A Model of Directed Walks with Random Self-Interactions.

B. DERRIDA, R. B. GRIFFITHS(*) and P. G. HIGGS

Service de Physique Théorique(**) *de Saclay, CEN-Saclay*
F-91191 Gif-sur-Yvette Cedex

Abstract. – We study a simple one-dimensional model of a folded polymer with random self-inter-
actions. A numerical study of the specific heat shows two regimes: at high temperature, the spe-
cific heat looks smooth and sample independent, whereas at low temperature it possesses many
narrow peaks which change with the sample considered. The model is simple enough to allow for
a full description of its ground states. We obtain numerical evidence for the presence of a «weak
freezing» transition and derive an upper bound for the transition temperature. Heuristic argu-
ments provide an estimate of a critical exponent $\gamma(T)$ which varies continuously with the tem-
perature in the low-temperature phase.

Many models have been proposed describing polymer chains with random self-interactions
(*e.g.* proteins) [1-5]. The purpose of this letter is to consider a very simple one-dimensional
model for the folding of random chains. The model is too simple to represent a real protein.
However, it possesses interesting properties, there is evidence for a phase transition and the
model may be simulated for rather long chains (10^5 monomers). Most other numerical
simulations of random chain models [6-8] are limited to very short chains, making the
comparison with theoretical predictions difficult.

Our model is a chain of N monomers, with the i-th monomer carrying a charge $q_i = \pm 1$.
These charges are chosen at random with equal probability for $+ 1$ and $- 1$, and are then
considered fixed for a given sample. The 2^{N-1} possible configurations $\{x_i\}$ are directed walks
one a one-dimensional lattice: x_i is the site occupied by monomer i, $x_1 = 1$, and $\varepsilon_i = x_{i+1} - x_i$

is either 0 or 1. The energy is

$$E(\{x_i\}) = \sum_{i < j} q_i q_j \delta_{x_i x_j} = \frac{1}{2} \sum_r [\bar{Q}(r)]^2 - \frac{1}{2} N, \qquad (1)$$

where $\delta_{xy} = 1$ if $x = y$, $= 0$ if $x \neq y$; and $\bar{Q}(r)$ is the sum of the charges of the monomers on site r. Clearly this model is over-simplified (zero-range interaction, directedness, no excluded volume). Nevertheless, it contains an important ingredient of several other models in that there is a quenched random sequence of charges.

Since the chain is directed, one can use transfer matrices in a simple way [9-13]. For similar problems with nondirected walks [6,7], the only reliable method seems to be exact enumeration. Using the right side of (1), one can write a recursion relation for the partition function of a chain of i monomers,

$$Z_i = \sum_{k=0}^{i-1} Z_k \exp\left[\frac{1}{2T}\left[i - k - \left(\sum_{j=k+1}^{i} q_j\right)^2\right]\right], \qquad (2)$$

with $Z_0 = Z_1 = 1$, where k is the last monomer on the site immediately preceding the site containing i. From (2), it is clear that the time required to compute Z_N is at most proportional to N^2.

By iterating numerically eq. (2) and its first derivatives [11] for randomly generated sequences we could measure the specific heat (fig. 1). One can see two ranges of temperature: i) at high temperature, the specific heat is smooth and does not depend on the sample; ii) at low temperature, it has peaks due to rearrangements of the chain which are very sample dependent. As N increases (fig. 1), the peaks look denser and steeper, but the range of the low-temperature regime seems to narrow down. We believe that for $N \to \infty$, the specific heat becomes self-averaging (sample independent). However, from the data of fig. 1, it is hard to know if the low-temperature regime indicates the existence of a freezing temperature for $N \to \infty$ or is only due to a crossover from high- to low-temperature behaviours. The average end-to-end distance per monomer, $\langle x_N - x_1 \rangle / N$, not shown here, varies smoothly with temperature when N is large, and shows no visible dependence on size or sample, and no sudden collapse. Here and below, $\langle \rangle$ denotes a thermal average with the q_i held fixed.

However, evidence for the existence of a freezing transition emerges from a study of the probability

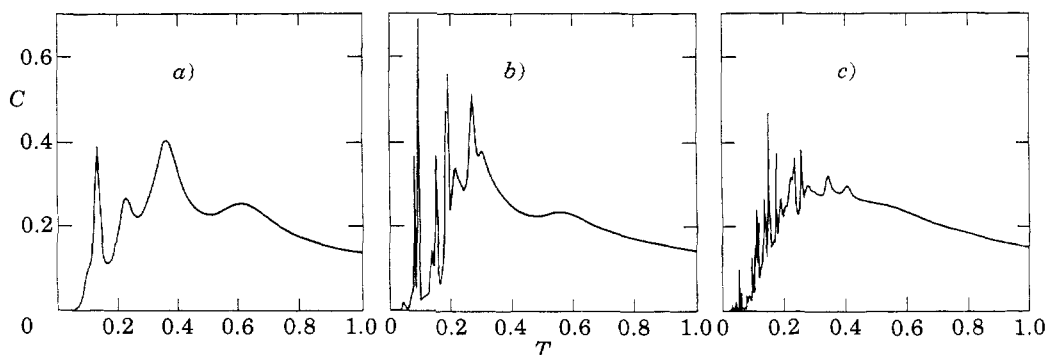$$p_i = \langle x_{i+1} - x_i \rangle = \langle \varepsilon_i \rangle \qquad (3)$$

Fig. 1. – Specific heat *vs.* temperature for a random chain of length $N = 500$ (*a*)), $N = 5000$ (*b*)) and $N = 50\,000$ (*c*)).
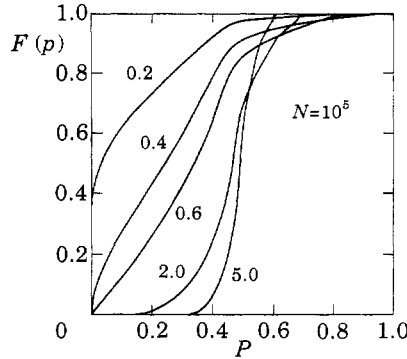
Fig. 2. – The integrated distribution $F(p)$ of the p_i (see eqs. (3) and (4)) at temperatures $T = 0.2$, 0.4, 0.6, 2.0, 5.0. In the low-temperature phase, $F(p)$ seems to be a power law near $p = 0$, whereas in the high-temperature phase $F(p)$ vanishes below a certain $p_{min}(T)$. Three samples of length 10^5 are plotted and the curves are indistinguishable.

that monomers i and $i + 1$ are not on the same site. Numerical simulations show that many p_i have a rapid and sometimes nonmonotonic temperature variation, not unlike the «chaotic» behaviour of the local magnetization m_i in a spin glass [14].

Figure 2 shows the integrated distribution $F(p)$, the fraction of monomers for which p_i is less than p, for 3 samples of 10^5 monomers each, at temperatures $T = 0.2$, 0.4, 0.6, 2.0 and 5.0:

$$F(p) = \frac{1}{N-1} \sum_{1 \leq i \leq N-1} \theta(p - p_i). \tag{4}$$

For increasing N we checked that $F(p)$ reaches a limit and becomes self-averaging. We see from fig. 2 that at high temperatures there seems to be a positive lower bound $p_{min}(T) \leq p_i$ for all i. At low temperatures there appear to be p_i arbitrarily close to zero. Plotting $F(p)$ on a log-log plot suggests that, at least over some temperature range, $F(p)$ is a power law close to $p = 0$,

$$F(p) \sim p^{\gamma(T)}, \tag{5}$$

where $\gamma(T)$ varies with the temperature. We estimate from the data: $\gamma(0.2) \simeq 0.15$; $\gamma(0.4) \simeq 0.5$; $\gamma(0.6) \simeq 1.0$; $\gamma(0.8) \simeq 1.5$; $\gamma(1.0) \simeq 2.1$.

Thus there appears to be a transition temperature T_c above which all the p_i are strictly positive and below which there exist bonds with p_i arbitrary close to zero (for an infinite chain). We call this a «weak freezing» transition since at low temperature there are bonds completely frozen in their collapsed state, although the whole chain is not yet in its ground state (the entropy is nonzero below T_c).

To discuss the model from a more analytical point of view, we need two new quantities: the cumulated charge Q_i defined by

$$Q_i = \sum_{i \leq j \leq i} q_j, \tag{6}$$

and the partition function Y_i of the right part of the chain (Y_i is the partition function of the chain of length $N - i + 1$ with charges $q_i, q_{i+1}, ..., q_N$). The Y_i satisfy a recursion similar to (2) with $Y_N = Y_{N+1} = 1$. It is then easy to check that

$$p_i = Z_i Y_{i+1}/Z_N = Z_i Y_{i+1}/(Z_i Y_{i+1} + R_i), \tag{7}$$

where

$$R_i = \sum_{0 \,\leqslant\, k \,\leqslant\, i\,-\,1} \sum_{i\,+\,2 \,\leqslant\, l \,\leqslant\, N\,+\,1} Z_k Y_l \exp\left[\frac{l - k - 1 - (Q_{l-1} - Q_k)^2}{2T}\right], \tag{8}$$

where k is the last monomer on the site preceeding the site of i and $i + 1$, and l is the first monomer on the site which follows that of i and $i + 1$.

To first order in a $1/T$ expansion, (7) gives

$$p_i = \frac{1}{2}\left(1 + \frac{1}{T} \sum_{k \,\leqslant\, i} \sum_{l \,\geqslant\, i\,+\,1} q_k q_l / 2^{l-k}\right) + O(1/T^2). \tag{9}$$

We see that at high temperature, p_i depends only on the local values of the charges around monomer i.

At $T = 0$, if one restricts our discussion (for simplicity) to the case where the total charge $Q_N = 0$, one can easily describe all the ground states: from (1), it is clear that in a ground state the charge on each site must be zero. The chain has the choice of moving to the next site only when $Q_i = 0$. Hence $p_i = 0$ if $Q_i \neq 0$ and $p_i = 1/2$, when $Q_i = 0$. Since Q_i as a function of i is a random walk, it should vanish a number of times of order $\sqrt{N}$. So the entropy at $T = 0$ increases like $\sqrt{N}$. In striking contrast with the high-T case, a p_i depends on charges which can be arbitrarily far apart along the chain at $T = 0$. In fact the numerical results shown above were done for chains with $Q_N = 0$, and we checked that there was no difference between this case and the general case of unrestricted random sequences.

We did not succeed in finding a proof of the existence of the transition indicated by fig. 2, but we obtained an upper bound of 2.08 for T_c by showing that for $T > 2.08$,

$$p_i \geqslant 1 \Big/ \left[1 + \frac{(\exp[1/2T] + \exp[1/T])^2}{[1 - \exp[1/T]/(1 + \exp[-1/T])]^2}\right] > 0, \tag{10}$$

a lower bound which is independent of i. Let us discuss briefly the origin of (10). Removing from (8) the term $(Q_{l-1} - Q_k)^2$, one gets

$$R_i \leqslant \left(\sum_{0 \,\leqslant\, k \,\leqslant\, i\,-\,1} Z_k \exp\left[\frac{i - k}{2T}\right]\right)\left(\sum_{i\,+\,2 \,\leqslant\, l \,\leqslant\, N\,+\,1} Y_l \exp\left[\frac{l - i - 1}{2T}\right]\right). \tag{11}$$

Now using the fact that $Z_{k-1} \leqslant Z_k$, and $Z_{k-2}(1 + \exp[-1/T]) \leqslant Z_k$ and similar inequalities for the Y_l one gets

$$R_i \leqslant Z_i Y_{i+1} \frac{(\exp[1/2T] + \exp[1/T])^2}{[1 - \exp[1/T]/(1 + \exp[-1/T])]^2}. \tag{12}$$

Substituting (12) into (7) gives the result (10). Notice that the bound for T_c is true for any choice of the q_i.

While we have not been able to prove that a low-temperature «frozen» phase (with arbitrarily small p_i) exists, the following argument is at least plausible. For a pure chain with all charges $q_i = +1$, the largest eigenvalue λ of eq. (2) satisfies

$$1 = \sum_{n \,\geqslant\, 1} \lambda^{-n} \exp[(n - n^2)/2T]. \tag{13}$$

One can show that a chain with $q_i = +1$ for $i \leqslant N/2$ and -1 for $i > N/2$ will collapse at a

temperature $T_c^{\text{pure}} = 1.43$ determined by the equation $\lambda = \exp[1/2T]$ since below T_c^{pure} the free energy of the completely collapsed state is lower than $-T \log \lambda$. Thus, for $T < T_c^{\text{pure}}$ almost all of the monomers are on one site, as the cancellation of positive and negative charges minimizes the energy.

When the q_i are random the polymer will still contain occasional segments of length $2l$ consistent of l positive charges followed by l negative charges, and it is plausible that at low temperature such segments will tend to collapse on a single site, making p_i small for i in the interior of the segment. When l is large, the fraction of monomers belonging to such segments is roughly proportional to 4^{-l}, while the corresponding p_i should be of the order of $\lambda^{2l} \exp[-l/T]$. Assuming such segments give the dominant contribution to $F(p)$ when p is small, $\gamma(T)$ in (5) is the solution (for $T < T_c^{\text{pure}}$) to

$$\sum_{n \geq 1} 2^{n/\gamma} \exp[-n^2/2T] = 1. \tag{14}$$

The $\gamma(T)$ obtained from (14) is in rough agreement ($\gamma(0.2) \simeq 0.3$, $\gamma(0.4) \simeq 0.6$, $\gamma(0.6) \simeq 1.0$, $\gamma(0.8) \simeq 1.7$, $\gamma(1.0) \simeq 2.3$) with our numerical simulations. Discrepancies could be due either to problems in extracting γ from the simulations or to the approximations used to obtain (14). Thus it seems plausible that weak freezing at low temperatures is produced by a mechanism in which some relatively long and relatively rare segments of the polymer chains find it energetically favourable to collapse onto a single site.

The random nature of the charge distribution is essential for this type of transition in that one can show that if the q_i are periodic, the p_i have a strictly positive (independent of i) lower bound at all temperatures. The role played by relatively rare «ordered» segments has analogies in other random systems [15, 16].

In summary, we have obtained rather convincing evidence of a weak freezing transition in this simple model. This includes an upper bound on T_c (which could surely be improved), and a heuristic argument relating the low-temperature exponent $\gamma(T)$ to the presence of some special and relatively rare segments in the polymer. While our methods could probably be extended to certain more complicated models, the directedness of the walk is essential when using a transfer matrix or recursion of the form (2), and thus there is little hope of extending our approach to the undirected case [7].

* * *

REFERENCES

[1] OBUKHOV S. P., *J. Phys. A*, **19** (1986) 3655.
[2] BRYNGELSON J. D. and WOLYNES P. G., *Proc. Natl. Acad. Sci.*, **84** (1987) 7524; *Biopolymers*, **30** (1990) 177.
[3] GAREL T. and ORLAND H., *Europhys. Lett.*, **6** (1988) 307, 597.
[4] SHALKHNOVICH E. I. and GUTIN A. M., *Europhys. Lett.*, **8** (1989) 327; *J. Phys. A*, **22** (1989) 1647.
[5] HIGGS P. G. and JOANNY J. F., *J. Chem. Phys.*, **94** (1991) 1543.
[6] LAU K. F. and DILL K. A., *Macromolecules*, **22** (1989) 3986.
[7] KANTOR Y. and KARDAR M., *Europhys. Lett.*, **14** (1991) 421.
[8] SHAKHNOVICH E., FARZTDINOV G., GUTIN A. M. and KARPLUS M., *Phys. Rev. Lett.*, **67** (1991) 1665.

[9] PRIVMAN V. and SVRAKIĆ N. M., *Lect. Notes Phys.*, **338** (Springer) 1989.

[10] LAURITZEN J. I. and ZWANZIG R., *J. Chem. Phys.*, **52** (1970) 3740.

[11] DERRIDA B. and GOLINELLI O., *Phys. Rev. A*, **41** (1990) 4160.

[12] FOSTER D. P. and YEOMANS J., *Physica A*, **177** (1991) 443.

[13] IGLOI F., *Phys. Rev. A*, **43** (1991) 3194.

[14] BRAY A. J. and MOORE M., *Phys. Rev. Lett.*, **58** (1987) 57.

[15] LIFSHITZ I. M., *Adv. Phys.*, **13** (1964) 483.

[16] GRIFFITHS R. B., *Phys. Rev. Lett.*, **23** (1969) 17.