

Canonical partition functions of freely jointed chains

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

1998 J. Phys. A: Math. Gen. 31 1949

(<http://iopscience.iop.org/0305-4470/31/8/008>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 171.66.16.104

The article was downloaded on 02/06/2010 at 07:23

Please note that [terms and conditions apply](#).

Canonical partition functions of freely jointed chains

Martial Mazars

Laboratoire de Physique Théorique et Hautes Energies[†], Université de Paris XI, Bâtiment 211,
91405 Orsay Cedex, France

Received 28 July 1997, in final form 7 November 1997

Abstract. The freely jointed chain is a simple off-lattice ideal model of a heteropolymer. We show how to compute the canonical partition function of this model exactly for all physical primary structures, degree of polymerization and space dimensions greater than or equal to 2. The canonical partition function of this model of a heteropolymer has an analytical expression involving a complicated multiple hypergeometric function. To obtain some accurate approximations we develop and use the independent motion approximation (IMA).

1. Introduction

In this paper we show that the canonical partition function of one off-lattice model of an ideal polymer, the freely jointed chain, has an exact analytical expression. The freely jointed chain model exhibits one of the principal features of the polymers, namely linear connection of the monomers. This model represents a chain of massive punctual monomers each jointed with two nearest neighbours. In this model, the monomers interact neither with solvent molecules, nor with other monomers of the chain; the unique effective interaction between the monomers is the interaction which imposes the condition that the distance between two nearest neighbours along the chain is constant. Because these interactions are simply holonomic constraints, the canonical partition function of the freely jointed chain can be expressed with the Maxwell–Boltzmann distribution of the ideal gas interpreted over a subset φ_r (the restricted phase space) of the phase space φ of the ideal gas. This very simple model, first studied in the case of homopolymers [1, 2], can be extended to build a model of heteropolymers.

A heteropolymer is firstly characterized by its primary structure, i.e. by the order in which monomers of different kinds are connected. If we define the kind of a monomer appearing in the polymer by a letter from the set $\{A, B, C, \dots\}$, the primary structure of the heteropolymer is then defined by a mapping σ of $[0, N]$ onto $\{A, B, C, \dots\}$ which associates a letter from the set $\{A, B, C, \dots\}$ with the n th position in the chain, giving the kind of the monomer numbered n . Formally the mapping σ is defined by

$$\begin{aligned} \sigma &: [0, N] \longrightarrow \{A, B, C, \dots\} \\ n &\longmapsto X_n. \end{aligned} \tag{1}$$

For instance, the mapping σ corresponding to a homopolymer is defined by: $\forall n \in [0, N], X_n \equiv A$, and for the primary structure of a diblock copolymer as: $\forall n \in [0, N]$, if $n < N_1$ $X_n \equiv A$ and if $n \geq N_1$, $X_n \equiv B$.

[†] Laboratoire associé au Centre National de la Recherche Scientifique (URA 63).

For the ideal heteropolymer described as a freely jointed chain there are only two relevant parameters: the mass of the monomers and the length of the bonds linking the two nearest neighbours. Thus to characterize the primary structure $\{X_n\}_{0 \leq n \leq N}$ in the freely jointed chain model one need use only two sequences: $\{m_n\}_{0 \leq n \leq N}$ the mass sequence and $\{a_n\}_{1 \leq n \leq N}$ the sequence of the bond lengths. Each element of the bond lengths sequence is defined by

$$a_n = a_{X_n X_{n+1}} = a_{\sigma(n)\sigma(n+1)} = a_{\sigma(n+1)\sigma(n)}. \quad (2)$$

The phase space $\varphi_{\{X_n\}}$ of the sequence $\{X_n\}_{0 \leq n \leq N}$ of linearly connected monomers is contained in φ_N , the phase space of the system of disconnected monomers. Thus the canonical partition function of any ideal heteropolymers, as defined previously, is given by

$$Q_{\{X_n\}}^{(d)}(\sigma) = \frac{1}{2!} \frac{1}{h^{(N+1)d}} \int_{\varphi_{\{X_n\}}} \prod_{n=0}^N dr_n dp_n \exp\left(-\frac{\beta}{2} \sum_{n=0}^N \frac{p_n^2}{m_n}\right). \quad (3)$$

The equilibrium properties of the heteropolymer, and in particular the canonical partition function, should be independents of the choice made in the labelling direction of the monomers. According to equation (1), one can formally define the primary structure as $\sigma^*([0, N]) = \{X_n\}_{0 \leq n \leq N}$; on the other hand, the mapping σ' which labels the monomers in the opposite direction ($\sigma'^*([0, N]) = \{X'_n\}_{0 \leq n \leq N}$, with $\forall n, X'_n \equiv X_{N-n}$) should define the same primary structure, i.e. the same heteropolymer. Thus the canonical partition function $Q_{\{X_n\}}^{(d)}(\sigma)$ must be invariant under the transformation $\sigma \rightarrow \sigma'$:

$$Q_{\{X_n\}}^{(d)}(\sigma) = Q_{\{X_n\}}^{(d)}(\sigma') = Q_{\{X_n\}}^{(d)}.$$

For homopolymers, the canonical partition function has been completely computed in a previous paper [3]. In this paper we extend the computation done in [3] and we show that the canonical partition function given by equation (3) can be computed exactly for all degree of polymerization N , for any primary structure $\{X_n\}_{0 \leq n \leq N}$ and for all dimension $d \geq 2$ of the space in which the polymer is embedded.

2. Computation of the canonical partition function

The computation of $Q_{\{X_n\}}^{(d)}$ from equation (3) is difficult because the topological structure of $\varphi_{\{X_n\}}$ is complicated. A parametrization of $\varphi_{\{X_n\}}$ with a set of continuous variables is not easy to construct. For an explicit computation of $Q_{\{X_n\}}^{(d)}$ we can express the integral of equation (3) as the average of a distribution $D\{X_n\}$ in the canonical ensemble of the ideal gas, such that the distribution $D\{X_n\}$ takes into account the holonomic constraints and primary structure of the heteropolymer. Thus equation (3) can be written as

$$Q_{\{X_n\}}^{(d)} = \frac{V^{N+1}}{C_{\{X_n\}}} \left(\prod_{n=0}^N \left(\frac{m_n kT}{2\pi\hbar^2} \right)^{\frac{1}{2}d} \right) \langle D\{X_n\} \rangle_{\varphi_N} \quad (4)$$

with

$$\langle D\{X_n\} \rangle_{\varphi_N} = \frac{\int_{\varphi_N} \prod_{n=0}^N dr_n dp_n D\{X_n\} \exp\left(-\frac{\beta}{2} \sum_{n=0}^N \frac{p_n^2}{m_n}\right)}{\int_{\varphi_N} \prod_{n=0}^N dr_n dp_n \exp\left(-\frac{\beta}{2} \sum_{n=0}^N \frac{p_n^2}{m_n}\right)} \quad (5)$$

where $C_{\{X_n\}}$ is the indiscernability factor of the disconnected gas of monomers. If there are p different kinds of monomers in the polymer, and if for each kind k there are N_k monomers in the primary structure, then $C_{\{X_n\}} = \prod_{k=1}^p N_k!$.

The distribution $D\{X_n\}$ is defined by the geometrical properties of the molecule. According to the model described in section 1, the first contribution to $D\{X_n\}$ is given by the constant length of the bonds in the linear connection of monomers. These geometrical constraints may be expressed as

$$\prod_{i=1}^N \alpha_i \delta((\mathbf{r}_{i-1} - \mathbf{r}_i)^2 - a_i^2)$$

where the α_i are constants which offset the physical dimension of the Dirac distributions. According to the general laws of classical mechanics, the relative velocity of the monomer i compared with the monomer $i - 1$ (i.e. $\mathbf{v}_{i-1} - \mathbf{v}_i$) should be perpendicular to the bond $\mathbf{r}_{i-1} - \mathbf{r}_i$ [3, 4]. These dynamical constraints produce a second contribution to $D\{X_n\}$ that one can also express in terms of a product

$$\prod_{i=1}^N \lambda_i \delta((\mathbf{v}_{i-1} - \mathbf{v}_i) \cdot (\mathbf{r}_{i-1} - \mathbf{r}_i))$$

where the λ_i play the same role as the α_i . Both sets of constants $\{\alpha_i\}$ and $\{\lambda_i\}$ should be determined by choosing a convention for the computation of the canonical partition function. In this paper we chose to take the partition function of a particle in a box of length L in a space of dimension d equal to the partition function of a particle moving on a hypersphere of radius L in a space of dimension $(d + 1)$. This convention and an explicit computation of the canonical partition function of the diatomic molecule ($N = 1$) gives

$$\alpha_i = a_i^2$$

$$\lambda_i = h \left(\frac{1}{m_{i-1}} + \frac{1}{m_i} \right) = \frac{h}{m_i^*}$$

where m_i^* is the reduced mass of the two-body problem. The expressions of the dynamical constraints via the Dirac distributions given above are in agreement with the classical mechanics only when $d \geq 2$. With both contributions, the distribution $D\{X_n\}$ takes the form

$$D\{X_n\} = \frac{C_{\{X_n\}}}{2!} \left(\prod_{i=1}^N a_i^2 \frac{h}{m_i^*} \right) \prod_{i=1}^N \delta((\mathbf{r}_{i-1} - \mathbf{r}_i)^2 - a_i^2) \delta((\mathbf{v}_{i-1} - \mathbf{v}_i) \cdot (\mathbf{r}_{i-1} - \mathbf{r}_i)). \quad (6)$$

With equations (4) and (5) it is possible to compute $Q_{\{X_n\}}^{(d)}$ in the phase space φ_N by using the distribution $D\{X_n\}$ given by equation (6). The natural variables appearing in the definition of $D\{X_n\}$ are the bond vectors and the velocities defined by

$$\mathbf{u}_n = a_n \hat{\mathbf{u}}_n = \mathbf{r}_{n-1} - \mathbf{r}_n \quad 1 \leq n \leq N$$

$$\mathbf{v}_n = \frac{1}{m_n} \mathbf{p}_n \quad 0 \leq n \leq N. \quad (7)$$

The dynamical constraints are expressed as

$$\delta((\mathbf{v}_n - \mathbf{v}_{n-1}) \cdot a_n \hat{\mathbf{u}}_n) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\Omega_n \exp(j\Omega_n a_n \hat{\mathbf{u}}_n \cdot (\mathbf{v}_n - \mathbf{v}_{n-1})) \quad (8)$$

and the integral over the velocities is a Gaussian integral. The integration over \mathbf{r}_0 is performed by considering that the volume of the box in which the polymer is contained

is very large compared with the radius of giration of the polymer. Because the dynamical constraints induced a non-quadratic term in the Gaussian weight, the computation of the Gaussian integral introduces a coupling between neighbouring bond vectors. Thus, $Q_{\{X_n\}}^{(d)}$ takes the form

$$Q_{\{X_n\}}^{(d)} = Z(\{m_n\}, \{a_n\}; T, V) \int \prod_{n=1}^N d\Omega_n \exp\left(-\frac{1}{2\beta} \sum_{n=1}^N \frac{a_n^2}{m_n^*} \Omega_n^2\right) \\ \times \int \prod_{n=1}^N d\hat{u}_n \delta(\hat{u}_n^2 - 1) \exp\left(\frac{1}{2\beta} \sum_{n=2}^N \frac{1}{m_{n-1}} (a_{n-1} \Omega_{n-1} \hat{u}_{n-1}) \cdot (a_n \Omega_n \hat{u}_n)\right) \quad (9)$$

where

$$Z(\{m_n\}, \{a_n\}; T, V) = \frac{1}{2} \left(\frac{1}{2\pi}\right)^{\frac{1}{2}d(N+1)} \frac{\hbar^N}{\hbar^{(N+1)d}} \left(\frac{m_0}{\beta}\right)^{\frac{1}{2}d} \left(\prod_{n=1}^N \frac{(m_n)^{\frac{1}{2}d}}{m_n^*} \left(\frac{a_n^2}{\beta}\right)^{\frac{1}{2}d}\right) V.$$

This coupling and the integral over the bond vectors are similar to the coupling that appears in the $O(d)$ spin model in one dimension with nearest-neighbour interactions studied by Stanley [5]. The integral over the set $\{\hat{u}_n\}$ in (9) can be computed following the procedure used by Stanley. In analogy with the spin model, we consider that the spin variables are the bond vectors \hat{u}_n and coupling constants defined as $\beta y_n = \Omega_{n-1} \Omega_n a_{n-1} a_n / m_{n-1}$, then the integral related to the canonical partition function of the spin model is

$$U_{\{X_n\}}^{(d)} = \int \prod_{n=1}^N d\hat{u}_n \delta(\hat{u}_n^2 - 1) \exp\left(\frac{1}{2\beta} \sum_{n=2}^N \frac{1}{m_{n-1}} (a_{n-1} \Omega_{n-1} \hat{u}_{n-1}) \cdot (a_n \Omega_n \hat{u}_n)\right). \quad (10)$$

This latter integral is exactly evaluated by the same computations as was done for the spin system [5]; it gives

$$U_{\{X_n\}}^{(d)} = S_d(1) \left(\frac{1}{2}\right)^{(N-1)} (2\pi)^{\frac{1}{2}d(N-1)} \prod_{n=2}^N y_n^{(1-\frac{1}{2}d)} I_{\frac{1}{2}d-1}(y_n) \quad (11)$$

with I_μ the modified Bessel function of the first kind.

Thus the canonical partition function of the model described in section 1 is given by

$$Q_{\{X_n\}}^{(d)} = S_d(1) \left(\frac{m_0}{4\pi^2 \beta \hbar^2}\right)^{\frac{1}{2}d} V \left(\prod_{n=1}^N \left(\left(\frac{1}{2}\right)^{1/(d-1)} \left(\frac{a_n^2}{\beta \hbar^2}\right)\right)^{\frac{1}{2}(d-1)} \frac{(m_n)^{\frac{1}{2}d}}{(m_n^*)^{\frac{1}{2}}}\right) J_{\{X_n\}}^{(d)} \quad (12)$$

where $J_{\{X_n\}}^{(d)}$ is defined as

$$J_{\{X_n\}}^{(d)} = \prod_{n=2}^N \left(\sqrt{\frac{4m_{n-1}^* m_n^*}{m_{n-1}^2}}\right)^{(1-\frac{1}{2}d)} \int \prod_{n=1}^N d\gamma_n \prod_{n=2}^N (\gamma_{n-1} \gamma_n)^{(1-\frac{1}{2}d)} \exp\left(-\sum_{n=1}^N \gamma_n^2\right) \\ \times I_{(\frac{1}{2}d-1)} \left(\sqrt{\frac{4m_{n-1}^* m_n^*}{m_{n-1}^2}} \gamma_{n-1} \gamma_n\right). \quad (13)$$

In what follows we will sometimes refer to the latter integral as the alphabetical integral, because it contains all the non-trivial dependence with the primary structure defined by the alphabetical ordering of the application σ . To cancel the coupling between γ_{n-1} and γ_n we make the transformation

$$\omega_n = \sqrt{\frac{4m_{n-1}^* m_n^*}{m_{n-1}^2}} \gamma_{n-1} \gamma_n \quad \text{for } 2 \leq n \leq N \\ \omega_1 = \gamma_1 \quad (14)$$

which is inverted as

$$\begin{aligned} \gamma_n &= A_n \omega_n & \text{for } 2 \leq n \leq N \\ \gamma_1 &= \omega_1 \end{aligned} \tag{15}$$

where we have set

$$A_n = \frac{(\sqrt{2m_1^*})^{(-1)^{(n-1)}} \prod_{p=1}^{(n-1)} \left(\frac{\omega_{n-p}}{m_{n-p}} \right)^{(-1)^p}}{\sqrt{2m_n^*}}. \tag{16}$$

The Jacobian matrix of this transformation is triangular (A_n depends only on ω_p with $p \leq n - 1$), thus the Jacobian is

$$\det(Ja(\omega_i)) = \prod_{n=2}^N A_n. \tag{17}$$

Using this transformation, the alphabetical integral is expressed as

$$J_{\{X_n\}}^{(d)} = \int d\omega_1 \exp(-\omega_1^2) \int \prod_{n=2}^N d\omega_n A_n \omega_n^{(1-\frac{1}{2}d)} I_{(\frac{1}{2}d-1)}(\omega_n) \exp(-A_n^2 \omega_n^2). \tag{18}$$

From the definition of A_n , we see easily that these factors verify the recurrence relation

$$\frac{1}{A_n^2} = 4 \left(\frac{m_{n-1}^* m_n^*}{m_{n-1}^2} \right) A_{n-1}^2 \omega_{n-1}^2 \quad \text{for } 3 \leq n \leq N. \tag{19}$$

All the correlations induced by the linear connection of the chain is implicitly contains in factor A_N because of the latter recurrence relation. This relation is important because it allows one to obtain an analytical result, which holds all the correlations due to the linear connection of the monomers.

To make the integration over the variables $\{\omega_n\}$, we integrate over ω_N , we use the recurrence relation to express A_N with A_{N-1} and ω_{N-1} , then we integrate over ω_{N-1} and the method is repeated until $N - p = 2$. We begin by extracting the integral over ω_N from equation (18) as

$$f_{X_{N-1}X_N}^{(d)} = A_N \int d\omega_N \omega_N^{(1-\frac{1}{2}d)} I_{(\frac{1}{2}d-1)}(\omega_N) \exp(-A_N^2 \omega_N^2) \tag{20}$$

the Bessel function is expressed with the integral representation

$$I_\mu(x) = \frac{(\frac{1}{2}x)^\mu}{\Gamma(\mu + \frac{1}{2})\Gamma(\frac{1}{2})} \int_{-1}^1 (1-t^2)^{(\mu-\frac{1}{2})} e^{xt} dt$$

thus we obtain

$$f_{X_{N-1}X_N}^{(d)} = \frac{(\frac{1}{2})^{(\frac{1}{2}d-2)}}{\Gamma(\frac{d-1}{2})\Gamma(\frac{1}{2})} A_N \int_{-\infty}^{+\infty} d\omega_N \int_{-1}^1 (1-t_N^2)^{\frac{1}{2}(d-3)} e^{-A_N^2 \omega_N^2 + \omega_N t_N} dt_N. \tag{21}$$

After a Gaussian integration over ω_N we have

$$f_{X_{N-1}X_N}^{(d)} = \frac{(\frac{1}{2})^{\frac{1}{2}d}}{\Gamma(\frac{d-1}{2})} \int_0^1 dt_N (1-t_N^2)^{\frac{1}{2}(d-3)} \exp\left(-\frac{1}{4A_N^2} t_N^2\right). \tag{22}$$

We carry out a series expansion of the exponential and perform the integration over t_N ; taking the recurrence relation (19), it follows that

$$f_{X_{N-1}X_N}^{(d)} = \left(\frac{1}{2}\right)^{(\frac{1}{2}d-1)} \sum_{n=0}^{\infty} \frac{1}{n!} \frac{\Gamma(n + \frac{1}{2})}{\Gamma(n + \frac{d}{2})} \left(\frac{m_{N-1}^* m_N^*}{m_{N-1}^2}\right)^n A_{N-1}^{2n} \omega_{N-1}^{2n}. \tag{23}$$

Substituting $f_{X_{N-1}X_N}^{(d)}$ in $J_{\{X_n\}}^{(d)}$, $f_{X_{N-2}X_{N-1}}$ is computed as

$$\begin{aligned} f_{X_{N-2}X_{N-1}}^{(d)} &= \left(\frac{1}{2}\right)^{\left(\frac{1}{2}d-1\right)} \sum_{n=0}^{\infty} \frac{1}{n!} \frac{\Gamma\left(n+\frac{1}{2}\right)}{\Gamma\left(n+\frac{d}{2}\right)} \left(\frac{m_{N-1}^* m_N^*}{m_{N-1}^2}\right)^n A_{N-1}^{2n} A_{N-1} \\ &\quad \times \frac{\left(\frac{1}{2}\right)^{\left(\frac{1}{2}d-1\right)}}{\Gamma\left(\frac{d-1}{2}\right)\Gamma\left(\frac{1}{2}\right)} \int_{-1}^1 dt_{N-1} (1-t_{N-1}^2)^{\frac{1}{2}(d-3)} \\ &\quad \times \int_{-\infty}^{+\infty} d\omega_{N-1} \omega_{N-1}^{2n} \exp(-A_{N-1}^2 \omega_{N-1}^2 + t_{N-1} \omega_{N-1}). \end{aligned} \quad (24)$$

Carrying out the same computations for $f_{X_{N-2}X_{N-1}}^{(d)}$ as those done for $f_{X_{N-1}X_N}^{(d)}$, it follows that

$$\begin{aligned} f_{X_{N-2}X_{N-1}}^{(d)} &= \left(\frac{1}{2}\right)^{\left(\frac{1}{2}d-1\right)} \sum_{n_1=0}^{\infty} \frac{1}{n_1!} \frac{\Gamma\left(n_1+\frac{1}{2}\right)}{\Gamma\left(n_1+\frac{d}{2}\right)} \left(\frac{m_{N-1}^* m_N^*}{m_{N-1}^2}\right)^{n_1} \\ &\quad \times \left(\frac{1}{2}\right)^{\left(\frac{1}{2}d-1\right)} \sum_{n_2=0}^{\infty} \frac{1}{n_2!} \frac{\Gamma\left(n_1+n_2+\frac{1}{2}\right)}{\Gamma\left(n_2+\frac{d}{2}\right)} \left(\frac{m_{N-2}^* m_{N-1}^*}{m_{N-2}^2}\right)^{n_2} A_{N-2}^{2n_2} \omega_{N-2}^{2n_2}. \end{aligned} \quad (25)$$

In the latter equation, one can note that the dependence of $f_{X_{N-2}X_{N-1}}^{(d)}$ on ω_{N-2} has the same structure as the dependence of $f_{X_{N-1}X_N}^{(d)}$ on ω_{N-1} (see equation (23)), thus we can repeat the method until $N-p=2$, and finally by the definition of $f_{X_0X_1}^{(d)}$ we find

$$J_{\{X_n\}}^{(d)} = f_{X_0X_1}^{(d)} = \int_{-\infty}^{+\infty} d\omega_1 f_{X_1X_2}^{(d)} \exp(-\omega_1^2). \quad (26)$$

This series expansion shows that the alphabetical integral $J_{\{X_n\}}^{(d)}$ is the value that a multiple variables function $g_N^{(d)}(\mathbf{x})$ takes at a point defined by

$$(x_n)_{1 \leq n \leq N-1} = \left(\frac{m_{N-n}^* m_{N-n+1}^*}{m_{N-n}^2} \right)_{1 \leq n \leq N-1}. \quad (27)$$

One can easily see from the definition of the x_i that we recover the result obtained for the homopolymer [3] by letting all masses be equal (i.e. $\forall i, m_i = m \Rightarrow x_i = \frac{1}{4}$). The function $g_N^{(d)}(\mathbf{x})$ which gives the value of the alphabetical integral $J_{\{X_n\}}^{(d)}$ can be formulated with the help of a multiple hypergeometric function ${}^{[N-1]}H_e^{[N]}(\frac{1}{2}; \frac{d}{2}; x_1, \dots, x_{N-1})$ by setting

$$J_{\{X_n\}}^{(d)} = g_N^{(d)}(\mathbf{x}) = g_N^{(d)}(\mathbf{0}) {}^{[N-1]}H_e^{[N]}(\frac{1}{2}; \frac{d}{2}; x_1, \dots, x_{N-1}). \quad (28)$$

The value taken by $g_N^{(d)}$ for $\mathbf{x} = \mathbf{0}$ is

$$g_N^{(d)}(\mathbf{0}) = \left(\frac{1}{2}\right)^{\left(\frac{1}{2}d-1\right)(N-1)} \Gamma\left(\frac{1}{2}\right) \left(\frac{\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)}\right)^{(N-1)} \quad (29)$$

and the multiple hypergeometric function ${}^{[N-1]}H_e^{[N]}$ can be defined as

$$\begin{aligned} {}^{[N-1]}H_e^{[N]}(\frac{1}{2}; \frac{d}{2}; \mathbf{x}) &= \sum_{n_1=0}^{\infty} \frac{1}{n_1!} \frac{(n_1, \frac{1}{2})}{(n_1, \frac{d}{2})} (x_1)^{n_1} \\ &\quad \times \sum_{n_2=0}^{\infty} \frac{1}{n_2!} \frac{(n_1+n_2, \frac{1}{2})}{(n_2, \frac{d}{2})} (x_2)^{n_2} \end{aligned}$$

$$\begin{aligned}
 & \vdots \\
 & \times \sum_{n_p=0}^{\infty} \frac{1}{n_p!} \frac{(n_{p-1} + n_p, \frac{1}{2})}{(n_p, \frac{d}{2})} (x_p)^{n_p} \\
 & \vdots \\
 & \times \sum_{n_{N-2}=0}^{\infty} \frac{1}{n_{N-2}!} \frac{(n_{N-3} + n_{N-2}, \frac{1}{2})}{(n_{N-2}, \frac{d}{2})} (x_{(N-2)})^{n_{N-2}} \\
 & \times \sum_{n_{N-1}=0}^{\infty} \frac{1}{n_{N-1}!} \frac{(n_{N-2} + n_{N-1}, \frac{1}{2})}{(n_{N-1}, \frac{d}{2})} (n_{N-1}, \frac{1}{2}) (x_{(N-1)})^{n_{N-1}} \quad (30)
 \end{aligned}$$

where x_p is the p th component of a vector \mathbf{x} of a Euclidean space of dimension $(N - 1)$, and $(n, a) = \Gamma(n + a) / \Gamma(a)$.

In summary, the canonical partition function of any freely jointed chain of primary structure $\{X_n\}_{0 \leq n \leq N}$ as defined in section 1 is

$$Q_{\{X_n\}}^{(d)} = Q_{\diamond}^{(d)}(V, T) \prod_{n=1}^N \left(\frac{T}{T_{X_{n-1}X_n}} \right)^{\frac{1}{2}(d-1)} [^{N-1}H_e^{[N]}(\frac{1}{2}; \frac{d}{2}; \mathbf{x}) \quad (31)$$

where

$$\left\{ \begin{aligned}
 Q_{\diamond}^{(d)}(V, T) &= \left(\frac{\sqrt{m_0 m_N}}{2\pi \hbar^2} kT \right)^{\frac{1}{2}d} V \\
 T_{X_{n-1}X_n} &= 2 \left(\frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{1}{2})} \frac{(m_n^*)^{\frac{1}{2}}}{(m_n m_{n-1})^{\frac{1}{4}d}} \right)^{2/(d-1)} \frac{\hbar^2}{k a_n^2} \\
 (x_n)_{1 \leq n \leq N-1} &= \left(\frac{m_{N-n}^* m_{N-n+1}^*}{m_{N-n}^2} \right)_{1 \leq n \leq N-1} .
 \end{aligned} \right. \quad (32)$$

With these definitions, it is a straightforward matter to verify that $Q_{\{X_n\}}^{(d)}$ is invariant when the numbering direction of monomers is reversed. This invariance gives a symmetry property to the multiple hypergeometric function $[^{N-1}H_e^{[N]}$, this function is invariant by the transformation $x_i \rightarrow x_{N-i}$. Because the integral (3) is absolutely convergent, and because we have used only analytical computations, the hypergeometric function $[^{N-1}H_e^{[N]}$ should be defined for all physical sequences. This argument is a physical one, but in order to show rigorously that the definition of $[^{N-1}H_e^{[N]}$ by the multiple power series (30) is meaningful, we need to show that for any physical sequence (i.e. $\{m_n\}_{0 \leq n \leq N} \in R^{+*}$) the point defined by equation (27) is in the region of convergence of $[^{N-1}H_e^{[N]}$. In section 3 we demonstrate this by using the Horn criterion on the multiple power series (30).

3. Region of convergence of the multiple hypergeometric series $[^{N-1}H_e^{[N]}$

The function $[^{N-1}H_e^{[N]}$, defined by the series (30), gives the contribution induced by linear connection of masses $\{m_n\}_{0 \leq n \leq N}$ to the canonical partition function of the sequence $\{X_n\}$. One needs to verify that for all physical sequences the value taken by $[^{N-1}H_e^{[N]}$ is defined. In this section we show that this is always true for all sequences $\{X_n\}_{0 \leq n \leq N}$ with the

associated mass sequence $\{m_n\}_{0 \leq n \leq N} \in R^{+*}$. We use the Horn criterion [6, 7] to define the region of convergence of a multiple power series

$$\sum_{k_1, \dots, k_n=0}^{\infty} A_{k_1, \dots, k_n} x_1^{k_1} \cdots x_n^{k_n}. \tag{33}$$

We call the positive quantities r_1, r_2, \dots, r_n the associated radii of convergence of the series, if this series is absolutely convergent when $|x_1| < r_1, \dots, |x_n| < r_n$ and divergent when $|x_1| > r_1, \dots, |x_n| > r_n$. In the absolute space (r_1, r_2, \dots, r_n) the points represented by the r_i lie on a hypersurface H , provided that H is interior to D , the unit hypercube in the positive hyperoctant with one corner coinciding with the origin in r_i -space, and with its edges parallel to the coordinate axis. H has the parametric representation

$$r_i = |\Phi_i(k_1, \dots, k_n)|^{-1} \quad k_i \in R^+$$

with

$$\Phi_i(k_1, \dots, k_n) = \lim_{\epsilon \rightarrow \infty} f_i(\epsilon k_1, \dots, \epsilon k_n)$$

where

$$f_i(k_1, \dots, k_n) = \frac{A_{k_1, \dots, k_{i-1}, k_i+1, k_{i+1}, \dots, k_n}}{A_{k_1, \dots, k_n}}.$$

For the multiple power series defined by equation (30) we have

$$\left\{ \begin{aligned} f_1(n_1, n_2) &= \frac{(n_1 + \frac{1}{2})(n_1 + n_2 + \frac{1}{2})}{(n_1 + 1)(n_1 + \frac{d}{2})} \\ f_i(n_{i-1}, n_i, n_{i+1}) &= \frac{(n_{i-1} + n_i + \frac{1}{2})(n_i + n_{i+1} + \frac{1}{2})}{(n_i + 1)(n_i + \frac{d}{2})} \quad \text{for } 2 \leq i \leq N - 2 \\ f_{N-1}(n_{N-2}, n_{N-1}) &= \frac{(n_{N-2} + n_{N-1} + \frac{1}{2})(n_{N-1} + \frac{1}{2})}{(n_{N-1} + 1)(n_{N-1} + \frac{d}{2})} \end{aligned} \right. \tag{34}$$

thus

$$\left\{ \begin{aligned} \Phi_1(n_1, n_2) &= 1 + \frac{n_2}{n_1} \\ \Phi_i(n_{i-1}, n_i, n_{i+1}) &= \left(1 + \frac{n_{i-1}}{n_i}\right) \left(1 + \frac{n_{i+1}}{n_i}\right) \quad \text{for } 2 \leq i \leq N - 2 \\ \Phi_{N-1}(n_{N-2}, n_{N-1}) &= 1 + \frac{n_{N-2}}{n_{N-1}}. \end{aligned} \right. \tag{35}$$

Setting $t_i = n_{i-1}/n_i$, the parametric representation of the hypersurface H is given by

$$\left\{ \begin{aligned} r_1 &= \frac{t_2}{1 + t_2} \\ r_i &= \frac{t_{i+1}}{(1 + t_i)(1 + t_{i+1})} \quad \text{for } 2 \leq i \leq N - 2 \\ r_{N-1} &= \frac{1}{1 + t_{N-1}}. \end{aligned} \right. \tag{36}$$

The Horn criterion tells us that the series are convergent at the point $(x_i)_{1 \leq i \leq N-1}$, if $\exists (t_i)_{2 \leq i \leq N-1} \in (R^+)^{(N-2)} : (\forall i, 1 \leq i \leq N - 1, x_i < r_i)$.

The point where the convergence of the series (30) is interesting is the point given by equation (27). Let $m_{n+1} = c_n m_n$, thus equation (27) becomes

$$\forall i, \quad 1 \leq i \leq N - 1, \quad x_i = \frac{c_i}{(1 + c_i)(1 + c_{i-1})}. \quad (37)$$

To show that $^{[N-1]}H_e^{[N]}$ is defined for all physical sequences (i.e. $\forall \{m_n\}_{0 \leq n \leq N} \in R^{+*}$), it is enough to show that the proposition

$$(\forall (c_n)_{0 \leq n \leq N-1} \in (R^{+*})^N)(\exists (t_n)_{2 \leq n \leq N-1} \in (R^+)^{N-2}) : (\forall i, 1 \leq i \leq N - 1, x_i < r_i) \quad (38)$$

is true.

In seeking to establish this property, we can see that

$$\text{if } \forall i, 2 \leq i \leq N - 1, \quad t_i = c_{i-1} \quad \text{then} \quad \begin{cases} x_1 < r_1 \\ x_i = r_i \quad \text{for } 2 \leq i \leq N - 2 \\ x_{N-1} < r_{N-1}. \end{cases} \quad (39)$$

The proposition (39) would be sufficient to show that the inertial point of the sequence $\{X_n\}_{0 \leq n \leq N}$ is in the region of convergence of the series, if we were able to define the interior of the region of convergence from equation (36). However, with the parametrized equation of H , one cannot easily define two separate regions of the space with the hypersurface, so it is necessary to verify the Horn criterion.

We define τ_2 by the relation

$$\frac{\tau_2}{1 + \tau_2} = \frac{c_1}{(1 + c_1)(1 + c_0)}.$$

Thus, if $\tau_2 < t_2 < c_1$ then $x_1 < r_1$, and

$$(\forall t_3 \in R^{+*}) \left(x_2 < \frac{t_3}{(1 + t_3)} \frac{1}{(1 + \tau_2)} = r_2 \right).$$

We define τ_3 in the same way with

$$\frac{\tau_3}{(1 + \tau_3)} = \frac{c_2}{(1 + c_2)} \frac{(1 + \tau_2)}{(1 + c_1)}.$$

Thus the construction of τ_i is done inductively as follows:

$$\frac{\tau_i}{(1 + \tau_i)} = \frac{c_{i-1}}{(1 + c_{i-1})} \frac{(1 + \tau_{i-1})}{(1 + c_{i-2})}.$$

If $\forall n, 2 \leq n \leq i, \tau_n < t_n < c_{n-1}$, then we have $\forall n, 2 \leq n \leq i - 1, x_n < r_n$ and by construction

$$(\forall t_{i+1} \in R^{+*}) \left(x_i < \frac{t_{i+1}}{(1 + t_{i+1})} \frac{1}{(1 + \tau_i)} = r_i \right).$$

If $\tau_{N-1} < t_{N-1} < c_{N-2}$ then $\forall n, 2 \leq n \leq N - 2, x_n < r_n$, and

$$r_{N-1} = \frac{1}{1 + t_{N-1}} > \frac{1}{1 + c_{N-2}} > \frac{1}{(1 + c_{N-2})} \frac{c_{N-1}}{(1 + c_{N-1})} = x_{N-1}.$$

Because one can always build the sequence τ_i inductively as shown previously, in summary we have the proposition

$$\text{if } \forall i, 2 \leq i \leq N - 1, \quad \tau_i < t_i < c_{i-1} \quad \text{then} \quad \begin{cases} x_1 < r_1 \\ x_i < r_i \quad \text{for } 2 \leq i \leq N - 2 \\ x_{N-1} < r_{N-1} \end{cases} \quad (40)$$

which establish that the proposition (38) is true.

The analytical expressions for the τ_i are complicated because they depend on $(c_n)_{0 \leq n \leq i-1}$.

This demonstration also shows that the region of convergence of the multiple hypergeometric function ${}^{[N-1]}H_e^{[N]}$ is not the empty set, nor does it reduce to one point, so the definition of ${}^{[N-1]}H_e^{[N]}$ from the multiple series (30) is meaningful.

This result is important because it shows that for all ideal heteropolymers defined as freely jointed chains (see section 1) one can compute the canonical partition function, which is given by equations (30), (31) and (32), analytically.

For small sequences, one can relate ${}^{[N-1]}H_e^{[N]}$ to the well known hypergeometric functions: for instance, for $N = 3, 4$ and 5 , we have

$$\left\{ \begin{array}{l} \{X_0, X_1, X_2\} \rightarrow {}^{[1]}H_e^{[2]}(\frac{1}{2}; \frac{d}{2}; \mathbf{x}) = {}_2F_1(\frac{1}{2}, \frac{1}{2}; \frac{d}{2}; x_1) \\ \{X_0, X_1, X_2, X_3\} \rightarrow {}^{[2]}H_e^{[3]}(\frac{1}{2}; \frac{d}{2}; \mathbf{x}) = F_2(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}; \frac{d}{2}, \frac{d}{2}; x_1, x_2) \\ \{X_0, X_1, X_2, X_3, X_4\} \rightarrow {}^{[3]}H_e^{[4]}(\frac{1}{2}; \frac{d}{2}; \mathbf{x}) = F_K(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}; \frac{d}{2}, \frac{d}{2}, \frac{d}{2}; x_1, x_2, x_3) \end{array} \right. \quad (41)$$

where the x_i are given by equation (32), F_2 is the second Appell double hypergeometric function and F_K the triple hypergeometric function of Lauricella–Saran.

The analytical expression of ${}^{[N-1]}H_e^{[N]}$ is complicated and not easy to use in applications. In section 4 we give a non-perturbative iterative scheme of approximations which gives very good estimates of the function ${}^{[N-1]}H_e^{[N]}$.

4. Independent motions approximation (IMA)

Since we know the multiple power series expansion of ${}^{[N-1]}H_e^{[N]}$, we know all the values taken by the derivatives of ${}^{[N-1]}H_e^{[N]}$ at the origin, and in particular we know the restrictions of this function to each axis exactly. These restrictions according to equation (30) are given by

$$\begin{aligned} 1 \leq i \leq N - 2 & \quad {}^{[N-1]}H_e^{[N]}(\frac{1}{2}; \frac{d}{2}; x_i \hat{e}_i) = {}_1F_1(\frac{1}{2}; \frac{d}{2}; x_i) \\ i = N - 1 & \quad {}^{[N-1]}H_e^{[N]}(\frac{1}{2}; \frac{d}{2}; x_{N-1} \hat{e}_{N-1}) = {}_2F_1(\frac{1}{2}, \frac{1}{2}; \frac{d}{2}; x_{N-1}). \end{aligned}$$

Thus a crude approximation of this function is given by

$${}^{[N-1]}H_e^{[N]}(\frac{1}{2}; \frac{d}{2}; x_1, \dots, x_{N-1}) \sim {}_2F_1(\frac{1}{2}, \frac{1}{2}; \frac{d}{2}; x_{N-1}) \left(\prod_{i=1}^{N-2} {}_1F_1(\frac{1}{2}; \frac{d}{2}; x_i) \right).$$

This approximation has two major weaknesses: the symmetry property of ${}^{[N-1]}H_e^{[N]}$ is not fulfilled and this procedure does not allow one to find a general scheme to obtain more accurate approximations [3].

To overcome both weaknesses, we use a diagrammatic method which simplifies the analytical computations and permits one to describe a non-perturbative scheme of approximation for ${}^{[N-1]}H_e^{[N]}$.

We define the following diagrammatic rules involving sites and lines. For sites with a multiplicity n :

$$\frac{1}{n!} \frac{1}{(n, \frac{d}{2})} x^n \equiv \begin{array}{c} | \quad x \\ | \\ \bigcirc \\ | \\ n \end{array} \quad (42a)$$

for propagators between two neighbouring sites:

$$(n + p, \frac{1}{2}) \equiv \begin{array}{c} \text{---} \\ | \\ n \end{array} \begin{array}{c} \text{---} \\ | \\ p \end{array} \equiv \begin{array}{c} \text{---} \\ | \\ p \end{array} \begin{array}{c} \text{---} \\ | \\ n \end{array} \quad (42b)$$

and for the ending sites:

$$\begin{array}{c} | \\ | \\ \bullet \\ | \\ n \end{array} \equiv (n, \frac{1}{2}) \begin{array}{c} | \\ | \\ \circ \\ | \\ n \end{array} \equiv \begin{array}{c} | \\ | \\ \circ \\ | \\ 0 \end{array} \begin{array}{c} | \\ | \\ \circ \\ | \\ n \end{array} \quad (42c)$$

With the above diagrammatic rules, the multiple hypergeometric function $^{[N-1]}H_e^{[N]}$ can be written as

$$^{[N-1]}H_e^{[N]}(\frac{1}{2}; \frac{d}{2}; \mathbf{x}) = \sum_{n_1=0}^{\infty} \cdots \sum_{n_{N-1}=0}^{\infty} \begin{array}{c} | \\ | \\ \bullet \\ | \\ n_1 \end{array} \begin{array}{c} | \\ | \\ \circ \\ | \\ n_2 \end{array} \cdots \begin{array}{c} | \\ | \\ \circ \\ | \\ n_{N-2} \end{array} \begin{array}{c} | \\ | \\ \bullet \\ | \\ n_{N-1} \end{array} \quad (43)$$

The technical difficulties encountered in the computation of $^{[N-1]}H_e^{[N]}$ are caused by the coupling between neighbouring sites. In the diagrammatic representation, the coupling enters the computation via propagators. The natural way to obtain a first approximation of this function is to cancel the coupling by disconnecting the sites. This can be done by cutting the propagators. One should remember that the coordinates x_i contains all the non-trivial information about the primary structure. The arbitrary choice of the labelling direction of these coordinates during the computation of the canonical partition function, which favours one direction of labelling for the monomers of the sequence, does not affect the indistinguishability of both direction of labelling in the canonical ensemble because of a special symmetry property of $^{[N-1]}H_e^{[N]}$. Therefore, the approximation that we obtain by cutting the propagators must have this symmetry property.

The diagrammatic rules (42c) and (42b) show how the cutting works. When we cut the propagator on the left, we take the value 0 for the multiplicity n on the left of the propagator and we keep the true value of n for the site on the left of the propagator; for the value p of the multiplicity on the right of the propagator, we keep the true value of p for both the right side of the propagator and the site. Thus, when one cuts all the propagators of (43) on the left, we obtain

$$^{[N-1]}H_e^{[N]}(\frac{1}{2}; \frac{d}{2}; \mathbf{x})|_{1(l)} = \prod_{i=1}^{(N-2)} \left(\sum_{n=0}^{\infty} \begin{array}{c} | \\ | \\ \bullet \\ | \\ n \end{array} \begin{array}{c} | \\ | \\ \circ \\ | \\ n \end{array} \right) \times \left(\sum_{n_{N-1}=0}^{\infty} \begin{array}{c} | \\ | \\ \circ \\ | \\ 0 \end{array} \begin{array}{c} | \\ | \\ \bullet \\ | \\ n_{N-1} \end{array} \right) \quad (44a)$$

Because of the contribution of x_{N-1} in equation (44a), the approximation $^{[N-1]}H_e^{[N]}(\frac{1}{2}; \frac{d}{2}; \mathbf{x})|_{1(l)}$ do not have the symmetry property of $^{[N-1]}H_e^{[N]}$. To restore this symmetry one has to cut also the propagators on the right:

$${}^{[N-1]}H_e^{[N]}(\frac{1}{2}; \frac{d}{2}; \mathbf{x})|_{1(r)} = \left(\sum_{n_1=0}^{\infty} \begin{array}{c} | \quad x_1 \\ | \\ \text{---} \bullet \\ | \\ n_1 \end{array} \right) \times \prod_{i=2}^{(N-1)} \left(\sum_{n=0}^{\infty} \begin{array}{c} | \quad x_i \\ | \\ \bullet \\ | \\ n \end{array} \right) \quad (44b)$$

and take the geometrical average of the expressions given by equations (44a) and (44b). Then we obtain a first approximation of ${}^{[N-1]}H_e^{[N]}$ as

$$\begin{aligned} {}^{[N-1]}H_e^{[N]}(\frac{1}{2}; \frac{d}{2}; \mathbf{x})|_1 &= \sqrt{{}^{[N-1]}H_e^{[N]}(\frac{1}{2}; \frac{d}{2}; \mathbf{x})|_{1(r)} {}^{[N-1]}H_e^{[N]}(\frac{1}{2}; \frac{d}{2}; \mathbf{x})|_{1(l)}} \\ &= \sqrt{\frac{{}_2F_1(x_1){}_2F_1(x_{N-1})}{{}_1F_1(x_1){}_1F_1(x_{N-1})}} \prod_{i=1}^{N-1} {}_1F_1(x_i) \end{aligned} \quad (45)$$

where we use the notation ${}_1F_1(x) = {}_1F_1(\frac{1}{2}; \frac{d}{2}; x)$ and ${}_2F_1(x) = {}_2F_1(\frac{1}{2}, \frac{1}{2}; \frac{d}{2}; x)$. For $N = 2$, one can verify that equation (45) is exactly the same as equation (41).

Because by cutting the propagators we cancel the coupling between sites, this approximation is equivalent to considering the movement of a monomer in the chain as being influenced only by the inertia of its closest neighbours in the chain. This is the reason we call this approximation the independent motions approximation (IMA).

The computations done to obtain equation (45) allow one to build a non-perturbative iterative scheme of approximations for ${}^{[N-1]}H_e^{[N]}$. In the previous computation we had cut all the propagators of the diagrammatic representation (43), thus this approximation is called the first-order IMA. The iterative scheme consist in redefining propagators in analogy with the conventional renormalization of the full propagator. For the second-order IMA we define the double propagator by

$$\sum_{n_2=0}^{\infty} \begin{array}{c} | \quad x_1 \quad | \quad x_2 \quad | \quad x_3 \\ | \quad | \quad | \\ \text{---} \circ \text{---} \circ \text{---} \circ \\ | \quad | \quad | \\ n_1 \quad n_2 \quad n_3 \end{array} \equiv \begin{array}{c} | \quad x_1 \quad | \quad x_3 \\ | \quad | \\ \text{---} \bullet \text{---} \bullet \\ | \quad | \\ n_1 \quad n_3 \end{array} \quad (46)$$

The latter relation allows one to introduce the diagrammatic rule for the double propagator

$$\frac{x}{n \quad p} \equiv {}_2F_1(\frac{1}{2} + n, \frac{1}{2} + p; \frac{d}{2}; x). \quad (47)$$

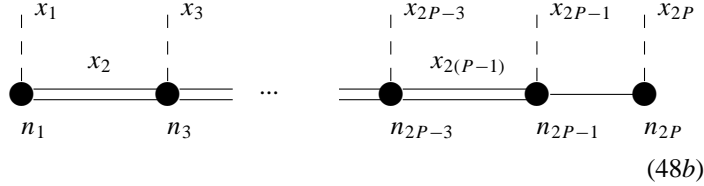
One can use double propagators to give a new diagrammatic representation of the function ${}^{[N-1]}H_e^{[N]}$, provided that we distinguish the N even case from the N odd case. For N even ($N = 2P$), the summation over all the n_{2p} with $1 \leq p \leq (P - 1)$ in (43) gives

$${}^{[2P-1]}H_e^{[2P]}(\frac{1}{2}; \frac{d}{2}; \mathbf{x}) = \sum_{n_1=0}^{\infty} \sum_{n_3=0}^{\infty} \dots \sum_{n_{2P-1}=0}^{\infty}$$

$$\begin{array}{c} | \quad x_1 \quad | \quad x_3 \quad \dots \quad | \quad x_{2P-3} \quad | \quad x_{2P-1} \\ | \quad | \quad | \quad \dots \quad | \quad | \\ \text{---} \bullet \text{---} \bullet \quad \dots \quad \text{---} \bullet \text{---} \bullet \\ | \quad | \quad | \quad \dots \quad | \quad | \\ n_1 \quad n_3 \quad \dots \quad n_{2P-3} \quad n_{2P-1} \end{array} \quad (48a)$$

while for N odd we have

$$[2P]H_e^{[2P+1]}(\frac{1}{2}; \frac{d}{2}; \mathbf{x}) = \sum_{n_1=0}^{\infty} \sum_{n_3=0}^{\infty} \cdots \sum_{n_{2P-1}=0}^{\infty} \sum_{n_{2P}=0}^{\infty}$$



By cutting the double propagators with the same rules as for the single propagators, and taking the geometric average to restore the symmetry, we obtain the following expressions for the second-order IMA.

For N even

$$[2P-1]H_e^{[2P]}(\frac{1}{2}; \frac{d}{2}; \mathbf{x})|_2 = \sqrt{{}_1F_1(x_{11}){}_1F_1(x_{2P-1})} \prod_{i=1}^{(P-1)} \sqrt{{}_2F_2(x_{2i}, x_{2i+1}){}_2F_2(x_{2i+1}, x_{2i+2})}$$

and for N odd

$$[2P]H_e^{[2P+1]}(\frac{1}{2}; \frac{d}{2}; \mathbf{x})|_2 = \sqrt{{}_1F_1(x_{11}){}_1F_1(x_{2P}){}_2F_1(x_{12}){}_2F_1(x_{2P})} \times \prod_{i=1}^{(P-1)} \sqrt{{}_2F_2(x_{2i}, x_{2i+1}){}_2F_2(x_{2i+1}, x_{2i+2})}$$

A generalization of this procedure is straightforward; the p th-order IMA is obtained by the following steps.

- (i) Choose a direction for the labelling of the sites.
- (ii) Group the sites by blocks of $p + 1$ sites according to the chosen direction of labelling, i.e. the first block contained the sites labelled $\{n_1, n_2, \dots, n_{p+1}\}$, the second the sites labelled $\{n_{p+2}, n_{p+3}, \dots, n_{2p+2}\}$, and so on (if the diagram is composed by a number of sites that is not divisible by $p + 1$, we put in the last block less than $p + 1$ sites).
- (iii) Define a p -uple propagator as done with equation (46) for the double propagator and define a new diagrammatic rule for the p -uple propagator.
- (iv) Then give the new diagrammatic representation of $^{[N-1]}H_e^{[N]}$ and cut all the p -uple propagators on the right to obtain $^{[N-1]}H_e^{[N]}|_{p(r)}$ and next cut them on the left to obtain $^{[N-1]}H_e^{[N]}|_{p(l)}$.
- (v) Take the geometrical average:

$$^{[N-1]}H_e^{[N]}|_{p(+)} = \sqrt{^{[N-1]}H_e^{[N]}|_{p(r)} \ ^{[N-1]}H_e^{[N]}|_{p(l)}} \tag{51}$$

(vi) Take the other direction of labelling and repeat steps (ii) to (v) to obtain $^{[N-1]}H_e^{[N]}|_{p(-)}$.

(vii) The p th-order IMA of $^{[N-1]}H_e^{[N]}$ is then given by the geometrical average:

$$^{[N-1]}H_e^{[N]}|_p = \sqrt{^{[N-1]}H_e^{[N]}|_{p(-)} \ ^{[N-1]}H_e^{[N]}|_{p(+)}} \tag{52}$$

One can observe that with this approximation scheme the $(N - 1)$ th-order IMA furnishes exactly ${}^{[N-1]}H_e^{[N]}$.

To test the numerical accuracy of this approximation scheme when $p < N - 1$ we compare the values of the alphabetical integral predicted by using IMA with numerical values obtained with a numerical computation using the Monte Carlo algorithm for a given primary structure. In equation (9) in section 2 we performed the integration over the bond vectors before the integration over the Fourier transform parameters $\{\Omega_n\}$. One can see that the integral over the set $\{\Omega_n\}$ is a Gaussian integral. If one performs the integral over the set $\{\Omega_n\}$ before the integral over the bond vectors, one finds that the alphabetical integral of the homopolymer can be written [2] as

$$J_{\text{homo}}^{(d)}(N) = \int_{[0,1]^{(N-1)}} \prod_{n=1}^{N-1} dt_n (1 - t_n^2)^{\frac{1}{2}(d-3)} \left(\frac{1}{\det M^{(N)}(\{t_n\})} \right)^{\frac{1}{2}}. \quad (53)$$

Rigorously, with the method described above, equation (53) is valid only for $d \geq 2$ and d integer. However, it is shown [3, section V] that this relation holds for non-integer value of d with $d \geq 2$. In [3] we performed a numerical computation of the latter integral using importance sampling with Monte Carlo algorithm for 28 values of d taken between 2 and 8, and for $N = 2$ to 500. A very good fit of the numerical results for $J_{\text{homo}}^{(d)}(N)$ takes the form

$$J_{\text{homo}}^{(d)}(N) = g_N^{(d)}(\mathbf{0}) B^{(d)}(A^{(d)})^N \quad (54)$$

where $B^{(d)}$ and $A^{(d)}$ are numerical values. Because the integral (3) is absolutely convergent, the final result must be independent of the order in which we perform the integrations. Thus these numerical evaluations give the value

$${}^{[N-1]}H_e^{[N]}(\frac{1}{2}; \frac{d}{2}; \frac{1}{4}) \simeq B^{(d)}(A^{(d)})^N. \quad (55)$$

On the other hand, the first-order IMA for the homopolymer gives

$${}^{[N-1]}H_e^{[N]}(\frac{1}{2}; \frac{d}{2}; \frac{1}{4})|_1 = {}_2F_1(\frac{1}{4})({}_1F_1(\frac{1}{4}))^{N-2} \quad (56)$$

therefore the numerical values predicted by first-order IMA are

$$A^{(d)} \simeq {}_1F_1(\frac{1}{4}) \quad \text{and} \quad B^{(d)} \simeq \frac{{}_2F_1(\frac{1}{4})}{({}_1F_1(\frac{1}{4}))^2}. \quad (57)$$

The values predicted by the second-order IMA are, for N even

$$A^{(d)} \simeq \sqrt{F_2(\frac{1}{4}, \frac{1}{4})} \quad \text{and} \quad B^{(d)} \simeq \frac{{}_1F_1(\frac{1}{4})}{F_2(\frac{1}{4}, \frac{1}{4})} \quad (58)$$

and for N odd

$$A^{(d)} \simeq \sqrt{F_2(\frac{1}{4}, \frac{1}{4})} \quad \text{and} \quad B^{(d)} \simeq \frac{{}_1F_1(\frac{1}{4}){}_2F_1(\frac{1}{4})}{(F_2(\frac{1}{4}, \frac{1}{4}))^{\frac{3}{2}}}. \quad (59)$$

The agreement between the numerical values and the values predicted by the IMA procedure is very good [3]. For instance, for $d = 3$ one finds $A^{(d)} \simeq 1.051(3)$ and we have ${}_1F_1(\frac{1}{4}) \simeq 1.089\,97$ and $\sqrt{(F_2(\frac{1}{4}, \frac{1}{4}))} \simeq 1.050\,03$, and for $d = 4$, $A^{(d)} \simeq 1.035(8)$ and ${}_1F_1(\frac{1}{4}) \simeq 1.066\,62$ and $\sqrt{(F_2(\frac{1}{4}, \frac{1}{4}))} \simeq 1.036\,09$.

In summary, the analysis of the numerical results obtained for 13972 partition functions computed in [3] shows that the accuracy of the first-order IMA is between 1–5%, while the second-order IMA improves the accuracy [8].

5. Conclusion

In this paper we had shown how to compute the canonical partition function of freely jointed chains exactly for any primary structures as defined in section 1. This computation take the classical Hamiltonian of the system as a starting point; thus the canonical partition function given by equations (30), (31) and (32) results from a microscopic model, and is valid for the large-volume limit (i.e. the free freely jointed chain). As shown in section 3, this result is analytic for any applications σ which define the primary structure of this ideal heteropolymer model.

This analytical result shows also that others models of ideal polymers are good approximations of the freely jointed chains. For instance the random walk on a regular lattice is a good approximation of the ideal homopolymer defined as a freely jointed chain. The number of conformations \mathfrak{R}_N of a random walk made of N steps on a regular lattice with z as coordination number is equivalent to the canonical partition function. A classical result gives

$$\mathfrak{R}_N = z^N. \tag{60}$$

This result agrees with the canonical partition function of equation (31) if we make the approximation ${}^{[N-1]}H_e^{[N]}(\frac{1}{2}; \frac{d}{2}; \frac{1}{4}) \sim 1$, which is a good approximation according to equation (55) and the numerical values of section 4 (see also [3]). To the accuracy of this approximation, one can consider that the freely jointed chain is a random walk on a regular lattice with an effective coordination number z_{eff} defined by

$$z_{\text{eff}} = \left(\frac{T}{T_0}\right)^{\frac{1}{2}(d-1)} \quad \text{with } T_0 = 2\left(\frac{1}{\sqrt{2}} \frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{1}{2})}\right)^{2/(d-1)} \frac{\hbar^2}{mka^2}. \tag{61}$$

One can also verify that the critical exponents γ and ν of the freely jointed chain, respectively defined by $Q_{\{A\}}^{(d)} = Q_0 N^{\gamma-1} z^N$ and $\langle (r_0 - r_N)^2 \rangle \propto a^2 N^{2\nu}$, are respectively exactly equal to the values 1 and $\frac{1}{2}$ of the random walks on regular lattice [3].

Another class of model of ideal polymers is given by the matrix transfer formalism applied to the end to end probability distribution considered as a Green function and often taken as a Gaussian distribution [9, 10]. In this formalism the Green function satisfies a diffusion-type equation and one can build an analogy with quantum mechanics; thus the Green function of the polymer can be represented as a bilinear series

$$G \left(\begin{array}{c|c} 0 & N \\ \mathbf{x}_0 & \mathbf{x}_N \end{array} \right) = \sum_m \Lambda_m^N \psi_m^+(\mathbf{x}_0) \psi_m(\mathbf{x}_N)$$

where the ψ_m are the eigenfunction and Λ_m the eigenvalues of the transfer operator, labelled in decreasing order. The partition function is then given by

$$Z_N = \int D\mathbf{x}_0 D\mathbf{x}_N G \left(\begin{array}{c|c} 0 & N \\ \mathbf{x}_0 & \mathbf{x}_N \end{array} \right).$$

In the case of a discrete spectrum, one can use the ground-state dominance approximation which gives

$$Z_N \sim \text{cte.} \Lambda_0^N \tag{62}$$

which is valid until $N \log(\Lambda_0/\Lambda_1) \gg 1$. Again we find that this model agrees with the freely jointed chain if we make the approximation ${}^{[N-1]}H_e^{[N]}(\frac{1}{2}; \frac{d}{2}; \frac{1}{4}) \sim 1$ and take

$$\Lambda_0 = \left(\frac{T}{T_0}\right)^{\frac{1}{2}(d-1)}. \tag{63}$$

Because the result obtained for the freely jointed chain is analytical, one can also compute the other eigenvalues Λ_m by using the IMA.

If one describes the polymers as Brownian paths (see [11, chapter X]), the analytical results obtained for this microscopic model might permit one to find the continuum limit used in these descriptions explicitly. In particular, these results might be useful to understand precisely how the microstructure of the polymer chain cancels the ‘ultraviolet’ divergences which appear in the perturbative expansion. This very interesting point had not been studied here, but will be the subject of a future paper.

The dimension of the physical space in which the polymer is embedded is necessarily greater than or equal to 2. This restriction is induced by the expression of the dynamical constraints obtained from the geometrical constraints [4]. The analytical continuation to all real values of d greater than 2 is done implicitly in all the computations of the paper. In [3] and [8] we showed that the IMA defined in section 4 is still a good approximation for non-integer values of d . This model for an ideal heteropolymer seems to be another model which generalizes the random walk to spaces with non-integer dimension. Comparison with the recently proposed model, *the D-dimensional spherically symmetric random walks in non-integer dimension* [12] has yet to be made. It is possible that both models are complementary, and eventually equivalent. The weakness of the freely jointed chain model as a random walk in non-integer dimension is that this model is defined only for $d \geq 2$; its strength is in the analytical expression obtained from the fundamental principles of statistical physics.

References

- [1] Kramers H 1946 *J. Chem. Phys.* **14** 415
- [2] Fixman M 1974 *Proc. Natl Acad. Sci. USA* **71** 3050
- [3] Mazars M 1996 *Phys. Rev. E* **53** 6297
- [4] Edwards S F and Goodyear A G 1972 *J. Phys. A: Math. Gen.* **5** 965
Edwards S F and Goodyear A G 1972 *J. Phys. A: Math. Gen.* **5** 1188
- [5] Stanley H E 1969 *Phys. Rev.* **179** 570
- [6] Exton H 1976 *Multiple Hypergeometric Functions and Applications* (Chichester: Ellis Horwood)
- [7] Horn J 1889 *Math. Ann.* **34** 544
- [8] Mazars M 1997 *Etudes Théoriques et Simulations Numériques de Systèmes de Molécules Linéaires Flexibles*
Thesis No 4693, Université de Paris-Sud XI
- [9] Grosberg A Y and Khokhlov A R 1994 *Statistical Physics of Macromolecules* ed R Larson and P A Pincus
(New York: American Institute of Physics)
- [10] Grosberg A Y and Shakhnovich E I 1986 *Zh. Eksp. Teor. Fiz.* **91** 2159 (Engl. Transl. 1986 *Sov. Phys.-JETP*
64 1284)
- [11] des Cloiseaux J and Jannink G 1987 *Polymères en Solution, leurs Modélisation et leurs Structure* (Paris: Editions de Physique)
- [12] Bender C M, Boettcher S and Mead L R 1994 *J. Math. Phys.* **35** 368
Bender C M, Boettcher S and Moshe M 1994 *J. Math. Phys.* **35** 4941