Journal of Statistical Physics, Vol. 9, No. 3, 1973

A New Analytic Solution for the Zwanzig-Lauritzen Model of Polymer Chain Folding

K. Sture J. Nordholm[1]

Received December 11, 1973; revised May 22, 1973

A two-dimensional model of polymer chain folding invented by Zwanzig and Lauritzen is here studied using a grand ensemble and transfer matrix method. Due to the character of the model, there are no extensive parameters in the grand ensemble and the dispersion in system size is large, raising doubts about the validity and usefulness of the ensemble. We find it possible to define a thermodynamic limit such that it leads to near equivalence between the canonical and grand ensembles in the limit of large systems. The transfer matrix in this case is a nonlocal operator on a space of L_2 functions, and the eigenvalue equation is a homogeneous Fredholm integral equation of the second kind which can be completely solved in terms of Bessel functions. The grand partition function can then be expressed as a sum of powers of the known eigenvalues. It is an easy matter to reproduce the second-order phase transition in the canonical ensemble found in the original work on the model. The investigation is extended to yield the probability densities describing the length of a segment and the correlations among segments. The concept of a local width of the folded chain is found to break down at higher temperatures, while critical correlations are characterized by infinite range, as expected. Apart from physical and methodological implications, the new solution provides striking illustrations of some basic ideas concerning phase transitions.

KEY WORDS: Polymer model; chain folding; phase transition; grand partition function; transfer matrix; integral equation; eigenvalue spectrum; fluctuations; correlations; end effects.

---

[1] The James Franck Institute, The University of Chicago, Chicago, Illinois.

1. INTRODUCTION

We shall here be concerned with a mathematical model of a polymer molecule invented by Zwanzig and Lauritzen(1) in an effort to understand the phenomenon of polymer crystallization by chain folding. The polymer molecule is restricted to a plane. It is assumed capable of creating 180° folds at a given cost in energy. The segments on either side of a particular fold attract each other in such a way that the energy gained is proportional to the length of the shorter of the two segments.

In two papers Zwanzig and Lauritzen(1,2) considered this simple model, allowing the chain segments first a continuous range of lengths, then a discrete range, as is reasonable in view of the submolecular structure of the polymer. Their work was in both cases based on the evaluation of the canonical partition function $Q(L)$, which is just the sum of the Boltzmann factors $e^{-\beta E_j(L)}$ for all configurations of a polymer molecule of total length L. It was shown that the model led to a second-order phase transition between an extended mode of infinite segment lengths below T_c and a folded mode of finite segment lengths above T_c.

From a physical point of view the transition can be said to result from the competition between configurational and energy contributions to the partition function. For low temperatures the Boltzmann factor varies rapidly with the energy $E_j(L)$ and the configurations with very long segments minimizing the energy dominate the sum. As temperature increases, the importance of energy decreases, sheer number of configurations become correspondingly more important, and the average segment length decreases.

The evaluation of the canonical partition function $Q(L)$ is not a simple matter, but Zwanzig and Lauritzen were able to show that in the thermodynamic limit of infinite chain length the average number of folds per unit length vanished below a temperature T_c and was finite above it. They found that a second-order phase transition took place at this temperature, the transition being characterized by an inverse-square-root singularity in the specific heat. This held true whether the segment length was taken to be a continuous or a discrete variable. The only qualitative difference between the two variations of the model occurred when the fold energy was taken to be negative, but not too small. Then the continuous model was found to have two transitions. At high temperatures the chain would be in the folded mode, for an intermediate range it would be in an extended mode, but then at some low temperature the chain would again go into a folded mode. This is physically reasonable since for the case of negative fold energy increasing the number of folds in the chain may well minimize the total energy. The discrete model did not show this phenomenon, which appears then to be an unphysical consequence of the continuous range of segment lengths.(2)

Although the Zwanzig–Lauritzen (ZL) model of a polymer has interesting physical content, its application to the explanation of polymer crystallization may not be immediately obvious. The great merit of the model is that it allows exact analytical study of very interesting thermodynamic behavior. Thus it can be used as a proving ground for mathematical techniques and thermodynamic hypotheses which may be useful in the study of more realistic but mathematically less tractable systems. This is the motivation of the present work. The aim has been to seek a new analytical solution to the continuous ZL model which would simplify the calculations and extend the range of results obtainable by analytical means. In particular, the probability density for the length of a given segment in the chain molecule and the correlations among the segments are of interest.

In this paper we introduce a symmetrized version of the original ZL model which eliminates the end effects in the molecule. A new solution will be obtained using a grand ensemble, which allows the total length of the molecules to vary, and the powerful transfer matrix method well known from work on the Ising model.(3) The simplifications obtained in a grand ensemble of this sort have been discussed by Lifson and Zimm(4) in the context of the helix–coil transition in biopolymers. But unlike the grand canonical ensemble familiar from the statistical mechanics of fluids, the grand ensemble lacks an extensive parameter corresponding to the volume of the system and the dispersion in the length of the molecules remains large in the thermodynamic limit. This means that the canonical and grand ensembles are not equivalent for large systems by the traditional arguments and as a result little use has been made of the latter ensemble. Here we reexamine the validity of the grand ensemble and find formal justification for the more extensive use we propose to make of it.

The transfer matrix formulation that we employ below is different from well-known previous applications in that the states available to each site in the linear lattice are not discrete but form a continuum, and consequently the eigenvalue equation is a homogeneous Fredholm integral equation of the second kind. Fortunately, this equation can be completely solved and we find the eigenvalues related to the positive zeros of Bessel functions, and the eigenfunctions are obtained in terms of the Bessel functions themselves.

Using these results, it becomes a simple matter to find an expression for the free energy density and reproduce the results of Zwanzig and Lauritzen concerning the phase transition. We point out, however, that the free energy density in the thermodynamic limit does not contain information about a large class of properties, including the probability density for the length of a segment, correlations among segments, and end effects in the molecule. These properties can be studied with our new methods. We choose to present a derivation of the probability density for one segment and for two or more segments separated by given numbers of folds. The concept of local crystal thickness is discussed and the range of critical correlations is found to be infinite. Further calculations have been carried out concerning end effects and other properties but we shall only state the more interesting results in the concluding discussion.

Section 2 presents the model in detail and the basic statistical mechanics used. In Section 3 we introduce the transfer matrix and solve for the eigenvalues and eigenfunctions. The form of the phase transition is obtained in Section 4 and in Section 5 we extend the investigation to segment length probability densities. We conclude with a discussion of the model on the basis of results obtained.

2. MODEL AND ENSEMBLE

2.1. The Model

Consider the case of a single, long polymer molecule confined to a plane. The ZL model is a specification of the configurations available to this two-dimensional molecule and the energy associated with each such configuration. In order to simplify the phenomenon of chain folding, the molecule is assumed capable of creating 180° folds in the forward direction only. The folds are of a uniform character occupying a length q of the molecule. The part of the molecule located between two folds is called a segment of the chain. It can have any length in the interval from zero to infinity (see Fig. 1).

The total energy of any configuration is given by associating an energy u with each fold and an interation energy $-J \min(x_l, x_{l+1})$ with any pair of neighboring segments of length x_l and x_{l+1}. Here $\min(x, y)$ stands for the smaller of the two numbers x and y. The idea behind this choice of energy is that the segments interact with short-ranged van der Waals forces. The

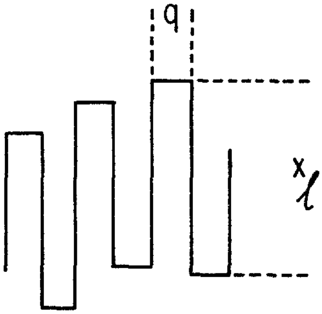

Fig. 1. Configuration of a polymer chain in the ZL model.

constant J will be taken to be positive, so that the forces between segments will be attractive. The fold energy will not be limited to positive values but, as mentioned above, the results for negative u may be of less physical interest.

Thus we can specify the configuration of the molecule by a sequence $x_1, x_2, x_3, \ldots, x_N$ and the corresponding total energy is

$$E(x_1, \ldots, x_N) = (N-1)u - J \sum_{i=1}^{N-1} \min(x_i, x_{i+1}) \tag{1}$$

$$L = (N-1)q + \sum_{i=1}^{N} x_i \tag{2}$$

From Fig. 1 and the total energy (1) it is clear that the end segments and the internal segments are fundamentally different. The internal segments appear in two nearest-neighbor pairs, while the end segments appear in only one such pair. This asymmetry in the model is physically well founded, but for mathematical reasons we shall here introduce a symmetrized version of the model. We simply let the two end segments form a nearest-neighbor pair with a fold in between. Then the total energy and length of the molecule become

$$E_s(x_1, \ldots, x_N) = Nu - J\left[\sum_{i=1}^{N-1} \min(x_i, x_{i+1}) + \min(x_N, x_1)\right] \tag{3}$$

$$L_s = Nu + \sum_{i=1}^{N} x_i \tag{4}$$

The symmetrized model has no simple relation to the type of chain molecule considered here. It does not result merely by bending the formerly considered configurations into a closed ring. Neither will the physical properties of the two models necessarily be the same. There may be circumstances such that the symmetrized model would be physically justifiable, but at this point the interaction between end segments is just an artifice introduced for mathematical purposes. We shall return in the concluding discussion to consider the differences between the two models, but the following presentation will be in terms of the somewhat simpler symmetrized model.

It should be clear that the relevance of the ZL model to the behavior of a real polymer is entirely hypothetical and intuitive. The model disregards the microscopic details of a real polymer. The velocities of various parts of the molecule are neglected and the spatial coordinates giving rise to the various configurations are treated in a coarse way. Thus the physical validity of the model must be determined by comparison of its predictions with the results of experiments.

2.2. The Canonical Ensemble

The canonical ensemble is characterized by the fact that the size of the system as well as any external parameters are fixed while the number of molecules in the available states is proportional to a weighting factor

$$W_c(\xi) = e^{-\beta E(\xi)} \tag{5}$$

Here $E(\xi)$ is the total energy in a state indexed by ξ and β^{-1} is equal to k_BT, the Boltzmann constant multiplied by the temperature. The canonical partition function is defined by

$$Q(S) = \sum_{\xi \in S} e^{-\beta E(\xi)} \tag{6}$$

where S denotes the set of all states satisfying the restrictions on the size and external parameters of the system.

Zwanzig and Lauritzen used the canonical ensemble and started their work from the following form for the partition function of the original model:

$$\begin{aligned} Q(L) = \sum_{N=1}^{\infty} \int_0^{\infty} dx_1 \cdots \int_0^{\infty} dx_N \, \delta\left(\sum_{i=1}^{N} x_i + (N-1)q - L \right) \\ \times \exp\left[\beta J \sum_{i=1}^{N-1} \min(x_i, x_{i+1}) - (N-1)\beta u \right] \end{aligned} \tag{7}$$

We note that the summation over available states ξ in (6) has been replaced by a summation over the number of segments N, integrations over all segment lengths, and a δ-function which picks out only those configurations having a fixed total length L. As it stands, (7) may be confusing since it appears as if $Q(L)$ is not dimensionless, as it should be and clearly is in (6). However, we note that the present model in which the states form a continuous set should be viewed as some limiting form of a discrete model. Thus the integration over segment length arises from a sum over discrete lengths

$$\sum_{x_i} \to \int_0^{\infty} dx_i \, w(x_i) \tag{8}$$

Here $w(x_i)$ is a density of states factor. In (7) it has been implicitly assumed that $w(x_i)$ is equal to unity, but it still carries the dimension inverse length, so that $Q(L)$ is in fact dimensionless.

It is generally assumed that a phase transition, as defined in terms of discontinuities or singularities in thermodynamic variables, will only occur in the limit of an infinitely large system, the thermodynamic limit. For the present model and the canonical ensemble the thermodynamic limit is obtained simply by letting L approach infinity.

2.3. The Grand Ensemble

The restriction to configurations of fixed total length complicates the evaluation of the canonical partition function $Q(L)$. Thus we shall want to use a generalized ensemble in which external parameters are fixed but the size of the system is allowed to vary. The number of molecules in a state ξ is given by the weighting factor

$$W_G(\xi) = e^{-\beta E(\xi) - \alpha M(\xi)} \tag{9}$$

where $M(\xi)$ is directly proportional to the size of the system. Usually $M(\xi)$ is taken to be the number of particles in the system, but in our continuous polymer model the total length L measures the size of the system. Summing over the available states, one obtains the grand partition function

$$Z(\alpha) = \sum_S e^{-\beta E(\xi) - \alpha M(\xi)} \tag{10}$$

Here the set of available states S includes the full range of system sizes, but any external parameter remains fixed.

The grand partition function for the symmetrized ZL model becomes

$$\begin{aligned} Z(\alpha) = \sum_{N=1}^{\infty} [\exp -N(\beta u + \alpha q)] \int_0^\infty dx_1 \cdots \int_0^\infty dx_N \\ \times \exp \left\{ \beta J \left[\sum_{i=1}^{N-1} \min(x_i, x_{i+1}) + \min(x_N, x_1) \right] - \alpha \sum_{i=1}^{N} x_i \right\} \end{aligned} \tag{11}$$

Comparing (11) to $Q(L)$ in (7), we see that the δ-function has been replaced by an exponential, $\exp[-\alpha L(x_1, \ldots, x_N)]$. Thus the two functions are related as follows:

$$Z(\alpha) = \int_0^\infty dL\, Q(L)\, e^{-\alpha L} \tag{12}$$

For a large class of systems the relation (12) or its analog leads to the equivalence of the canonical and grand ensembles for large systems. The reason is that $Q(L)e^{-\alpha L}$ is sharply peaked at some value $L(\alpha)$, so that the dispersion in system size is small and one gets

$$Z(\alpha) \cong Q(L(\alpha)) e^{-\alpha L(\alpha)} \tag{13}$$

This approximation will usually become increasingly accurate as $L(\alpha)$ becomes larger.

In the present case we shall not assume the equivalence of the canonical and grand ensemble. It should be noted that our model differs from a more typical system in that there is no spatial limitation to take the place of volume. Thus the total length L is the only extensive parameter in the canonical ensemble, and the grand ensemble is entirely without extensive variables. Such ensembles lacking extensive parameters are often called generalized ensembles.(5) A grand ensemble similar to the one we shall use has recently been discussed by Lifson and Zimm(4) in the context of the helix–coil transition in biopolymers. The dispersion in system size is known to remain significant even for large systems, and the grand and canonical ensembles are not equivalent in the usual sense. In view of this we shall not rely in what follows upon any assumption of *a priori* validity of the grand ensemble. Instead we shall base our discussion on the commonly accepted canonical ensemble and carefully prove that a large class of properties will in fact be the same in the two ensembles in the thermodynamic limit.

The question arises as to how the thermodynamic limit should be defined in the grand ensemble in the absence of any extensive parameters. Noting that α is the parameter conjugate to the total length,

$$\langle L\rangle_\alpha = -Z^{-1}(\alpha)(\partial/\partial\alpha)Z(\alpha) \tag{14}$$

we suggest the definition

$$\lim_{\alpha\to\alpha_0+} \quad \text{where} \quad \lim_{\alpha\to\alpha_0+} \langle L\rangle_\alpha = \infty \quad \text{and} \quad \langle L\rangle_\alpha < \infty \qquad \text{for} \quad \alpha > \alpha_0 \tag{15}$$

It can be seen directly from (12) that $Z(\alpha)$ is a monotone decreasing function of α, as is $\langle L\rangle_\alpha$. If $Z(\alpha)$ is not discontinuous, the thermodynamic limit can then be equivalently defined by

$$\lim_{\alpha\to\alpha_0+}, \qquad \lim_{\alpha\to\alpha_0+} Z(\alpha) = \infty, \quad Z(\alpha) < \infty \qquad \text{for} \quad \alpha > \alpha_0 \tag{16}$$

Since we have

$$\int_0^{L_0} dL\, Q(L)\, e^{-\alpha L} < \infty, \qquad L_0 < \infty, \quad \text{all } \alpha \tag{17}$$

α_0 is the largest α such that the infinitely long molecules play a dominant role in the ensemble.

3. THE TRANSFER MATRIX METHOD

3.1. The Method

Statistical mechanics provides a simple procedure for the theoretical investigation of a system in equilibrium. First one choose the proper statistical ensemble to describe the conditions under which the system is studied. Then the corresponding partition function is written down as a sum of weighting factors for all states of the system. The properties sought can usually be expressed in terms of the partition function and its derivative with respect to some parameters of the problem. Thus the third step is to find a form for the partition function simple enough so that these expressions for the properties can be evaluated. This is the difficult part of the procedure, and in the presence of nontrivial interactions among the constituents of the system exact analytical methods will only rarely suffice to produce the desired information.

Leaving aside for the moment the question of equivalence, we shall here replace the traditional canonical ensemble by the grand ensemble introduced above, in the hopes of finding the mathematical procedure thereby simplified. Thus we turn now to consider the grand partition function for our polymer model. The evaluation of Z as defined in (11) appears to be easier by virtue of the lack of the δ-function representing the length restriction in the canonical partition function, but it still remains a nontrivial task. We shall therefore attempt to generalize the transfer matrix method, well known from much work on the Ising model,⁽³⁾ so that it may apply to the present problem also.

Let the polymer molecule form a linear lattice with a segment at each site. The possible lengths of the segment then define the set of states available to each site. Since we are dealing with a continuum of states, we shall use Dirac's notation, familiar from quantum mechanics. Thus we let $| x\rangle$ denote the state corresponding to segment length x, and we assume the usual orthonormality to hold,

$$\langle x \mid y\rangle = \delta(x - y) \tag{18}$$

It is then not difficult to verify that the partition function Z can be written as

$$\begin{aligned} Z &= \sum_{N=1}^{\infty} e^{-N(\beta u+\alpha q)} \int_0^{\infty} dx_1 \cdots \int_0^{\infty} dx_N \, \langle x_1 \mid M \mid x_2\rangle \\ &\quad \times \langle x_2 \mid M \mid x_3\rangle \cdots \langle x_N \mid M \mid x_1\rangle \\ &= \sum_{N=1}^{\infty} e^{-N(\beta u+\alpha q)} \int_0^{\infty} dx_1 \, \langle x_1 \mid M^N \mid x_1\rangle \\ &= \sum_{N=1}^{\infty} e^{-N(\beta u+\alpha q)} \operatorname{Tr}(M^N) \end{aligned} \tag{19}$$

Here M is an operator defined by the matrix element

$$\langle x \mid M \mid y\rangle = \exp[\beta J \min(x, y) - \tfrac{1}{2}\alpha(x + y)] \tag{20}$$

It operates on the space of states $| \psi\rangle$ spanned by the set $\{| x\rangle\}$, $0 < x < \infty$,

$$| \psi\rangle = \int_0^\infty dx\, \psi(x) \,| x\rangle \tag{21}$$

$$M \,| \psi\rangle = \int_0^\infty dx\, \psi(x) M \,| x\rangle = \int_0^\infty dx\, \phi(x) \,| x\rangle \tag{22}$$

Taking the component of (22) in the $|y\rangle$ direction we find that

$$\phi(y) = \int_0^\infty dx\, \langle y \,|\, M \,|\, x\rangle\, \psi(x) \tag{23}$$

Thus M also defines a nonlocal operator on the space of real functions on the interval zero to infinity. In fact, it follows from (20) and the relation

$$\begin{aligned}\iint_0^\infty dx\, dy\, \langle y \,|\, M \,|\, x\rangle^2 &= \iint_0^\infty dx\, dy\, e^{2\beta J \min(x,y) - \alpha(x+y)} \\ &= 1/\alpha(\alpha - \beta J), \qquad \alpha > \beta J \end{aligned} \tag{24}$$

that for $\alpha > \beta J$, $\langle y \,|\, M \,|\, x\rangle$ is an L_2 kernel and M as defined by (23) is a symmetric operator on the space of real L_2 functions $L_2[0, \infty]$.

Symmetric L_2 operators of this kind have been investigated in connection with the theory of integral equations and we draw upon the text by Tricomi(6) for the following results.

1. M possesses real eigenvalues $\{\lambda_i\}$ and corresponding orthonormal eigenfunctions $\{\psi_i\}$ which may be obtained from the eigenvalue equation

$$\lambda \psi(y) = \int_0^\infty dx\, \langle y \,|\, M \,|\, x\rangle\, \psi(x) \tag{25}$$

2. The trace of M^N, $N = 2, 3, \ldots$, is related to the eigenvalues in the usual way,

$$\mathrm{Tr}(M^N) = \int_0^\infty dx\, \langle x \,|\, M^N \,|\, x\rangle = \sum_i \lambda_i^N \tag{26}$$

where the sum extends over all nonzero eigenvalues with their full degeneracy.

3. We have

$$\iint_0^\infty dx\, dy\, \langle y \,|\, M^N \,|\, x\rangle\, \phi(x)\, \phi(y) = \sum c_i^2 \lambda_i^N \tag{27}$$

where

$$c_i = \int_0^\infty dx\, \phi(x)\, \psi_i(x) \tag{28}$$

These results can be obtained with the use of the Hilbert–Schmidt theorem. Furthermore, we shall, after finding the eigenvalues, explicitly verify that (26) holds for $N = 1$ also.

It follows directly from (26) that

$$Z = \sum_{N=1}^{\infty} e^{-N(\beta u+\alpha q)} \sum_i \lambda_i^N = \sum_{N=1}^{\infty} \sum_i (\lambda_i')^N \tag{29}$$

where

$$\lambda_i' = \lambda_i e^{-(\beta u+\alpha q)} \tag{30}$$

For finite systems Z is finite and all the λ_i' must be smaller than unity in magnitude. We can then interchange the summations in (30), perform the summation over N, and obtain

$$Z = \sum_i \lambda_i'/(1 - \lambda_i') \tag{31}$$

Thus we see that the grand partition function of the symmetrized ZL model is directly related to the eigenvalues of the operator M describing the interaction between neighboring segments in the polymer chain. We shall want to base our calculation of the properties of the ZL model on the spectral form (31) for the grand partition function. However, we note that this relation has been found for α larger than βJ only. This limitation on the use of the transfer matrix method turns out to be of little consequence. The reason is that Z diverges for $\alpha < \beta J$ and all finite systems should therefore fall into the interval $\alpha \geqslant \beta J$.

Clearly the usefulness of the results obtained above rests on our ability to obtain the necessary spectral properties of M. This becomes our next task.

3.2. The Transfer Matrix

Let us now proceed to calculate the eigenvalues and eigenfunctions of M by solving the integral equation

$$\lambda\psi(y) = \int_0^{\infty} dx \, \{\exp[\beta J \min(x, y) - \tfrac{1}{2}\alpha(x + y)]\} \, \psi(x) \tag{32}$$

Using the relation

$$\min(x, y) = \tfrac{1}{2}(x + y) - \tfrac{1}{2} \mid x - y \mid \tag{33}$$

and the definitions

$$\eta = \alpha - \beta J, \qquad \gamma = \tfrac{1}{2}\beta J \tag{34}$$

we multiply (32) by $\exp(\frac{1}{2}\eta y)$ and obtain

$$\lambda f(y) = \int_0^\infty dx\, e^{-\gamma|x-y|} e^{-\eta x} f(x) \tag{35}$$

By differentiating this integral equation twice with respect to y using

$$(\partial/\partial y)|\,x - y\,| = \operatorname{sign}(y - x) \tag{36}$$

$$(\partial/\partial y)\operatorname{sign}(y - x) = 2\delta(y - x) \tag{37}$$

we turn it into a differential equation

$$(\partial^2/\partial y^2) f(y) = [\gamma^2 - (2\gamma/\lambda)e^{-\eta y}] f(y) \tag{38}$$

In order to get rid of the exponential function in (38), we try a coordinate transformation to the new variable

$$z = ae^{-\eta y/2} \tag{39}$$

Defining the new function $g(z)$ by

$$g(z) = f(y) \tag{40}$$

and differentiating both sides twice with respect to z using (38) and (39), we obtain

$$\frac{\partial^2}{\partial z^2} g(z) = \left(\gamma^2 - \frac{2\gamma}{\lambda a^2} z^2\right) \frac{4}{\eta^2} \frac{1}{z^2} g(z) - \frac{1}{z} \frac{\partial}{\partial z} g(z) \tag{41}$$

At this point we shall limit ourselves to the positive eigenvalues, $\lambda > 0$, and define the constant a by

$$a = (8\gamma/\lambda\eta^2)^{1/2} \tag{42}$$

Equation (41) can now be rewritten in the form of the Bessel equation

$$z^2 \frac{\partial^2}{\partial z^2} g(z) + z \frac{\partial}{\partial z} g(z) + (z^2 - \nu^2)\, g(z) = 0 \tag{43}$$

Here the constant ν is defined by

$$\nu = 2\gamma/\eta = \beta J/(\alpha - \beta J) \tag{44}$$

Recalling our restriction to α larger than βJ, we see that ν is positive and generally of nonintegral order. The set of solutions is then immediately known and is summarized by

$$g(z) = C_1 \mathscr{J}_\nu(z) + C_2 \mathscr{J}_{-\nu}(z) \tag{45}$$

where $\mathscr{J}_\nu$ and $\mathscr{J}_{-\nu}$ are Bessel functions tabulated in standard references. The constants C_1 and C_2 are allowed to vary over the real axis to span the full set of real-valued solutions to Eq. (42).

At this point we must note that in turning an integral equation into a differential equation, one generally expands the set of solutions. Thus we shall now insert the solutions

$$f(y) = C_1\mathscr{J}_\nu(ae^{-\eta y/2}) + C_2\mathscr{J}_{-\nu}(ae^{-\eta y/2}) \tag{46}$$

obtained from (45) into the original integral Eq. (35) in order to pick out the eigenfunctions we are looking for. We get

$$\begin{aligned}\lambda(C_1\mathscr{J}_\nu(ae^{-\eta y/2}) &+ C_2\mathscr{J}_{-\nu}(ae^{-\eta y/2}))\\ &= \int_0^\infty dx\, e^{-\gamma|x-y|}e^{-\eta x}[C_1\mathscr{J}_\nu(ae^{-\eta x/2}) + C_2\mathscr{J}_{-\nu}(ae^{-\eta x/2})]\end{aligned} \tag{47}$$

In terms of the variable z this identity can be written

$$\begin{aligned}\tfrac{1}{2}\lambda a^2\eta[C_1\mathscr{J}_\nu(z) + C_2\mathscr{J}_{-\nu}(z)] &= z^\nu\int_z^a dy\, y^{-\nu+1}[C_1\mathscr{J}_\nu(y) + C_2\mathscr{J}_{-\nu}(y)]\\ &\quad + z^{-\nu}\int_0^z dy\, y^{\nu+1}[C_1\mathscr{J}_\nu(y) + C_2\mathscr{J}_{-\nu}(y)]\end{aligned} \tag{48}$$

Here we have split up the integration range to avoid the absolute value operation.

In performing the integrations of (48), we make use of the following identities found in standard references [Ref. 7, formulas (11.3.20) and (11.3.21)]:

$$\int_0^z dt\, t^\nu\mathscr{J}_{\nu-1}(t) = z^\nu\mathscr{J}_\nu(z), \qquad \operatorname{Re}\nu > 0, \tag{49}$$

$$\int_0^z dt\, t^{-\nu}\mathscr{J}_{\nu+1}(t) = [1/2^\nu\Gamma(\nu+1)] - z^{-\nu}\mathscr{J}_\nu(z), \qquad \operatorname{Re}\nu > 0 \tag{50}$$

The integrated identity becomes

$$\begin{aligned}&\tfrac{1}{2}\lambda a^2\eta[C_1\mathscr{J}_\nu(z) + C_2\mathscr{J}_{-\nu}(z)]\\ &\quad = C_1[z\mathscr{J}_{\nu-1}(z) - z^\nu a^{-\nu+1}\mathscr{J}_{\nu-1}(a) + z\mathscr{J}_{\nu+1}(z)]\\ &\qquad + C_2\left[z^\nu a^{-\nu+1}\mathscr{J}_{-\nu+1}(a) - z\mathscr{J}_{-\nu+1}(z) + \frac{z^{-\nu}}{2^{-\nu-1}\Gamma(-\nu)} - z\mathscr{J}_{-\nu-1}(z)\right]\end{aligned} \tag{51}$$

It is clear that for $\nu \neq 0$ the identity above will require C_2 to vanish due to the presence of the simple power $z^{-\nu}$ on the right. Thus (51) becomes

$$\tfrac{1}{2}\lambda a^2\eta \mathscr{J}_\nu(z) = z\mathscr{J}_{\nu-1}(z) - z^\nu a^{-\nu+1}\mathscr{J}_{\nu-1}(a) + z\mathscr{J}_{\nu+1}(z) \tag{52}$$

The coefficient multiplying z^ν on the right must also vanish. Since we require a to be nonzero, we then get

$$\mathscr{J}_{\nu-1}(a) = 0 \tag{53}$$

Thus a must be one of the infinitely many positive zeros of the Bessel function $\mathscr{J}_{\nu-1}$. The remaining identity is

$$\tfrac{1}{2}\lambda a^2\eta \mathscr{J}_\nu(z) = z\mathscr{J}_{\nu-1}(z) + z\mathscr{J}_{\nu+1}(z) \tag{54}$$

Recalling the definition of a, we can rewrite (54) in the form

$$\mathscr{J}_{\nu-1}(z) + \mathscr{J}_{\nu+1}(z) = (2\nu/z)\mathscr{J}_\nu(z) \tag{55}$$

which is a well-known identity satisfied by the Bessel functions.

We conclude that the eigenfunctions of M corresponding to positive-definite eigenvalues can be summarized as

$$\psi_i(y) = C_i e^{-\eta y/2}\mathscr{J}_\nu(a_i e^{-\eta y/2}) \tag{56}$$

where a_i is one of the positive, definite-zeros of $\mathscr{J}_{\nu-1}(y)$. The condition on a_i determines the corresponding eigenvalues and we get

$$\lambda_i = \frac{8\gamma}{\eta^2 a_i^2} = \frac{4\beta J}{(\alpha-\beta J)^2}\frac{1}{a_i^2} = \frac{4}{\beta J}\frac{\nu^2}{a_i^2} \tag{57}$$

Since the eigenvalues are all different, we know that the corresponding eigenfunctions are orthogonal, but we still have to determine the normalization constants C_i by the condition

$$\int_0^\infty dx\, [\psi_i(x)]^2 = 1 \tag{58}$$

Thus we have

$$C_i^{-2} = \int_0^\infty dx\, e^{-\eta x}[\mathscr{J}_\nu(a_i e^{-\eta x/2})]^2 \tag{59}$$

Transforming to the variable

$$t = e^{-\eta x/2} \tag{60}$$

we get

$$C_i^{-2} = \int_0^1 dt\, (2/\eta)\, t\mathscr{J}_\nu^2(a_i t) \tag{61}$$

This integration can be carried out directly [use (11.4.5) and (9.1.27) in the *Handbook of Mathematical Functions*[7]] and we find that

$$C_i = \eta^{1/2}[\mathscr{J}_\nu(a_i)]^{-1} \tag{62}$$

Thus we conclude that the normalized eigenfunction is

$$\psi_i(y) = \eta^{1/2}[\mathscr{J}_\nu(a_i)]^{-1}\, e^{-\eta y/2} \mathscr{J}_\nu(a_i e^{-\eta y/2}) \tag{63}$$

At the outset we limited our calculations to the set of positive definite eigenvalues and corresponding eigenfunctions. Clearly only nonzero eigenvalues have any relevance for Z, but we still need to prove that there are no negative eigenvalues. This can be done by explicitly verifying that (26) is satisfied for $N = 1$ and $N = 2$, where the sum is taken over the positive eigenvalues found above.

By direct integration we easily find that

$$\operatorname{Tr} M = \int_0^\infty dx\, e^{(\beta J-\alpha)x} = (\alpha - \beta J)^{-1}, \qquad \alpha > \beta J \tag{64}$$

$$\operatorname{Tr} M^2 = \int_0^\infty dx \int_0^\infty dy\, e^{2\beta J \min(x,y) - \alpha(x+y)} = \alpha^{-1}(\alpha - \beta J)^{-1}, \qquad \alpha > \beta J \tag{65}$$

The corresponding sums of eigenvalues λ_i and λ_i^2 given by (61) are

$$\sum_{i=1}^\infty \lambda_i = \frac{4\nu^2}{\beta J} \sum_{i=1}^\infty \frac{1}{a_i^2} = \frac{4\nu^2}{\beta J} S_{2,\nu-1} \tag{66}$$

$$\sum_{i=1}^\infty \lambda_i^2 = \frac{16\nu^4}{(\beta J)^2} \sum_{i=1}^\infty \frac{1}{a_i^4} = \frac{16\nu^4}{(\beta J)^2} S_{4,\nu-1} \tag{67}$$

Fortunately, the sums $S_{2,\nu-1}$ and $S_{4,\nu-1}$ of inverse powers of positive zeros of Bessel functions can be found in the literature.(8) They are

$$S_{2,\nu-1} = 2^{-2}\nu^{-1} \tag{68}$$

$$S_{4,\nu-1} = 2^{-4}\nu^{-2}(\nu + 1)^{-1} \tag{69}$$

Substituting these values into (66) and (67), we find the expected agreement with (64) and (65). This concludes our evaluation of the grand partition function in terms of the eigenvalues of the operator M.

4. THE PHASE TRANSITION

In their original article Zwanzig and Lauritzen studied the properties of the polymer molecule in the thermodynamic limit based on an evaluation of the quantity $\lim_{L\to\infty} L^{-1} \ln Q(L)$, where $Q(L)$ is the canonical partition function given in (7). We note that the Helmholtz free energy per unit length is

$$f(T, L) = -(\beta L)^{-1} \ln Q(\beta, L) \tag{70}$$

and the energy and entropy per unit length as well as other thermodynamic properties can be obtained in terms of simple derivatives of $f(T, L)$.

Here we shall present a simple rederivation of the original analytical results using the transfer matrix method and a relation between canonical and grand partition functions pointed out by Lifson.(9) It follows easily from (12) that if $Q(L)$ is finite for finite L and

$$\lim_{L\to\infty} L^{-1} \ln Q(L) = y_1 \tag{71}$$

then the grand partition function $Z(\alpha)$ will be finite for $\alpha > y_1$ and infinite for $\alpha < y_1$. Thus we find that

$$\lim_{L\to\infty} L^{-1} \ln Q(L) = \alpha_0 \tag{72}$$

where α_0 is defined by (16).

We proceed to determine α_0 as a function of temperature by examining the divergence of Z. Recall that for finite systems Z is given by

$$Z = \sum_{i=1}^{\infty} \lambda_i'/(1 - \lambda_i'), \qquad \alpha > \beta J \tag{73}$$

From the original form of $Z(\alpha)$ as given by (11) it is easily seen that it is infinite for $\alpha < \beta J$. Thus we can conclude that $\alpha_0 \geqslant \beta J$ and we can limit our investigation to the case when $\alpha > \beta J$. This is precisely the regime where the transfer matrix method works.

Since λ_1' is a function of α and β, we determine α_0 as an implicit function of β by requiring λ_1' to be unity

$$\lambda_1'(\alpha_0, \beta) = e^{-(\beta u + \alpha q)}(4/\beta J)(\nu_0{}^2/a^2_{\nu_0-1,1}) = 1 \tag{74}$$

Here we have made use of (57) and ν_0 is defined by

$$\nu_0 = \beta J/(\alpha_0 - \beta J) \tag{75}$$

The sequence of all positive zeros of $\mathscr{J}_{\nu_0-1}(t)$ is given by $a_{\nu_0-1,1} < a_{\nu_0-1,2} < \cdots$, where the dependence on the order of the Bessel function has been explicitly indicated. If we define σ by

$$\sigma = 2(\beta J)^{-1/2}\, e^{-\beta(u+Jq)/2} \tag{76}$$

and note that

$$e^{-\alpha q} = e^{-\beta J q} e^{-\beta J q/\nu} \tag{77}$$

and take the square root of (74), we see that

$$a_{\nu_0 - 1,1} = \sigma e^{-\beta J q/2\nu_0} \nu_0 \tag{78}$$

This relation determines ν_0 and through (75) also α_0 as an implicit function of β. When $a_{\nu-1,1}$ is larger than $\sigma\nu \exp -(\beta J q/2\nu)$, λ_1' is smaller than unity. If there is a ν_0 satisfying (78), then there is a corresponding α_0 larger than βJ. However, if

$$a_{\nu-1,1} > \sigma\nu e^{-\beta J q/2\nu}, \qquad 0 < \nu < \infty \tag{79}$$

then α_0 must be set equal to βJ.

From the fact that Z is a monotone increasing function of ν for a given temperature, it follows that λ_1' is monotone increasing also. Thus there will be a unique solution ν_0 of (78) or none. Furthermore, if there is a solution, then

$$a_{\nu-1,1} < \sigma\nu e^{-\beta J q/2\nu}, \qquad \nu \to \infty \tag{80}$$

The behavior of $a_{\nu-1,1}$ for large orders is available in the literature.(10) We have

$$a_{\nu-1,1} = \nu - 1 + d_1 \nu^{1/3} + O(\nu^{-1/3}), \qquad d_1 \cong 1.856 \tag{81}$$

Similarly, we note that for large ν

$$\sigma\nu e^{-\beta J q/2\nu} = \sigma(\nu - \tfrac{1}{2}\beta J q) + O(\nu^{-1}) \tag{82}$$

Comparing (81) and (82), we can conclude that for $\sigma < 1$ there is no solution to (78), for $\sigma = 1$ there is an asymptotic solution $\nu_0 = \infty$, and for $\sigma > 1$ there is a unique solution $0 < \nu_0 < \infty$.

When the fold energy u is positive, as we assume here, σ is a monotone decreasing function of β and thus a monotone increasing function of temperature. Thus σ is unity at a temperature T_c determined by

$$2(kT_c/J)^{1/2} e^{-(1/2kT_c)(u+Jq)} = 1 \tag{83}$$

Below T_c, σ is less than unity and above T_c it is larger than unity. It follows that

$$\lim_{L\to\infty} (1/L) \ln Q(L) = \alpha_0 = \begin{cases} \beta J, & T < T_c \\ \beta J(1 + \nu_0^{-1}), & T > T_c \end{cases} \tag{84}$$

where ν_0 approaches infinity as T approaches T_c from above.

We shall now show that T_c in fact marks a second-order phase transition in the polymer molecule. This will follow from a study of the energy density $L^{-1}\langle E\rangle_L$ and the corresponding specific heat $L^{-1}(\partial/\partial T)\langle E\rangle_L$ in the thermodynamic limit. Here $\langle E\rangle_L$ is the average energy of a molecule of length L in the canonical ensemble, and we note that

$$L^{-1}\langle E\rangle_L = -L^{-1}(\partial/\partial\beta)\ln Q(\beta, L) \tag{85}$$

From the results above it then follows that

$$\lim_{L\to\infty} L^{-1}\langle E\rangle_L = -(\partial/\partial\beta)(\alpha_0) = \begin{cases} -J, & T < T_c \\ -J - (\partial/\partial\beta)(\beta J/\nu_0), & T > T_c \end{cases} \tag{86}$$

Both sides of (78) are continuous functions of β and ν, so that there will be no discontinuity in the energy density for $T > T_c$, but we must check what happens at $T = T_c$. In the limit as T approaches T_c from above, $T \to T_c{}^+$, ν_0 approaches infinity, and (78) can be written as

$$\sigma\nu_0 \cong \nu_0 + d_1\nu_0^{1/3} \tag{87}$$

It follows that

$$\nu_0 \cong d_1^{3/2}(\sigma - 1)^{-3/2}, \qquad T \sim T_c + \tag{88}$$

$$\nu_0^{-1} \cong d_1^{-3/2}(\sigma - 1)^{3/2}, \qquad T \sim T_c + \tag{89}$$

and

$$(\partial/\partial\beta)\,\nu_0^{-1} \cong d_1^{-3/2}\tfrac{3}{2}(\sigma - 1)^{1/2}(\partial/\partial\beta)\sigma, \qquad T \sim T_c + \tag{90}$$

From (82) we get

$$(\partial/\partial\beta)\sigma = -[\tfrac{1}{2}\beta^{-1} + \tfrac{1}{2}(u + Jq)]\sigma \tag{91}$$

and it easily follows that $(\partial/\partial\beta)\nu_0^{-1}$ vanishes at $T = T_c$ where $\sigma = 1$. Thus the energy density is continuous in the thermodynamic limit at $T = T_c$.

The specific heat in the thermodynamic limit is

$$C(T) = \lim_{L\to\infty} L^{-1}(\partial/\partial T)\langle E\rangle_L = \begin{cases} -(\partial/\partial T)(\partial/\partial\beta)(\beta J/\nu_0), & T > T_c \\ 0, & T < T_c \end{cases} \tag{92}$$

Both above and below T_c it is continuous, so we can limit our study to $T = T_c$. Dropping terms that obviously vanish as $T \to T_c{}^+$, we have

$$C(T) \cong -\beta J(\partial/\partial T)(\partial/\partial\beta)\,\nu_0^{-1}, \qquad T \sim T_c + \tag{93}$$

Using (90) and (91) and noting that $(\partial/\partial T)\,(\partial/\partial\beta)\sigma$ is finite at $T = T_c$, we get

$$C(T) = (J/kT)\, d_1^{-3/2}\tfrac{3}{4}(kT + u + Jq)(\partial/\partial T)(\sigma - 1)^{1/2}, \qquad T \sim T_c+ \tag{94}$$

From the relation

$$\frac{\partial}{\partial T}\,\sigma(T)\bigg|_{T=T_c} = \frac{1}{2T_c} + \frac{1}{2kT_c^{\,2}}\,(u + Jq) \tag{95}$$

it follows that

$$\frac{\partial}{\partial T}\,[\sigma(T) - 1]^{1/2} = (8T_c)^{-1/2}\left(1 + \frac{u + Jq}{kT_c}\right)^{1/2}(T - T_c)^{-1/2}, \qquad T \sim T_c+ \tag{96}$$

and we finally obtain

$$C(T) = \frac{3}{8}\, d_1^{-3/2} J \left(1 + \frac{u + Jq}{kT_c}\right)^{3/2}(2T_c)^{-1/2}(T - T_c)^{-1/2}, \qquad T \sim T_c+ \tag{97}$$

We have found that $C(T)$ has an inverse square root singularity at $T = T_c$. Thus the polymer molecule goes through a second-order phase transition at T_c. This result was obtained by Zwanzig and Lauritzen in their original work for the special case of vanishing fold length. Setting q equal to zero in (97) and transforming to the reduced variables of their paper, our results coincide with theirs, as they should.

We can, of course, go on and generalize all the analytical results of Zwanzig and Lauritzen to the case of nonvanishing fold length. In particular, it is easily seen that for negative fold energy in the interval $-(J/4e) - qJ < u < -qJ$, σ will be infinite for both $T = 0$ and $T = \infty$, having a single minimum below unity for some finite temperature. It follows that σ will intersect unity at two temperatures and there will be two phase transitions. The folded mode will be stable for both high and low temperatures, while the extended mode associated with $\sigma < 1$ will be stable over some intermediate range of temperatures. When u is less than $-(J/4e) - qJ$ the minimum value of σ will be above unity and the folded mode will be stable for all temperatures.

Despite their obvious power, both the solution presented above and that of Zwanzig and Lauritzen are rather limited in scope. In calculating $\lim_{L\to\infty} L^{-1} \ln Q(L)$, we suppress much of the more detailed behavior of the partition function in the thermodynamic limit. Thus the above solution will not allow us to discuss the probability density for the length of a segment or the correlations among the segments. For example, we note that the probability density for the length of a segment indexed by 1 is

$$P^{(1)}(x_1) = Q(L, x_1)/Q(L) \tag{98}$$

Here $Q(L, x_1)$ is the partition function of a molecule with the length of segment 1 fixed at x_1. It is not hard to verify that

$$\lim_{L\to\infty} L^{-1} \ln Q(L, x_1) = \lim_{L\to\infty} L^{-1} \ln Q(L) = y_1 \tag{99}$$

so that

$$Q(L) = Q_0(L) y_1^L, \qquad Q(L, x_1) = Q_0(L, x_1) y_1^L \tag{100}$$

It follows that $P^{(1)}(x_1)$ depends precisely on those factors $Q_0(L)$ and $Q_0(L, x_1)$ which are not determined by the above solution in the thermodynamic limit.

5. CALCULATIONS IN THE GRAND ENSEMBLE

5.1. Validity of the Grand Ensemble

Despite its apparent mathematical advantages, the grand ensemble has only rarely been used to represent the properties of one-dimensional systems such as the polymer model discussed here. The reason for this neglect may perhaps be found in the facts that (a) the model lacks an external parameter such as volume describing the physical size of the system, and (b) the fluctuations in the size of the system are large in the grand ensemble. The lack of any extensive parameters in the grand ensemble suggests a close relation to the generalized ensembles introduced by Guggenheim^(5,11) and known to require some care in their use. Furthermore, there is the question of how to define the thermodynamic limit. The latter fact (b) indicates that the canonical and grand ensembles are not equivalent in all respects. The canonical ensemble being commonly accepted, one then would need to explicitly justify the use of the grand ensemble.

In an article on the helix–coil transition in DNA Lifson and Zimm^(4) made use of the grand ensemble to calculate the fraction of bonded pairs in the polymer. While noting the mathematical simplicity achieved in this ensemble, they took a negative view of its validity. Lifson^(9) subsequently used the grand ensemble purely as a mathematical artifice in his method of calculating the free energy density in the thermodynamic limit discussed in the preceding section.

However, we have already pointed out that the thermodynamic limit of the free energy density does not contain the desired information about the probability density for the length of a given segment, correlations among the segments, and end effects in the polymer chain. In order to obtain such information, we shall have to make more extensive use of the grand ensemble. We turn now to examine the validity of the grand ensemble and to provide justification for the use we shall make of it.

Without entering into a discussion of the philosophy underlying the currently popular choices of statistical ensembles, let us note that the canonical ensemble can be obtained on the basis of a minimum information principle.(12) That is, if we specify that all of the polymer molecules in our ensemble should have a fixed length L and that the average energy should take on the value $\langle E\rangle$, then the canonical ensemble minimizes the total information. Similarly, if we let energy and length vary but demand that the averages $\langle E\rangle$ and $\langle L\rangle$ take on given values, then the grand ensemble as defined in Section 23 minimizes the information. Although a theoretician is free to study either one of these two ensembles, it should be noted that there is no reason to believe that his results would be independent of his choice of ensemble. We take the point of view that the canonical ensemble is the more relevant here for obvious reasons. It is predominant in studies of a similar nature and it was used in the original studies of the ZL model. In fact the ZL model appears to be restricted to describe properties of polymer chains of fixed length since it does not consider any mechanism by which a molecule could alter its total length.

Thus we shall in the following seek to obtain the properties of the ZL model in the canonical ensemble, allowing the use of the grand ensemble only if we can explicitly show that the results in the two ensembles must be the same. In view of the differences between the ensembles already noted, it may seem unlikely that we can make any further use of the grand ensemble. However, we shall limit our attention here to properties of the system in the thermodynamic limit. We have already noted that the commonly accepted definition of a phase transition is believed to make it a property of infinite systems only. Any physical system will, of course, be finite, but the idea is that we are really only interested in properties which become independent of system size, or dependent on size in some predictable way, in the thermodynamic limit. One then assumes that the physical system studied is large enough that its behavior cannot be distinguished from that of an infinite system. If this is the case, then we would expect the difference between the canonical and grand ensembles to be small even if the dispersion in system size in the latter ensemble is large. In fact, we shall now go on to state and prove a theorem of central importance to this work which shows that for most practical purposes the difference between the ensembles will vanish in the thermodynamic limit.

Theorem. Let f be a property of the polymer molecule such that its canonical average is bounded and convergent in the thermodynamic limit for a given temperature,

$$|\langle f\rangle_{C,L}| < f_M < \infty, \qquad 0 < L < \infty \tag{101}$$

$$\lim_{L\to\infty} \langle f\rangle_{C,L} = \langle f\rangle_{C,\infty} \tag{102}$$

and let the thermodynamic limit in the grand ensemble exist as defined in (16); then the thermodynamic limit of the grand average must exist and be identical with the corresponding canonical average,

$$\lim_{\alpha\to\alpha_0+} \langle f\rangle_{G,\alpha} = \langle f\rangle_{G,\alpha_0} = \langle f\rangle_{C,\infty} \tag{103}$$

Proof. Note first the following relation between the grand and the canonical averages:

$$\begin{aligned}\langle f\rangle_{G,\alpha} &= \int_0^\infty dL\, Q(L)\, e^{-\alpha L}\langle f\rangle_{C,L}\, Z^{-1}(\alpha) \\ &= \int_0^\infty dL\, \mu(L;\alpha)\langle f\rangle_{C,L}\end{aligned} \tag{104}$$

Here $\mu(L;\alpha)$ is a probability density for the total length L of the molecule. It follows from (16) that $\mu(L;\alpha)$ satisfies

$$\lim_{\alpha\to\alpha_0+} \int_0^{L_0} dL\, \mu(L;\alpha) = \begin{cases}0, & L_0 < \infty \\ 1, & L_0 = \infty\end{cases} \tag{105}$$

But then we have that for each $\epsilon > 0$ there exists an $L(\epsilon)$ such that

$$|\langle f\rangle_{C,L} - \langle f\rangle_{C,\infty}| < \epsilon/2, \qquad L \geqslant L(\epsilon) \tag{106}$$

and an $\alpha'(\epsilon, L(\epsilon))$ such that

$$\int_0^{L(\epsilon)} dL\, \mu(L;\alpha) < \epsilon/4f_M, \qquad \alpha_0 < \alpha < \alpha'(\epsilon, L(\epsilon)) \tag{107}$$

and it follows immediately that

$$|\langle f\rangle_{G,\alpha} - \langle f\rangle_{C,\infty}| < \epsilon, \qquad \alpha_0 < \alpha < \alpha'(\epsilon, L(\epsilon)) \tag{108}$$

Since ϵ can be chosen arbitrarily small, we conclude that

$$\lim_{\alpha\to\alpha_0+} \langle f\rangle_{G,\alpha} = \langle f\rangle_{G,\alpha_0} = \langle f\rangle_{C,\infty} \tag{109}$$

In the present model the size of the system is measured by the continuous parameter L. It should be clear, however, that the above theorem will hold also when the size of the system is measured by a discrete variable, and therefore it applies to a wide range of models. The theorem is quite strong and inclusive, the only limitations being that the thermodynamic limit in the grand ensemble must exist as defined in (16) and the canonical average must be bounded and have a thermodynamic limit. It appears that these conditions will be satisfied in all but rather exceptional circumstances.

The above theorem establishes a near equivalence between the canonical and grand ensembles in the thermodynamic limit which we shall rely upon to justify the following investigation of the properties of the ZL model.

5.2. Probability Density for Segment Length and Correlations Among Segments

In order to obtain deeper insight into the properties of the symmetrized ZL model of polymer chain folding, we now consider the probability density for the length of a particular segment and the joint probability density for the lengths of two or more segments separated by given numbers of folds. The calculations will be performed in the grand ensemble using the spectral resolution for the grand partition function obtained in Section 3 by the transfer matrix method. One easily verifies that the thermodynamic limit exists as defined in (16) and the probability densities correspond to canonical averages bounded by unity. Thus we can conclude from our equivalence theorem that if the thermodynamic limit of these probability densities exists in the canonical ensemble, then it must coincide with the corresponding result that we shall obtain here in the grand ensemble. Further results concerning the behavior of the model will be reported without proof in the concluding section. Here we shall merely note that below the transition temperature when the polymer molecule is in its extended mode the average segment length is infinite in the thermodynamic limit and we can expect the probability density for segment length to vanish at finite lengths. Thus our investigation will be limited to the folded mode of the polymer above T_c.

Let us first recall that all the molecules in our ensemble form closed rings in a mathematical sense, with the number of segments varying from one to infinity. If we consider the subensemble of rings of a given number of segments, then there is complete symmetry among the segments. Enumerating the segments in each ring by 1, 2,..., N, we then define $P^{(1)}(x)$ to be the probability density for the length of segment 1 in the whole ensemble. It is not hard to see that $P^{(1)}(x)$ will be given by

$$P^{(1)}(x) = Z(x)/Z = Z^{-1} \sum_{N=1}^{\infty} e^{-(\beta u+\alpha q)N} \langle x \mid M^N \mid x \rangle \tag{110}$$

Note now that any integral power of the transfer matrix may be written

$$M^N = \sum_{i=1}^{\infty} \mid \psi_i \rangle \langle \psi_i \mid \lambda_i^N, \qquad N = 1, 2, \ldots \tag{111}$$

Substituting this form into (110), we get

$$P^{(1)}(x) = Z^{-1} \sum_{i=1}^{\infty} [\psi_i(x)]^2 \, \lambda_i'/(1 - \lambda_i') \tag{112}$$

In the thermodynamic limit and above T_c (112) simplifies to

$$P_0^{(1)}(x) = \psi_1^2(x) = (\alpha_0 - \beta J)\,\mathscr{J}_\nu^{-2}(a_1)\,e^{-(\alpha_0-\beta J)x} \times \mathscr{J}_\nu^2(a_1 e^{-(\alpha_0-\beta J)x/2}), \qquad T > T_c \tag{113}$$

Below T_c, $P_0^{(1)}(x)$ vanishes for finite x, as expected.

Integrating $P^{(1)}(x)$, we find

$$\int_0^\infty dx\, P^{(1)}(x) = Z^{-1} \sum_{i=1}^\infty \int_0^\infty dx\, \psi_i^2(x)\, \lambda_i'/(1 - \lambda_i') = 1 \tag{114}$$

Thus the proper normalization of $P^{(1)}(x)$ is a consequence of the normalization of the eigenfunctions $\{\psi_i(x)\}_1^\infty$.

In a similar manner we can define a two-segment joint probability density $P^{(2)}(x_1, n_1, x_2)$ describing the likelihood of finding segment 1 having length x_1 and a segment $n_1 + 1$ in the same molecule having length x_2. We readily obtain the following expression for $P^{(2)}$:

$$\begin{aligned} P^{(2)}(x_1, n_1, x_2) &= Z^{-1} \sum_{N=n_1+1}^\infty e^{-N(\beta u+\alpha q)} \langle x_1 \mid M^{n_1} \mid x_2\rangle\langle x_2 \mid M^{N-n_1} \mid x_1\rangle \\ &= Z^{-1}\left[\sum_{i=1}^\infty \psi_i(x_1)\,\psi_i(x_2)(\lambda_i')^{n_1}\right] \\ &\quad \times \sum_{i=1}^\infty \psi_i(x_1)\,\psi_i(x_2)\lambda_i'/(1 - \lambda_i') \end{aligned} \tag{115}$$

In the thermodynamic limit and above T_c we get

$$P_0^{(2)}(x_1, n_1, x_2) = \psi_1(x_1)\,\psi_1(x_2) \sum_{i=1}^\infty \psi_i(x_1)\,\psi_i(x_2)(\lambda_i')^{n_1} \qquad T > T_c \tag{116}$$

It should now be clear how to generate these probability densities for three and more segments. We obtain

$$\begin{aligned} P^{(3)}(x_1, n_1, x_2, n_2, x_3) &= Z^{-1}\left[\sum_{i=1}^\infty (\lambda_i')^{n_1}\,\psi_i(x_1)\,\psi_i(x_2)\right] \\ &\quad \times \left[\sum_{j=1}^\infty (\lambda_j')^{n_2}\,\psi_j(x_2)\,\psi_j(x_3)\right] \\ &\quad \times \sum_{k=1}^\infty \psi_k(x_3)\,\psi_k(x_1)\lambda_k'/(1 - \lambda_k') \end{aligned} \tag{117}$$

and so on for higher orders. It is again easy to see that these densities will be properly normalized in the thermodynamic limit above T_c where all molecules have the needed number of segments. Note that if we integrate (117) over x_1, x_2, and x_3, we get the probability of finding a molecule with at least $n_1 + n_2 + 1$ segments.

We now turn to illustrate the usefulness of the above probability densities in two simple applications. Let us first recall that the interest in the ZL model was originally based on its representation of the crystallization of long polymer molecules into a folded mode of essentially uniform segment lengths or crystal thickness. The quantity $\rho = L/N$ carries information about the average length of a segment, but it does not answer the question whether there is a well-defined local crystal thickness. This question is better answered by $P_0^{(1)}(x)$. Let us therefore briefly examine the qualitative behavior of this function in the thermodynamic limit and above T_c. Note first that $P_0^{(1)}(x)$ can be written

$$P_0^{(1)}(x) = (\alpha_0 - \beta J)\, a_{\nu-1,1}^{-2} \mathscr{J}_\nu^{-2}(a_{\nu-1,1})\, y^2 \mathscr{J}_\nu^2(y), \qquad T > T_c \qquad (118)$$

where y is equal to $a_{\nu-1,1} \exp[-\frac{1}{2}(\alpha_0 - \beta J)x]$. When x varies between zero and infinity, y varies between $a_{\nu-1,1}$ and zero. Note also that the smallest zero of $\mathscr{J}_\nu$, $a_{\nu,1}$, is larger than $a_{\nu-1,1}$ and for $\nu > 0$, $\mathscr{J}_\nu(0)$ vanishes and $\mathscr{J}_\nu(t)$ has a single maximum between zero and $a_{\nu,1}$

If the folded polymer molecule is to have a well-defined local thickness, it appears reasonable to demand that at least $P_0^{(1)}(x)$ have a maximum, if not a very sharp one. To see if it does, we need to study the zeros of $(\partial/\partial x)$ $P_0^{(1)}(x)$ or equivalently of $(\partial/\partial x)$ $[y^2 \mathscr{J}_\nu^2(y)]$ in the interval $0 < y < a_{\nu-1,1}$. Thus we seek the solutions of

$$\begin{aligned}(\partial/\partial x)[y^2 \mathscr{J}_\nu^2(y)] = {} & [2y\mathscr{J}^2(y) + 2y^2 \mathscr{J}_\nu(y) \mathscr{J}_\nu'(y)] \\ & \times [-\tfrac{1}{2}(\alpha_0 - \beta J)y] = 0 \end{aligned} \qquad (119)$$

or equivalently

$$y^2 \mathscr{J}_\nu(y)[\mathscr{J}_\nu(y) + \mathscr{J}_\nu'(y)] = 0 \qquad (120)$$

The factor $y^2 \mathscr{J}_\nu(y)$ is larger than zero within the interesting range $0 < y < a_{\nu-1,1}$, so we consider the equation

$$\mathscr{J}_\nu(y) + y\mathscr{J}_\nu'(y) = 0 \qquad (121)$$

Using the recurrence relation

$$\mathscr{J}_\nu'(y) = \mathscr{J}_{\nu-1}(y) - (\nu/y)\, \mathscr{J}_\nu(y) \qquad (122)$$

(121) can be written

$$y\mathscr{J}_{\nu-1}(y) - (\nu - 1)\, \mathscr{J}_\nu(y) = 0 \qquad (123)$$

For $\nu < 1$ it is clear that the left-hand side of (123) is always larger than zero in the interesting range of y. For $\nu > 1$, however, there is a solution. This is clear since, for $y = a_{\nu-1,1}$, $y\mathscr{J}_{\nu-1}(y)$ vanishes and the left-hand side is negative, while for small y we have[7]

$$\mathscr{J}_\nu(y) \cong (\tfrac{1}{2}y)^\nu[\Gamma(\nu+1)]^{-1}, \qquad y \to 0 \tag{124}$$

$$y\mathscr{J}_{\nu-1}(y) - (\nu-1)\,\mathscr{J}_\nu(y) \cong y^\nu 2^{-\nu+1}[\Gamma(\nu)]^{-1}\left(1 - \frac{1}{2}\frac{\nu-1}{\nu}\right) > 0, \qquad y \to 0 \tag{125}$$

It follows that for $\nu_0 > 1$, $(\partial/\partial x)\, P_0^{(1)}(x)$ starts out positive for small x, changes sign at some finite value, and is negative for large x. Thus $P_0^{(1)}(x)$ must have a maximum for some length $x_{\max}$, $0 < x_{\max} < \infty$. For $\nu_0 < 1$, on the other hand, $P_0^{(1)}(x)$ takes on its maximum value $\alpha_0 - \beta J$ at zero length and decreases in a monotone fashion as x increases.

The temperature dividing the two regions $\nu_0 > 1$ and $\nu_0 < 1$ can be found by noting that $\nu_0 = 1$ means that $\alpha_0 = 2\beta J$. The condition that λ_1' be equal to unity then leads to

$$a_{0,1} = 2(\beta J)^{-1/2}\, e^{-\beta(u+2Jq)/2} \tag{126}$$

The right-hand side is a continuous monotone decreasing function of β varying between zero and infinity. Thus (126) determines a unique β_0 or temperature T_0 above which $P_0^{(1)}(x)$ ceases to have a maximum and the concept of a local crystal thickness in the model breaks down.

Finally, we shall consider the decay of correlation between two segments as the number of folds between them goes to infinity. We define a segment pair correlation function by

$$g(x_1, n_1, x_2) = P^{(2)}(x_1, n_1, x_2)[P^{(1)}(x_1)P^{(1)}(x_2)]^{-1} - 1 \tag{127}$$

In the thermodynamic limit and above T_c it becomes

$$g_0(x_1, n_1, x_2) = \left[\sum_{i=1}^{\infty} (\lambda_i')^{n_1}\, \psi_i(x_1)\, \psi_i(x_2)\right][\psi_1(x_1)\, \psi_1(x_2)]^{-1} - 1 \tag{128}$$

Noting that $\lambda_1' > \lambda_2' > \lambda_3' > \cdots$, etc., we find that in the limit of large separation n_1 we have

$$g_0(x_1, n_1, x_2) \sim F(x_1, x_2)\exp\{-[\ln(1/\lambda_2')]n_1\}, \qquad n_1 \to \infty \tag{129}$$

$$F(x_1, x_2) = \psi_2(x_1)\psi_2(x_2)[\psi_1(x_1)\psi_1(x_2)]^{-1} \tag{130}$$

Thus we see that the correlation drops off exponentially at a rate determined by $\ln(1/\lambda_2')$ as the number of folds in between the chosen segments approaches infinity. In order to examine the decay of the correlation in the critical region $T \sim T_c$, we recall that as T approaches T_c, ν_0 goes to infinity. Using the following relations for zeros of Bessel functions of large order⁽¹⁰⁾

$$a_{\nu,1} \sim \nu + 1.856\nu^{1/3} + O(\nu^{-1/3}) \tag{131}$$

$$a_{\nu,2} \sim \nu + 3.244\nu^{1/3} + O(\nu^{-1/3}) \tag{132}$$

we then find

$$\lim_{T\to T_c+} \ln(1/\lambda_2') = \lim_{T\to T_c+} \ln(\lambda_1'/\lambda_2') = 0 \tag{133}$$

We conclude that the decay rate of the correlation at large separation vanishes as we approach the phase transition. This result supports the commonly accepted rule that critical correlations have infinite range. Furthermore, we note that the infinite range is due to the onset of asymptotic degeneracy in the largest eigenvalues as the temperature approaches T_c. Such behavior was first predicted by Ashkin and Lamb.⁽¹³⁾

6. SUMMARY AND DISCUSSION

The very strong influence of a few exactly soluble models, notably the Ising model in one and two dimensions, on the theory of phase transitions is well known. There is, of course, a limitation on the generalizability of the results obtained within any particular model. It becomes important to seek a wider range of soluble models in order to reduce excessive influence of model-specific behavior on the general theory. This is the basic reason for our interest in the continuous Zwanzig–Lauritzen model of polymer crystallization by chain folding. It does contain physically interesting interactions quite distinct from those of the Ising model while allowing the study of a second-order phase transition and other thermodynamic properties by analytical methods. The presence and form of the phase transition were established already by Zwanzig and Lauritzen in their original article, but their solution is rather limited in scope and not easily extended. We have therefore sought a new method of solution both simpler and more informative than the original one.

It was observed that the loose ends of the original model led to an asymmetry which complicated the mathematical analysis. Thus we introduced a symmetrized version of the model by letting the two end segments interact as if the molecule formed a closed ring. All of the calculations presented in the preceding sections concern this symmetrized model. In order to obtain a measure of the end effects of the molecule, calculations were carried out by the same methods for the original model also. The mathematics is substantially overlapping between the two models and somewhat more complicated in the original loose-end version. In order not to overburden this paper with mathematical detail, we shall state the most interesting results without proof.

Searching for both simplicity and power, we decided to make use of a grand ensemble, eliminating the troublesome length restriction in the canonical ensemble and the transfer matrix method well known from its successes in the work on Ising models. We quickly became aware of the unusual properties of the grand ensemble which has kept it from use in earlier work on similar models. It has no extensive parameters, there is substantial dispersion in system size even for large systems, and it obviously is not equivalent to the canonical ensemble in the customary way. Nevertheless, the grand ensemble can be derived from a principle of minimum information and, more importantly, we found that in the thermodynamic limit the canonical and grand averages of practically all interesting properties must coincide if the former averages exist.

The application of the transfer matrix method contains some unusual features also. In the one-dimensional Ising model a single particle is assigned to each site in a linear lattice and the states available to each site are discrete. Here, using the grand ensemble, we found it convenient to assign a varying length of the molecule to each site, and the set of states available to each site became continuous. Thus the transfer matrix became a kernel and the eigenvalue equation a homogeneous Fredholm integral equation of the second kind. However, it was still possible to carry out the program as established for finite-order transfer matrices. All the nonvanishing eigenvalues and corresponding eigenfunctions were found and after verifying that the kernel was symmetric and square-integrable we could employ the Hilbert–Schmidt theorem to see that the grand partition function was related to the eigenvalue spectrum in the manner suggested by the discrete examples.

The results made available by these new methods are plentiful and certainly not exhausted in this study. We easily reproduced the results concerning the phase transition obtained already in the original work on the model and slightly extended them by allowing the length of a fold to be greater than zero. The specific heat density diverges as the inverse square root of $T - T_c$ as T approaches T_c from above. The polymer molecule is in a folded mode above T_c characterized by a finite average segment length. As T decreases, the segments grow longer, to become infinite at $T = T_c$ and below T_c in the extended mode. Analytic forms were obtained for the probability density of the length of a segment and the joint probability densities describing correlations among two or more segments. Examining the one-segment probability density, we concluded that the concept of a local crystal thickness in the model breaks down at higher temperatures in the sense that the density no longer has a maximum but decreases in a monotone fashion. We also looked at the decay rate of critical correlations among two segments as a function of the number of segments separating them and we found, in accord with previous predictions, that the range of the critical correlations became infinite due to the onset of degeneracy among the two largest eigenvalues of the transfer matrix.

The above results were obtained in the thermodynamic limit for the symmetrized version of the ZL model. We subsequently extended the calculations to the original model in order to check on the end effects. There were no significant differences for long molecules in the folded mode above T_c. Below T_c, however, the loose ends caused a discontinuity in both the grand partition function and the average molecular lengths, which meant that the thermodynamic limit in the grand ensemble did not exist as defined.

Reviewing our results, we note that the ZL model provides an interesting illustration of some basic ideas of the general theory of phase transitions. First we recall the rule that one-dimensional systems do not have phase transitions unless the range of interaction is infinite. The ZL model can be thought of as being one-dimensional and it does have a phase transition. The interaction range is directly related to segment length. Noting that the average segment length increases as T decreases to become infinite at $T = T_c$, we see that the effective range of interaction becomes infinite at the phase transition.

It has been proposed(14) that the mathematical mechanism of a phase transition in the transfer matrix method is the onset of degeneracy in the largest eigenvalue. This is indeed the case in the present calculation on the ZL model. We found in a calculation not reported here that the whole eigenvalue spectrum collapses into a point as T approaches T_c from above. We have already noted that the critical correlations are of infinite range, as expected.

In this connection we should remark that a set of critical indices of the ZL model has recently been reported.(15) They were found to be different from the critical indices of other known transitions, but they did satisfy the familiar relations between critical exponents as exact identities.

In the light of our new solution we feel the ZL model must rank among the most soluble of the nontrivial models displaying a phase transition. As such, it should have great pedagogical value. The relevance of the model to the behavior of long polymer molecules is less clear, but the model may perhaps be dressed up to become more realistic without losing all of the nice mathematical properties that made the present study possible.

ACKNOWLEDGMENTS

This work was started at the University of Maryland while I was a student of Professor Robert Zwanzig. I am indebted to Professor Zwanzig, who suggested this problem and directed my efforts at the outset. My thanks also go to Professor Stuart A. Rice, who provided the opportunity to complete this work at the James Franck Institute and who read the manuscript.

During the course of this work I have benefitted from support by the National Science Foundation at the University of Maryland and by the National Science Foundation at the University of Chicago.

REFERENCES

1. R. Zwanzig and J. I. Lauritzen, Jr., *J. Chem. Phys.* **48**:3351 (1968).
2. J. I. Lauritzen, Jr. and R. Zwanzig, *J. Chem. Phys.* **52**:3740 (1970).
3. G. F. Newell and E. W. Montroll, *Rev. Mod. Phys.* **25**:353 (1953).
4. S. Lifson and B. H. Zimm, *Biopolymers* **1**:15 (1963).
5. T. L. Hill, *Statistical Mechanics*, McGraw-Hill, New York (1956), Chapter 3.
6. F. G. Tricomi, *Integral Equations*, Interscience, New York (1957).
7. M. Abramowitz and J. A. Stegun, *Handbook of Mathematical Functions*, U. S. Government Printing Office, Washington, D. C.
8. A. Erdelyi *et al.*, *Higher Transcendental Functions*, McGraw-Hill, New York (1953), Vol. 2, p. 61.
9. S. Lifson, *J. Chem. Phys.* **40**:3705 (1964).
10. F. W. J. Olver (ed.), *Royal Society Mathematical Tables*, Cambridge University Press, Cambridge, England, (1966), Vol. 7, part III, p. xviii.
11. E. A. Guggenheim, *J. Chem. Phys.* **7**:103 (1939).
12. E. T. Jaynes, *Phys. Rev.* **106**:620 (1957).
13. J. Ashkin and W. E. Lamb, Jr., *Phys. Rev.* **64**:159 (1943).
14. M. Kac, in *Statistical Physics, Phase Transitions, and Superfluidity*, M. Chretien, E. P. Gross, and S. Deser, eds., Gordon and Breach, New York (1968), Vol. 1, p. 241.
15. M. J. Cooper, *J. Chem. Phys.* **52**:473 (1970).