_0)^3}{w}. \tag{5.13}$$

Therefore the quantities to be computed are the integral (5.12) and its derivative with respect to α. These integrals run only over the variable w, then (5.12) is easily computed numerically.

Acknowledgments

First of all, I want to thank my supervisor, prof. Amos Maritan, for having introduced me in this field of research, and dr. Marco Pettini, for having continuously and friendly assisted me in this work.

I wish to thank Paolo De Los Rios, Michele Vendruscolo and Flavio Seno for helpful comments and discussions.

I am grateful to Paolo e Michele also because I have been "camping" in their room for days and days.

Furthermore, an heartfelt acknowledgment is addressed to Lapo Casetti, Lando Caiani, Monica Cerruti-Sola, Alessandro Laio, Claudio Tebaldi, Guglielmo Iacomelli, Simone Bianchi, Francesca Colaiori, Alessandro Flammini, Marco Saitta, Alice Ruini and Sandro Scandolo for their advice and suggestions.

Special thanks for Andrea, Riccardo, Fabia and Sabrina from the secretariat.

Bibliography

[1] H. S. Chan and K. A. Dill, Physics Today **46**, 24 (1993).

[2] T. E. Creighton *Proteins: Structures and Molecular Properties*, W. H. Freeman & Company, New York (1993).

[3] A. M. Gutin, V. I. Abkevich and E. I. Shackanovich, Proc. Natl. Acad. Sci. USA **92**, 1282 (1995).

[4] A. Sali, E. I. Shackanovic and M. Karplus, Nature **369**, 248 (1994).

[5] A. Sali, E. I. Shackanovic and M. Karplus, Journal of Molecular Biology **235**, 1614 (1994).

[6] J. Brygelson, J. N. Onuchic, J. N. Socci and P. G. Wolynes, Proteins: Structures, Function and Genetics **21**, 167 (1995).

[7] E. I. Shakhnovich and A. M. Gutin, Nature **346**, 773 (1990).

[8] R. Livi, M. Pettini, S. Ruffo, M. Sparpaglione and A. Vulpiani, Phys. Rev. A **31**, 1039 (1985).

[9] R. Livi, M. Pettini, S. Ruffo and A. Vulpiani, Phys. Rev. **A 31**, 2740 (1985).

[10] M. Pettini and M. Landolfi, Phys. Rev. **A 41**, 768 (1990).

[11] M. Pettini and M. Cerruti-Sola, Phys. Rev. **A 44**, 768 (1991).

[12] M. Pettini, Phys. Rev **E 47**, 828 (1993).

[13] L. Casetti and M. Pettini, Phys. Rev **E 48**, 4320 (1993).

[14] M. Cerruti-Sola and M. Pettini, Phys. Rev. **E 51**, 53 (1995).

[15] P. G. de Gennes, *Scaling Concept in Polymer Physics* Cornell University Press, Ithaca and London (1979).

[16] P. G. de Gennes, J. Physique Lett. **36** L55, (1975).

[17] P. G. de Gennes, J. Physique Lett. **39** L299, (1978).

[18] B. Duplantier and H. Saleur, Phys. Rev. Lett. **59**, 539 (1987).

[19] C. Vanderzande, F. Seno and A. L. Stella, Phys. Rev. Lett. **61**, 1520 (1988).

[20] K. De Bell and T. Lookman, Review of Modern Physics **65**, 87 (1993).

[21] A. Baumgärtner, Journal of Chemical Physics **72**, 871 (1980).

[22] G. Iori, E. Marinari and G. Parisi, Journal of Physics **A 24**, 5349 (1992).

[23] G. Iori, E. Marinari, G. Parisi and M. V. Struglia, Physica **A 185**, 98 (1992).

[24] L. Casetti, R. Livi and M. Pettini, Phys. Rev. Lett. **74**, 375 (1995).

[25] L. Casetti, C. Clementi and M. Pettini, Accepted for pubblication in Phys. Rev. **E**.

[26] J. L. Lebowitz, J. K. Percus and L. Verlet, Phys. Rev. **153**, 250 (1967).

[27] G. Benettin, L. Galgani and J. M. Strelcyn, Phys. Rev. **A 14**, 2338 (1976).

[28] P. Flory, *Principles of Polymer Chemistry*, Cornell University Press, Ithaca, N. Y. (1971).

[29] F. Seno and A. L. Stella, J. Phys. France **49**, 739 (1988).

[30] C. Dellago, H. A. Posch and W. G. Hoover, Phys. Rev **E 53**, 1485 (1996).

[31] C. Dellago and H. A. Posch, to be published in Physica A (1996).

[32] C. Clementi, Laurea Thesis in Physics, Università di Firenze, Italy (1995).

[33] L. Caiani, Laurea Thesis in Physics, Università di Firenze, Italy (1995).

[34] L. Casetti, L. Caiani and C. Clementi, Proceedings of the E. Fermi Summer School (1996), *The Physics of Complex System*, to be published in Nuovo Cimento G.

[35] L. Casetti, L. Caiani, C. Clementi and M. Pettini, in preparation.

[36] M. Rasetti, in *Differential geometric methods in mathematical physics*, C. Döbner ed., Springer-Verlag, New York, NY (1989).

[37] L. Casetti, Physica Scripta, **51**, 29 (1995).

[38] H. Goldstein, *Classical Mechanics*, 2nd edition, (Addison-Wesley, New York (1980).

[39] E. Forest and R. D. Ruth, Physica **D 43**, 105 (1990).

[40] R. Abraham, J. E. Marsden, *Foundations of mechanics*, Addison-Wesley, Redwood City, CA (1987).

[41] N. S. Krylov, *Works on the foundations of statistical mechanics*, Princeton University Press, Princeton, NJ (1979).

[42] Y. G. Sinai, *Dynamical Systems II*, Encyclopaedia of Mathematical Sciences, Vol. 2, Springer-Verlag, Berlin (1989).

[43] L. P. Eisenhart, Math. Ann. **30**, 591 (1929).

[44] M. P. do Carmo, *Riemannian geometry*, Birkhäuser, Boston-Basel (1992)