Necklace-cloverleaf transition in associating RNA-like diblock copolymers

M. V. Tamm[1,2] and S. K. Nechaev[1,3]
[1]*LPTMS, Université Paris Sud, 91405 Orsay Cedex, France*
[2]*Physics Department, Moscow State University 119992 Moscow, Russia*
[3]*Landau Institute for Theoretical Physics, 117334 Moscow, Russia*
(Received 20 June 2006; revised manuscript received 9 October 2006; published 7 March 2007)

We consider a A_mB_n diblock copolymer, whose links are capable of forming local reversible bonds with each other. We assume that the resulting structure of the bonds is RNA like—i.e., topologically isomorphic to a tree. We show that, depending on the relative strengths of *A*-*A*, *A*-*B*, and *B*-*B* contacts, such a polymer can be in one of two different states. Namely, if a self-association is preferable (i.e., *A*-*A* and *B*-*B* bonds are comparatively stronger than *A*-*B* contacts), then the polymer forms a typical randomly branched cloverleaf structure with the so-called roughness exponent $\gamma=1/2$. On the contrary, if alternating association is preferable (i.e., *A*-*B* bonds are stronger than *A*-*A* and *B*-*B* contacts), then the polymer tends to form a generally linear necklace structure with $\gamma=1$. The transition between cloverleaf and necklace states is studied in detail, and it is shown that it is a second-order phase transition.

DOI: 10.1103/PhysRevE.75.031904 PACS number(s): 87.15.−v, 02.30.Fn, 82.35.Jk

I. INTRODUCTION

It would not be a strong exaggeration to say that since the very beginning, the statistical physics of macromolecules was driven mainly by biological motivations. The problems of phase rearrangements in biopolymers, such as proteins and DNA and RNA molecules, have stimulated an intensive development of new theoretical approaches aiming to describe and to predict the structural, conformational, and functional changes in biopolymers under variation of external conditions (see, for example, [1–3]).

A very important biological role has a wide class of so-called "associating" polymers. The associating polymers, besides the strong covalent interactions responsible for the very chainlike structure of a macromolecule with frozen sequence ("primary structure") of monomer units in the chain, are capable of forming additional weaker reversible temperature-dependent (i.e., "thermoreversible") bonds between different monomer units. Many biologically important macromolecules, like proteins and nucleic acids, belong to the class of associating polymers. It is known that just the presence of thermoreversible bonds is crucial for formation of various equilibrium spatial conformations of such macromolecules (so-called ternary structures) and, hence, affects the biological activity of these macromolecules [4].

The physics of associating polymers is typical for the statistical mechanics of complex systems and consists in an interplay between several entropic and energetic factors. In the particular case of associating polymers one can pick out the following three major factors, which, together with the primary structure of the chain, give rise to all the variety of possible thermodynamic states in these systems: (i) the energy gain due to the formation of thermoreversible contacts, (ii) the combinatoric entropy due to the choice of which particular monomers (among those able to participate in bonds formation) actually do create bonds, and (iii) the loss of conformational entropy of the polymer chain due to bond formation and, in particular, the entropic penalty of loop creation between two bonded monomers.

Among a variety of macromolecular systems with thermoreversible interactions we pay special attention to a class of so-called "RNA-like" polymers. These polymers are distinguished from other biologically active associating polymers, such as, for instance, proteins, by a capability of forming only "cloverleaflike" (or "cactuslike") secondary structures. Indeed, the formation of a thermoreversible contact between two distant bonds in a RNA molecule (or in a single-stranded DNA molecule) imposes a nonlocal constraint on a number of other possible bonds: all bonds in a RNA chain are known to be arranged in a way to allow only hierarchical cactuslike folded conformations topologically isomorphic to a tree. The pairs of bonds which do not obey such a structure are called "pseudoknots" and are forbidden for RNA-type molecules. A quantitative definition of conformations allowed and forbidden in RNA-like polymers is described in Sec. II. Note that for proteins such a constraint is relaxed.

The physical reason for the aforementioned constraint is not quite clear. It is often believed that the absence of "pseudoknots" is due to the strong cooperativity effects in real RNA and DNA. This is schematically explained in Figs. 1(a) and 1(b). The presence of pseudoknots in a strong cooperativity regime [Fig. 1(b)] leads due to topological reasons to the creation of a "domain wall" separating saturated (*A*) and nonsaturated (*B*) bonds. Such domain walls are absent in cactuslike configurations [see Fig. 1(a)]. The loss of complimentarity costs some energy and hence has an influ-

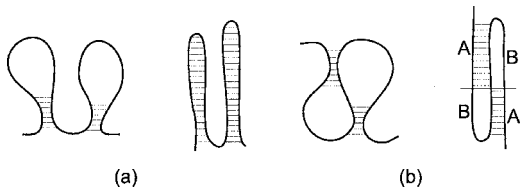

FIG. 1. (a) The treelike topology of the chain in the presence of strong cooperativity. (b) The pseudoknot in the presence of strong cooperativity.

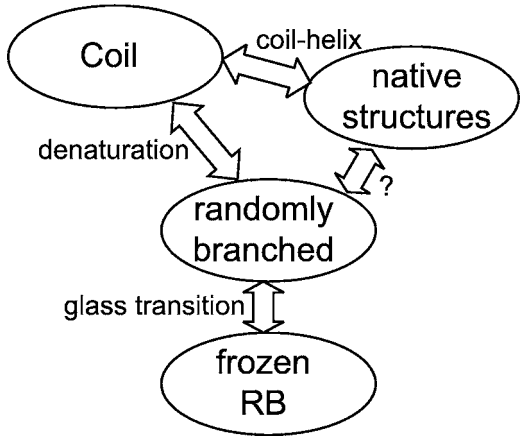

FIG. 2. Schematic picture of possible thermodynamic states of RNA-like heteropolymers.

ence on the thermodynamics of the considered system. However, in many theoretical works (see, for example, [5–10]) a discussion of the physical reasons which cause the treelike structure of RNA-type molecules is skipped. We shall follow this line and accept the absence of pseudoknots as just a matter of fact.

There are plenty of works concerning the statistical properties of associating polymers without any constraints on the topology of thermoreversible contacts [11,12]. For such systems one can construct a satisfactory mean-field theory [11]. However, RNA-like polymers are much less investigated.

It is known that depending on the temperature (which comes into play via the temperature dependence of the statistical weights of the bonds) and on the primary structure of the chain, the RNA-like polymer may be found in the following four distinct states—see Fig. 2. First, at high enough temperature any RNA-like molecule forms a usual Gaussian or non-self-intersecting (depending on the quality of the solvent) coil with no, or almost no, reversible bonds formed. Second, at lower temperature it may form a highly branched "genuine cactuslike" state topologically isomorphic to a randomly branched tree. Third, there is strong evidence that at low enough temperatures the RNA molecules with highly irregular primary structure tend to form glassy states. Finally, in the case of some special primary structures, there exists an energy gap between the ground (so-called "native") energy state of the RNA and all the other states. The RNA molecule can, therefore, in this case form a highly deterministic secondary structure (as compared to annealed random and quenched random states in the previous two cases).

In particular, if the primary structure of a RNA molecule consists of two (or several) complimentary parts, it tends to form a double-folded (or, in the case of several parts, a crosslike, as in typical tRNA) structure. Such a structure, contrary to a randomly branched one, is highly deterministic—one can predict easily which particular monomers will aggregate with each other—and has a different asymptotic (for large chain lengths N) mean-square end-to-end distance R. For example, if one neglects the volume interactions (i.e., the nonsaturating interactions between monomers which are far from each other along the chain), $R_{\mathrm{nc}} \sim N^{1/2}$ for necklace (or "double-folded") conformation, rather than $R_{\mathrm{rb}} \sim N^{1/4}$ in the randomly branching case.

Let us give a sketch of what is known about the transitions between the four states depicted in the Fig. 2. First of all, the transition between the single-stranded and double-stranded RNA (or DNA) molecules commonly known as a helix-coil transition, which is a typical representative of the transition between the coil and the deterministic ground state, is rather well understood [13]; the modern developments in this field can be found in [14]. It is known that the particular form of loop penalties plays crucial role in whether it is a genuine phase transition or just a collective one. The contribution to the polymer partition function emerging from the loops is usually supposed to have the scaling form N^{α} (where N is the loop length)—see, for example, [13]. If $0 \leq \alpha < 1$ ($\alpha=0$ corresponds to the absence of any loop contribution), the passage from a coil to a helix is not a phase transition, but a cooperative phenomenon with a more or less (depending on the value of α and on the cooperativity of bond formation) sharp crossover. On the contrary, if $\alpha \geq 1$, there is a genuine phase transition. The situation concerning randomly branched RNA molecules is more complicated. Despite the fact that the denaturation of a cactuslike RNA molecule has been investigated for several decades, starting from classical works [3,15], it was shown only recently [16] that the character of the transition in RNA resembles in some aspects the coil-helix transition in DNA. Namely, for large loop penalties ($\alpha \geq 2$ for RNA) one has a genuine phase transition, while for smaller loop penalties there is just a crossover behavior.

Note that in both aforementioned cases the heteropolymer nature of real RNA structure played a minor role in the transition under discussion: one could have considered all monomers as identical and still have a proper insight into what goes on. On the contrary, in the problem of glass transitions in RNA-like polymers, the frozen heteropolymer structure of a chain plays a crucial role: the presence of frustrations in the system is itself a peculiarity of heterogenous systems [6]. There is, hence, a barest necessity for a statistical theory of *heteropolymers* forming RNA-like structures. The investigation of some thermodynamic properties of random heteropolymers is addressed in a few recent theoretical papers [6–8,17]. On the other hand, the theoretical results are usually obtained not for "real" heteropolymers with a finite number of link types, but for polymers with a frozen Gaussian distribution of association constants, sometimes even excluding the loop factor. It is believed (see, for example, [6]) that the loop penalties are less significant for understanding the physical properties of RNA-like molecules at low temperature (i.e., close to the glass transition) when the typical sizes of the loops become rather small.

We would like to stress that to the best of our knowledge, a theoretical consideration of the possible transition between regular (for example, double-stranded) and randomly branched cactuslike states of RNA is absent. It seems to be instructive, therefore, to consider a model of a RNA-like polymer which shares the peculiarity of a chain nonuniformity in a way which can give rise to the formation of linear double-stranded structure, with an advantage of being exactly solvable. Presenting here such a model, we believe that

besides solving a specific model of the transition between double-stranded and randomly branched states of RNA, we also provide better insight into the structure of possible states of heteropolymer RNAs molecules, thus contributing in the construction of the general theory of heteropolymers forming RNA-like structures.

Specifically, we consider in this paper an

$$A_m B_n \equiv \overbrace{AA\ldots A}^{m \text{ times}}\ \overbrace{BB\ldots B}^{n \text{ times}}$$

diblock copolymer, whose links can form local reversible bonds with each other. We show that, depending on the relative strengths of *A*-*A*, *A*-*B*, and *B*-*B* bonds, such a polymer can be in one of two different states. Namely, if a self-association is preferable (i.e., *A*-*A* and *B*-*B* bonds are comparatively stronger than *A*-*B* ones), then the polymer forms a typical randomly branched cloverleaf structure. On the contrary, if alternating association is preferable (i.e., *A*-*B* bonds are stronger than *A*-*A* and *B*-*B* contacts), then the polymer tends to form a generally linear necklace structure somewhat reminiscent of a double-stranded DNA (with, probably, some rear side branches and loops, which, however, do not influence the overall characteristics of the chain). These two states can be clearly distinguished by the value of the so-called roughness exponent γ to be defined in Sec. II. In this paper we consider the transition between these two states (cloverleaf and necklace) in detail and show that it turns out to be a second-order phase transition.

We would like to mention that when we had already finished the draft of this paper two articles [9,10] considering systems rather similar to one described above came to our knowledge. The authors of [9] have found a similar necklace-cloverleaf transition in the problem of associating two RNA-like homopolymers of similar length. Actually, their result exactly corresponds to the $\alpha=1/2$ case of our Eqs. (15) and (16). In turn, the authors of [10] consider the formation of a complex of two single-stranded DNA molecules, which are assumed to be not capable of self-association. This system corresponds exactly to putting $v_1=v_2=0$ in our formulas. One can consider the current paper as a generalization of the results of these two articles.

The paper is organized as follows. In Sec. II we discuss the model under consideration and derive the basic equations for the partition function. In Sec. III, which plays the central role in the paper, we solve these equations in the approximation of no loop penalties and study the case of large equal lengths of blocks, $m=n\gg 1$. We show that a necklace-cloverleaf transition is a second-order phase transition, compute the corresponding phase diagram, and discuss the behavior of the free energy near the transition point. In Sec. IV we solve the basic equations with "ideal" loop weights ($\alpha=3/2$) and show that the presence of these loop factors does not have a strong influence on the phase transition. Finally, in Sec. V we summarize the results and briefly discuss the perspectives.

II. THE MODEL

Consider an A_nB_m diblock copolymer whose monomers *A* and *B* are capable of forming thermoreversible saturating bonds of all possible types *A*-*A*, *B*-*B*, and *A*-*B*. ("Saturating" in this context means that each monomer can participate at most in one bond.) Let the statistical weights of bonds be equal to v_1, v_2, and u, respectively. Having applications to RNA molecules in mind, assume also that the structures formed by these thermoreversible bonds are always of a tree-like type, as shown in Fig. 3(a). It means that we restrict ourselves to the situation in which the chain conformations with "pseudoknots" shown in Fig. 3(b) are prohibited.

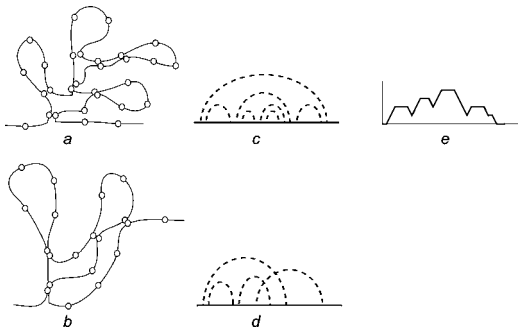

FIG. 3. (a),(b) Schematic picture of allowed (a) cactuslike and prohibited (b) pseudoknot configurations of the bonds; (c),(d) Arc diagrams corresponding to configurations (a) and (b), respectively [note the intersection of arcs in (d)]; (e) The height diagram corresponding to the bond configuration (a).

This definition can be formalized as follows. Take a chain $\mathcal{C}$ and enumerate its links: $1,2,\ldots,n$. Consider two different thermoreversible bonds $b(i_1,i_2)$ and $b(j_1,j_2)$, where $b(i_1,i_2)$ corresponds to a contact between two monomers i_1 and i_2 (where $1\le i_1<i_2\le n$) and $b(j_1,j_2)$ is formed by monomers j_1 and j_2 (where $1\le j_1<j_2\le n$). We say that the chain $\mathcal{C}$ forms an RNA-like structure if and only if *for any two different* bonds $b(i_1,i_2)$ and $b(j_1,j_2)$ the following inequality holds:

$$(j_1-i_1)(j_2-i_1)(j_1-i_2)(j_2-i_2)>0. \tag{1}$$

To make the difference between allowed and not allowed structures geometrically more transparent, redraw the structures shown in Figs. 3(a) and 3(b) in the following way. Represent a polymer under consideration as a straight line with active monomers situated along it in the natural order, and depict the thermoreversible bonds by dashed arcs connecting the corresponding monomers. Now, the absence of pseudoknots simply means that the arcs do not intersect—see Figs. 3(c) and 3(d).

Moreover, we assume, for simplicity, that except pseudoknots, all other bond configurations are allowed. This means, in particular, that we do not require any minimal loop length, as well as we do not yet take into account the cooperativity effect (the fact that if two links are connected with each other, then the two adjacent links have larger probability to be also connected). These assumptions are known to be false for real RNA molecules (for example, there are no loops shorter than three monomers in RNA chains [16]). However, one can speculate that—see, for example, [18]—if the links of the chain are considered as renormalized quasi-

monomers consisting of several "bare" monomer units, the assumptions made seem to be plausible.

Let us also mention another possible representation of the secondary structures (configurations of bonds) allowed in the system under discussion. Consider again the arc diagram representation. Let then unconnected monomers, the monomers at which any arc starts and those where it ends, correspond to vectors $(1,0)$, $(1,1)$, and $(1,-1)$, respectively. Then any bond configuration is mapped to some sequence of vectors which can be considered as a random walk in $1+1$ dimensions [$(1+1)$D] [see 3(e)] situated in the upper half-plane. The walk begins and ends on the x axis; the end points are separated by the distance N—the total length of the polymer. Such a representation is often called the "height diagram" of the secondary structure. Note that the height of the point in the phase diagram is exactly equal to the number of arcs going above the corresponding point on the arc diagram—i.e., to the number of bonds one has to break to reach the corresponding monomer from the starting point of the chain. The critical exponent γ connecting the mean height of such a diagram with the length of the polymer ($\langle h\rangle \sim N^{\gamma}$, $0 \leqslant \gamma \leqslant 1$), called the roughness exponent, is an important characteristic of the state of the system. In the randomly branched homopolymer state $\gamma_{\mathrm{br}}=1/2$, while in the glassy state of random RNA it was found [7] numerically $\gamma_{\mathrm{gl}} \simeq 2/3$, and in the recent work [8] the analytical estimate $\gamma_{\mathrm{gl}} \simeq 5/8$ has been obtained for the same system via the renormalization group procedure.

Consider now the topology of the bonds formed in the AB copolymer. One can easily see that A-B bonds separate the independent regions along the chain. Let us start moving along the chain from the left end and find the first A-B bond. As the arcs cannot intersect each other (see Fig. 3), all the monomers under the given arc can be connected only with each other. Moreover, the A monomers in the beginning of the chain can form only A-A bonds (as we are speaking about the *first* A-B bond) and the B monomers in the rest of the chain can form only B-B bonds with each other, too. This separation of a chain into three independent parts allows us to write down the following Dyson-type equation for the partition function of the whole chain:

$$G_2(n,m,\mathbf{q}|v_1,v_2,u) = G(n,\mathbf{q}|v_1)G(m,\mathbf{q}|v_2) + u\sum_{i,j} G(i,\mathbf{q}|v_1)G(j,\mathbf{q}|v_2)G_2^{(\mathrm{loop})}(n-i-1,m-j-1|v_1,v_2,u), \quad (2)$$

where $G_2(n,m,\mathbf{q}|v_1,v_2,u)$ is the desired partition function (written in the $\mathbf{q}$ space), $G(k,\mathbf{q}|v)$ is the partition function of a self-associating chain of A or B monomers of length k (we assume for simplicity that the only difference between A and B monomers is in the weight of bond formation—respectively, v_1 and v_2), and $G_2^{(\mathrm{loop})}(n,m|v_1,v_2,u)$ is a partition function of a diblock copolymer which is forced to form a loop (i.e., its end monomers are forced to form a bond). Finally, the first term in the sum on the right-hand side (RHS) of Eq. (2) is due to the fact that there can be no A-B bond at all. Note that to solve Eq. (2), apart from defining the homopolymer partition functions G, one should allow for some particular connection between the partition functions of "simple," G_2, and "looped," $G_2^{(\mathrm{loop})}$, chains.

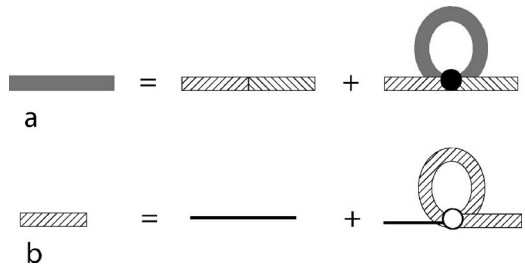

FIG. 4. Diagrammatic form of Dyson-type equations: (a) for Eq. (4) and (b) for Eq. (5). Thin lines correspond to free propagators $g(s,\mathbf{q})$, thick hatched lines to the homopolymer propagators $G(s,\mathbf{q}|v)$, and thick gray lines to the diblock copolymer propagators $G_2(s_1,s_2,\mathbf{q}|v_1,v_2,u)$. Finally, open and solid circles correspond to the intraspecies and interspecies bonds, respectively.

Introducing the generating functions for all relevant partition functions,

$$G_2(s_1,s_2,\mathbf{q}|v_1,v_2,u) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} G_2(m,n,\mathbf{q}|v_1,v_2,u)s_1^m s_2^n,$$

$$G(s,\mathbf{q}|v) = \sum_{m=0}^{\infty} G(m,\mathbf{q}|v)s^m,$$

$$G_2^{(\mathrm{loop})}(s_1,s_2|v_1,v_2,u) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} G_2^{(\mathrm{loop})}(m,n|v_1,v_2,u)s_1^m s_2^n, \quad (3)$$

we can rewrite Eq. (2) in the following form:

$$G_2(s_1,s_2,\mathbf{q}|v_1,v_2,u) = G(s_1,\mathbf{q}|v_1)G(s_2,\mathbf{q}|v_2) + us_1s_2G(s_1,\mathbf{q}|v_1)G(s_2,\mathbf{q}|v_2) \times G_2^{(\mathrm{loop})}(s_1,s_2|v_1,v_2,u) \quad (4)$$

[see Fig. 4(a) for the diagrammatic form of this equation]. Moreover, the partition functions of a homopolymer forming cloverleaf structures (i.e., structures without pseudoknots) are known (see, for example, [5,15,16]) to satisfy the equation

$$G(s,\mathbf{q}|v) = g(s,\mathbf{q}) + vg(s,\mathbf{q})G(s,\mathbf{q}|v)G^{(\mathrm{loop})}(s|v) \quad (5)$$

[see Fig. 4(b) for its diagrammatic form] and the function $g(s,\mathrm{q})$ is a partition function of a free chain (i.e., chain without any thermoreversible bonds or volume interactions).

The particular form of $g(s,\mathbf{q})$ is known to be irrelevant for long enough chains (see, for example, [13]). In our case, when a chain can form short-ranged loops, the particular form of the bare propagator influences importantly the entropy penalties for closing a chain into a loop. Note, however, that we introduce the loop penalties explicitly via the connection between G and $G^{(\mathrm{loop})}$ in Eq. (5) and therefore we assume the form of the function $g(s,q)$ to be irrelevant. We can thus set without the loss of generality that

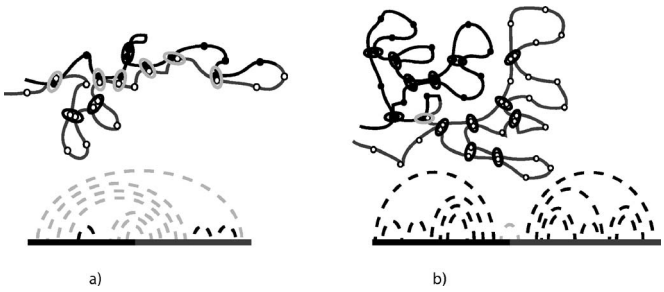

FIG. 5. Sketches of typical conformations of diblock copolymers in the necklace (a) and cloverleaf (b) phases. *A* block is represented by solid circles on the black string, *B* block by open circles on the dark gray string. Intraspecies (*A*-*A* and *B*-*B*) and interspecies (*A*-*B*) bonds are shown by black and light gray, respectively.

$$g(N,R) = \left(\frac{3}{2\pi N a^2}\right)^{3/2} \exp\left(-\frac{3R^2}{Na^2}\right), \tag{6}$$

where a is a mean length of a covalent bond between two subsequent monomers. After the Fourier-Laplace transform one gets

$$g(s,\mathbf{q}) = \frac{1}{e^{\mathbf{q}^2 a^2/6} - s}. \tag{7}$$

In what follows we chose the length units to be $a^2/6=1$. Note that while the Gaussian form of the free chain propagator is a matter of assumption, the value of the propagator at $q=0$ [namely, $(1-s)^{-1}$] is independent of its particular spatial dependence. Indeed, the partition function of a free chain with a given number of monomers $Z(N)=\int Z(N,r)d^3r$ is by definition equal to unity (see [13]). Therefore, the value of $g(q,s)$ at $q=0$, which is just $\Sigma_{N=0}^{\infty} Z(N)s^N$, equals always $g(0,s)=1/(1-s)$.

The crucial difference between Eqs. (4) and (5) is that the former is linear. This corresponds to the aforementioned fact that the *A*-*B* bonds separate the chain into independent parts. Therefore the global structure of a bonded chain is just a linear sequence of independent "blobs" separated by *A*-*B* bonds; only *A*-*A* and *B*-*B* bonds are allowed inside each blob. This notion leads one to an understanding of the main feature of the system under consideration: the possibility of a necklace-cactus transition. Indeed, for $v_{1,2} \gg u$ one can neglect the alternating *A*-*B* bonds and the chain forms just two independent "cactuses" consisting of *A* and *B* monomers, the roughness exponent γ being equal to 1/2 similarly to the homopolymer case. On the other hand, if $u \gg v_{1,2}$, the alternating bonds dominate in the structure and the chain forms a "hairpinlike" structure with $\gamma=1$. Even a stronger conjecture is true: if the fraction of *A*-*B* bonds is nonzero in the thermodynamic limit, the roughness exponent γ is equal to 1. In Fig. 5 we show the typical conformations of a "hairpin" and a "cactuslike" phase.

We are going now to analyze the described transition in more detail. To close the system of equations we should now address the aforementioned problem of the connection between "simple" and "looped" partition functions. In the two following sections we deal with two most usually considered closures, those being no loop weights (Sec. III) and "ideal" loop weights (Sec. IV).

To be specific, in Sec. III we assume the most simple possible situation. Namely, we postulate

$$G_2^{(\mathrm{loop})}(s_1,s_2) = G_2(s_1,s_2,\mathbf{q}=0) = \int G_2(s_1,s_2,\mathbf{r})d\mathbf{r},$$

$$G^{(\mathrm{loop})}(s) = G(s,\mathbf{q}=0) = \int G(s,\mathbf{r})d\mathbf{r}. \tag{8}$$

Here the functions $G(s,\mathbf{r})$ and $G(s_1,s_2,\mathbf{r})$ are the partition functions of the homopolymer and copolymer chains in $\mathbf{r}$ space. The assumption (8) corresponds obviously to omitting any entropic penalties for loop formation. We suppose that such an approximation can be valid at least for low enough temperatures; for the justification of this assertion, see the beginning of Sec. III.

In Sec. IV we assume a more natural connection between the partition functions under discussion:

$$G_2^{(\mathrm{loop})}(s_1,s_2) = \frac{1}{(2\pi)^3}\int G_2(s_1,s_2,\mathbf{q})d\mathbf{q} = G_2(s_1,s_2,\mathbf{r}=\mathbf{0}),$$

$$G^{(\mathrm{loop})}(s) = \frac{1}{(2\pi)^3}\int G(s,\mathbf{q})d\mathbf{q} = G(s,\mathbf{r}=\mathbf{0}), \tag{9}$$

where the RHS is written also in terms of the partition function in $\mathbf{r}$ space. Thus, we simply assume here that the partition function of a loop (i.e., chain, actually bonded at the end points) equals the partition function of a similar chain without end-to-end bond, but whose end monomers are situated at the same point in space. This assumption at first glance looks almost obvious, and it is hard to understand why it could fail. Note, however, that the above equations do not allow for volume interactions—i.e., interactions of nonsaturating nature between monomers, situated far from each other along the chain, but close to each other in the real space. These interactions, however, do not influence the topology of possible cloverleaf conformations. Having in mind the discussion of the major factors governing the problem of a RNA-type folding, which we have presented in Sec. I, one can hope to take the volume interactions into account via altering the loop weight function or, in our formalism, by choosing the proper connection between the "looped" and "nonlooped" partition functions in the form different from the ideal one (9).

III. NO LOOP WEIGHTS

In this section we calculate the partition function of the chain with no penalties for loop formation. To justify this simplification let us recall that the loop weight is supposed to be relevant only if the loops formed in the system are long enough. Indeed, the loop weight—i.e., the probability that a polymer chain of length l forms a loop—is, at least in the first approximation, proportional to R_l^{-D}, were R_l is the typical end-to-end distance of such a chain and D is the space

dimensionality. In turn, R_l is proportional to some positive power of l depending on the state of the system. Therefore, the loop weights are of order of unity for short loops and much less than unity for long ones. In other words, the loop weights effectively suppress the long loops. If such loops are already suppressed (which is indeed the case if $u, v_{1,2} \gg 1$, when almost all the monomers are involved into bonds), the particular form of loop weights is not very relevant. Hence, for large enough values of association constants (i.e., for low enough temperature) the no-loop-weights approximation seems to be realistic.

Keeping this in mind, we can simplify Eqs. (4) and (5) as follows. We substitute the total partition functions $G(\mathbf{q}=0)$ and $G_2(\mathbf{q}=0)$ instead of loop ones $G^{(\mathrm{loop})}$ and $G_2^{(\mathrm{loop})}$ and then set $\mathbf{q}$ equal to zero. We arrive therefore at a set of two algebraic equations

$$G(s|v) = g(0,s) + v g(0,s) s^2 G^2(s|v),$$

$$G_2(s_1,s_2|v_1,v_2,u) = G(s_1|v_1)G(s_2|v_2) + u s_1 s_2 G(s_1|v_1)G(s_2|v_2) \times G_2(s_1,s_2|v_1,v_2,u) \tag{10}$$

The solutions to (10) read

$$G(s,v) = \frac{1 - s - \sqrt{(1-s)^2 - 4vs^2}}{2vs^2} \tag{11}$$

for the homopolymer partition function (we have chosen the branch of solutions which approaches $G=(1-s)^{-1}$ as $v \to 0$) and

$$G_2(s_1,s_2) = \frac{1}{G^{-1}(s_1,v_1)G^{-1}(s_2,v_2) - u s_1 s_2} = \left(\frac{4 v_1 v_2 s_1^2 s_2^2}{[1 - s_1 - \sqrt{(1-s_1)^2 - 4v_1 s_1^2}][1 - s_2 - \sqrt{(1-s_2)^2 - 4v_2 s_2^2}]} - u s_1 s_2 \right)^{-1} \tag{12}$$

for the block-copolymer partition function.

Note that although we have introduced the generating functions $G(s)$ and $G_2(s_1,s_2)$ as purely formal objects, one can ascribe a physical meaning to the result obtained in Eq. (12). Indeed, imagine a polydisperse system of AB copolymers in the equilibrium with a reservoir of A and B monomers. The length of A and B blocks will then be controlled by the corresponding fugacities s_1 and s_2, and the partition function of such a block copolymer is given by Eq. (12). An important note is that this situation would be fully annealed with respect to the lengths of the blocks, while in most (at least biologically relevant) experimental cases the length and primary structure of a polymer are quenched. In what follows we are going to consider the quenched case when the length of the blocks is fixed and does not depend on the particular form of the interaction between monomers.

It seems now instructive to study the analytic structure and determine singularities of Eq. (12), because the particular form of these singular points is responsible for the equilibrium conformations of the system under consideration. Indeed, one can easily see that there are three possible singularities of Eq. (12): two square-root singularities at the points

$$v_1 = \left(\frac{1 - s_1}{2 s_1} \right)^2 \quad v_2 = \left(\frac{1 - s_2}{2 s_2} \right)^2 \tag{13}$$

and the simple pole singularity at the point

$$\frac{4 v_1 v_2 s_1 s_2}{u} = \left[1 - s_1 - \sqrt{(1-s_1)^2 - 4 v_1 s_1^2} \right] \left[1 - s_2 - \sqrt{(1-s_2)^2 - 4 v_2 s_2^2} \right]. \tag{14}$$

The square-root singularity corresponds (see, for example, [16]) to the formation of a usual randomly branched cloverleaf structure with $\gamma = 1/2$ due to the self-association of monomers of type A (or B), while the simple pole signals the formation of a linear ($\gamma = 1$) double-folded conformation due to the alternating A-B association. Thus, if the square-root singularity "wins" (i.e., becomes more important), the global conformation of a chain is randomly branched, while if the pole singularity "wins," the global conformation is linear. One can imagine also a situation when (say, for a very long A block and short B block) there can arise also a mixed situation, when both singularities are of equal importance or, saying it in more physical words, when within a single polymer there is a phase equilibrium between a necklace (double-stranded) phase and a homopolymer (in this case A) cloverleaf phase. However, as we show later on, this situation actually never exists in the particular system under discussion.

If a system is controlled by one single activity s, the meaning of "winning singularity" is well defined: in the thermodynamic limit the singularity closest to $s=0$ is dominant. In the case under consideration we have two independent activities s_1 and s_2, which makes the situation somewhat more tricky. However, one can handle this problem in the case of two activities considering the thermodynamic limit $m = \alpha N$, $n = (1-\alpha)N$ when $N \to \infty$ and $\alpha = \mathrm{const}$.

Here we present only the results, while the details of the algorithm we are using are explained in Appendix A and the details of particular computations in Appendix B.

Depending on the particular value of the association constants $\{u, v_1, v_2\}$ either the square-root or the pole singularity becomes dominant. So we have two options.

(i) If $u < u_{\mathrm{tr}} = \sqrt{v_1 v_2}$, then the square-root singularity is dominant and the system is in the cloverleaf phase with the

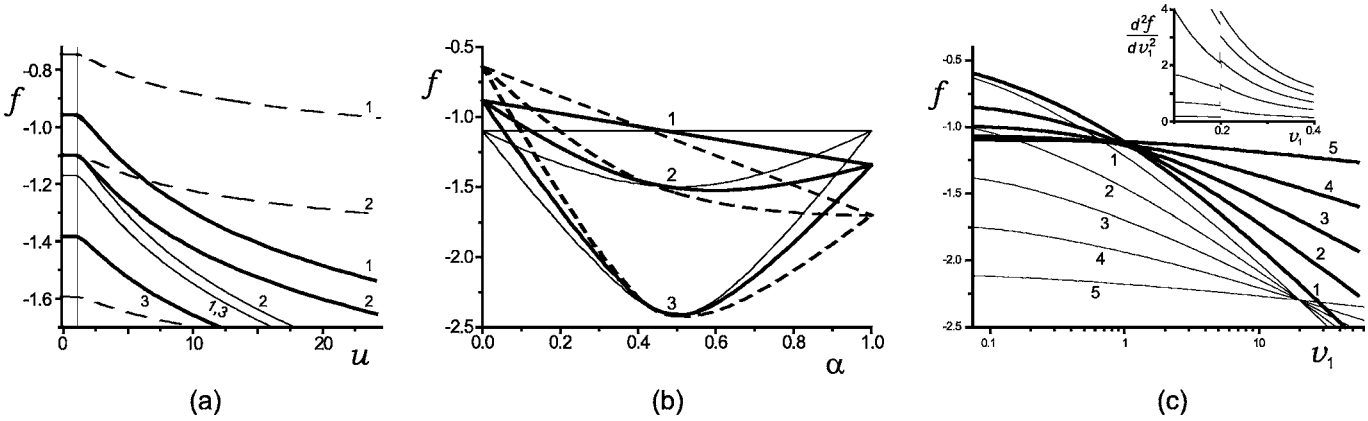

FIG. 6. Free energy of the system defined by Eqs. (15) and (16) as a function of u (a), υ_1 (b), and α (c). (a) (b) Thin lines correspond to $\upsilon_1=\upsilon_2=1$, thick lines to $\upsilon_2=2$, $\upsilon_1=0.5$, and dashed lines to $\upsilon_2=5$, $\upsilon_1=0.2$. (a) Curves corresponding to $\alpha=0.1$, 0.5, and 0.9 are marked by numbers 1, 2, and 3, respectively. (b) Curves corresponding to $u=1$ (the transition point), 10, and 100 are marked by numbers 1, 2, and 3, respectively. (c) Thick lines correspond to $u=2$, $\upsilon_2=1$, thin lines, both in the main figure and the inset, to $u=2$, $\upsilon_2=20$. Curves marked by 1, 2, 3, 4, and 5 correspond to $\alpha=0.9$, 0.7, 0.5, 0.3, and 0.1, respectively. Note that the thin lines intersect at the same point $\upsilon_1=\upsilon_2$ which, in this case, belongs to the cloverleaf phase. This is not true for the thick lines since the point $\upsilon_1=\upsilon_2$ belong to the necklace phase in this case.

free energy per link of the chain f asymptotically equal (when $N\to\infty$) to Eq. (B3):

$$f_{\text{clov}} = \ln s^*_{\text{sqr}} = -\alpha \ln(1+2\sqrt{\upsilon_1}) - (1-\alpha)\ln(1+2\sqrt{\upsilon_2}). \tag{15}$$

If $u > u_{\text{tr}}$, then the pole singularity is dominant and the polymer is in the necklace conformation. The corresponding free energy (B9) reads

$$f_{\text{neck}} = \ln s^*_{\text{pole}} = -\alpha \ln[1+2\sqrt{\upsilon_1}\cosh(\vartheta-\xi)] - (1-\alpha)\ln[1 + 2\sqrt{\upsilon_2}\cosh(\vartheta+\xi)], \tag{16}$$

where

$$\vartheta = \frac{1}{4}\ln\frac{u^2}{\upsilon_1\upsilon_2} \tag{17}$$

and ξ is the solution of the equation

$$4\sqrt{\upsilon_1\upsilon_2}\sinh 2\xi + 4(1-2\alpha)\sqrt{\upsilon_1\upsilon_2}\sinh 2\vartheta - \alpha\sqrt{\upsilon_1}\sinh(\vartheta-\xi) + (1-\alpha)\sqrt{\upsilon_2}\sinh(\vartheta+\xi) = 0 \tag{18}$$

on the interval $(-\vartheta,\vartheta)$.

In Fig. 6 we have plotted the free energy of the system defined by Eqs. (15) and (16) as a function of the variables u, υ_1, and α — see Figs. 6(a)–6(c), respectively.

The free energy does not depend on u in the cloverleaf phase, and it diverges quadratically at the transition point from the cloverleaf value, thus giving rise to the discontinuity of its second derivative. For $f(\upsilon_1)$ dependences this discontinuity is less evident. To make it more clear we have plotted in the inset the function $\frac{d^2f(\upsilon_1)}{d\upsilon_1^2}$. The transition from the cloverleaf to the necklace phase is, thus, a second-order phase transition. Moreover, it is very important that the function $f(\alpha)$ is always a convex curve. This fact provides support for our assertion that there is no phase separation inside the system under consideration. Indeed, if there were regions where the free energy as a function of α is concave, the system would have minimized its energy in the usual way by separating into two necklace phases with different effective values of α. However, as we see, it is not the case. Together with our consideration in the end of Appendix B where we prove that there is no necklace-cloverleaf phase coexistence in the system, this leads us to the conclusion that the homogenous phase in the system under consideration is always more stable (i.e., has lower free energy) than two phases.

IV. IDEAL LOOP WEIGHTS

In this section we analyze the exact solution of Eqs. (4) and (5) with the loop weights defined by Eqs. (9) and with the propagator of a free chain given by Eq. (7). Substituting Eqs. (7) and (9) into Eq. (5) one gets a nonlinear integral equation which is (see [5,15]) solvable by means of integration over $\mathbf{q}$. One gets

$$G(s,\upsilon,\mathbf{q}) = [\exp(\mathbf{q}^2) - X(s,\upsilon)]^{-1}, \tag{19}$$

where $X(s,\upsilon)$ satisfies the equation

$$(4\pi)^{3/2}X(X-s) = \upsilon s^2 \text{Li}_{3/2}(X) \tag{20}$$

and $\text{Li}_{3/2}(X)$ is a polylogarithm function of order 3/2. Equation (20) always has a solution for small enough s. When s is increasing, Eq. (20) loses the solution at the point defined by the derivative of Eq. (20) with respect to X:

$$(4\pi)^{3/2}X(2X-s) = \upsilon s^2 \text{Li}_{1/2}(X). \tag{21}$$

Let us note that at this point the solution $X(s,\upsilon)$ [and, therefore $G(s,\upsilon,\mathbf{q})$] has a square-root singularity similar to that of Eq. (13). Substituting Eq. (19) together with the first equation of Eqs. (9) into (4), one arrives at the partition function of a diblock copolymer with ideal loop weights:

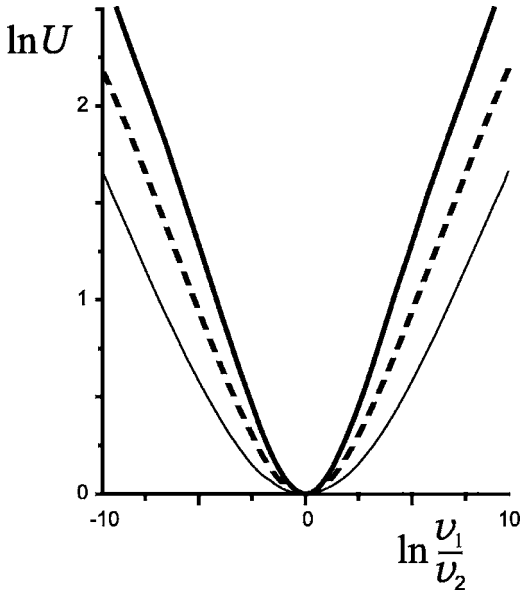

FIG. 7. Phase diagram of the cloverleaf-necklace phase transition with ideal loop weights. Thin, dashed, and thick lines correspond to the product v_1v_2= 100, 10, and 1, respectively.

$$A_2(s_1,s_2) = \frac{1}{(2\pi)^3}\int G_2(s_1,s_2,\mathbf{q})d\mathbf{q} = \left(\frac{X(s_1,v_1)-X(s_2,v_2)}{\dfrac{X(s_1,v_1)-s_1}{v_1s_1^2}-\dfrac{X(s_2,v_2)-s_2}{v_2s_2^2}} - us_1s_2\right)^{-1}. \tag{22}$$

Equation (22) has a simple pole singularity at the point when the whole expression in brackets on the RHS equals zero. Note that there is no singularity due to the denominator of the ratio inside the brackets, since the nominator and denominator always tend to zero simultaneously, while the ratio as a whole stays finite. Thus, one sees again that the conformation of the block copolymer under discussion is governed by the interplay between a square-root singularity (corresponding to the cloverleaf phase) and a simple pole (corresponding to the necklace phase).

Similarly to the previous section, we sketch here the results for the free energy and the transitions in the system, putting all the details of computations into Appendix C.

The transition point now is given by

$$U = \frac{u}{\sqrt{v_1v_2}} = \frac{\sqrt{v_1v_2}(X_1^*-X_2^*)}{v_2s_2^*[X_1^*(s_1^*)^{-1}-1]-v_1s_1^*[X_2^*(s_2^*)^{-1}-1]}, \tag{23}$$

where (s_1^*,X_1^*) and (s_2^*,X_2^*) are common solutions of Eqs. (20) and (21) for $v=v_1$ and v_2, respectively. In Fig. 7 we plot the phase transition line $U_{tr}(v_1,v_2)$ as a function of v_1/v_2 for different values of the mean homopolymer association constant $\sqrt{v_1v_2}$. Note that in this case the dependence $U_{tr}(v_1/v_2)$ is more rich than in the no-loop-weight case, where the transition point was at $U_{tr}=\frac{u_{tr}}{\sqrt{v_1v_2}}=1$ independently of the ratio of

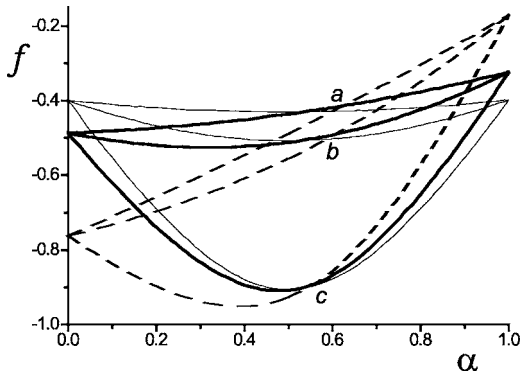

FIG. 8. Free energy of the system with ideal loop weights as a function of α. Thin, thick, and dashed lines correspond to $v_1=v_2=14.1$; $v_1=10$, $v_2=20$; and $v_1=4$, $v_2=50$, respectively. The curves corresponding to $u=40$, $u=80$, and $u=400$ are marked by a, b, and c, respectively. Note that $u=40$ (curves a) correspond to a point in the necklace phase just above the transition. In the cloverleaf phase [see Eq. (C2)] the free energy is linear in α.

the homopolymer association constants. In particular, this nontrivial dependence means that there is the possibility of a reentrant cloverleaf-necklace-cloverleaf transition in the system under discussion. Indeed, it is natural to assume (compare to [19–21]) that the association constants depend on temperature in the usual Boltzmann way:

$$u = u_0\exp[-\epsilon_0/(k_BT)], \quad v_{1,2} = v_0^{1,2}\exp[-\epsilon_{1,2}/(k_BT)]. \tag{24}$$

This means that the change of temperature in any particular system with given chemical structure and given values of u_0, $v_0^{1,2}$, and $\epsilon_{0,1,2}$ corresponds to a movement along some straight line on the log-log plot in Fig. 7. We easily see that such a straight line can easily have at least two intersections with the phase transition line thus giving rise to the aforementioned reentrant transition.[1]

Moreover, the free energy of the system with ideal loop weights reads

$$f = \alpha \ln s_1^a + (1-\alpha)\ln s_2^a, \tag{25}$$

where in the cloverleaf phase the acting singularity values $s_{1,2}^a$ are equal to the $s_{1,2}^*$ defined in Eq. (23), while in the necklace phase they are the solutions of the set of equations (20), (C1), and (C7) lying within the rectangle $0<s_1^a \le s_1^*; 0<s_2^a \le s_2^*$.

In Fig. 8 we plot this free energy f as a function of α for several different values of v_1, v_2, and u. One can notice the similarity between this plot and Fig. 6(b). The curves are once again always convex, so there is once again no internal phase separation in the system. Moreover, one can easily check that the necklace-cloverleaf transition is again of second order. Thus we see that the difference between no-loop-

[1]Note that, generally speaking, the mean homopolymer constant $\sqrt{v_1v_2}$ is changing itself when the temperature changes. However, this change only slightly perturbs the transition curve and does not influence the aforementioned result.

weight and ideal-loop-weight cases is mainly qualitative, except the peculiarities in the behavior of the transition point $U_{tr}(v_1,v_2)$ and the possibility of the reentrant transition in the latter case.

It may be instructive to mention here that recently the possibility of internal phase separation has been studied [10] in a somewhat similar system. Indeed, expressing the result in the notations of current paper, the authors of [10] have shown that in the special case of $v_1=v_2=0$ the internal phase separation is possible for the loop weight exponent $\alpha>2$. We expect this result to hold in the more general case of nonzero constants $v_{1,2}$ but a rigorous proof of this fact needs some further investigation.

V. CONCLUSIONS

We have shown that a diblock copolymer capable of forming RNA-like structures due to reversible bond formation can exhibit a phase transition between two different thermodynamically equilibrium states, a "necklace" with roughness exponent $\gamma_{nc}=1$, and a "cloverleaf" with $\gamma_{rb}=1/2$. The transition is governed by the values of equilibrium association constants and the weights of loop formation in the system.

We have demonstrated that a cloverleaf-necklace transition in the thermodynamic limit is a second-order phase transition both in the case of no ideal loop weight and ideal loop weights (in these cases we have calculated the free energy of the system exactly). One can expect that allowing larger penalties for loop formation, the new phase, that of a coil, may arise in the system, similarly to what predicted in [16] and the order of the necklace-cloverleaf transition itself may change.

The interesting feature of the systems under discussion in the case of ideal loop weights is the possibility of a reentrant necklace-cloverleaf transition for some particular values of parameters determining the temperature dependence of the association constants.

Moreover, the proposed approach allows to calculate the mean free energy of a frozen ensemble of the diblock RNA-like copolymers with some given distribution of bond lengths.

It is worth noting that the mean-field thermodynamics of a globule formed by saturating bonds is investigated in the literature only in the homopolymer [11] and random heteropolymer [18] cases. It may be instructive to study a mean-field globule formed by a diblock copolymer without any constraint on the topology of the bond clusters (i.e., with "pseudoknots" allowed) and compare the results with those obtained in this paper. The corresponding work is currently in progress and will be published elsewhere [22,23].

Finally, we would like to emphasize that the method of calculating the partition function of a diblock copolymer based on Eqs. (4) and (5) [or Eqs. (10) in the case of no loop weights] can be straightforwardly generalized to the case of alternating polyblock RNA-like copolymers of type $A_{t_1}B_{t_2}A_{t_3}B_{t_4}\cdots A_{t_{k-1}}B_{t_k}$ with all lengths of blocks being different. This situation models a particular example of an alternating heteropolymer RNA-like molecule with quenched primary structure defined by the sequence of length $\{t_1,t_2,\ldots,t_k\}$. For example, the case of a four-block copolymer seems to be of particular interest, as in this case one can expect equilibrium crosslike conformations reminiscent of those of tRNA.

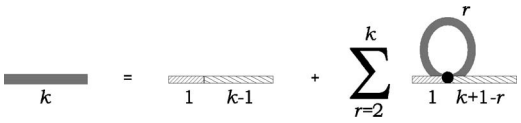

FIG. 9. Diagrammatic structure of Eq. (26).

The corresponding general Dyson-type equation for the generation function $G_n(s_1,s_2,\ldots,s_n)$, where

$$G_n(s_1,s_2,\ldots,s_k)=\sum_{t_1=0}^{\infty}\sum_{t_2=0}^{\infty}\cdots\sum_{t_k=0}^{\infty}G_k(t_1,t_2,\ldots,t_k)s_1^{t_1}s_2^{t_2}\cdots s_k^{t_k},$$

takes the following form [compare to Eq. (4)]:

$$\begin{aligned}G_k(s_1,\ldots,s_k)&=G_1(s_1)G_{k-1}(s_2,\ldots,s_k)\\&+\sum_{r=2}^{k}u_{1,r}s_1s_rG_1(s_1)G_r(s_1,\ldots,s_r)\\&\times G_{k+1-r}(s_r,\ldots,s_k),\end{aligned}\tag{26}$$

which has the diagrammatic structure shown in Fig. 9 [compare to Fig. 4(a)]. Equation (5), as well as it diagrammatic structure shown in Fig. 4(b), remains without changes.

We believe that Eq. (26) can serve as a basis for a renormalization group procedure, thus connecting the exact partition function $G(4)$ of an four-block copolymer with an approximative expression for $G(4)$ constructed of two diblock copolymers $G(2)$—i.e., $G(4)\approx G(2)\times G(2)$. It is obvious that if all coupling constants are identical, then we have a RNA-like homopolymer and the last relation is exact. However, if the A-A, B-B, and A-B constants are slightly different, we should have $G(4)-G(2)\times G(2)=\epsilon(u,v_1,v_2)$. Considering ϵ as a perturbation we may look how ϵ is transformed under the standard coarse-graining procedure. This work is in progress, and the corresponding results will be reported elsewhere [24].

ACKNOWLEDGMENTS

The authors are grateful to I.E. Erukhimovich and M. Mezard for valuable discussions. M.V.T. gratefully acknowledges the warm hospitality he received throughout his stay at the LPTMS where this work was done. The work is partially supported by Grant No. ACI-NIM-2004-243 "Nouvelles Interfaces des Mathématiques" (France). We are grateful to S. Lando and P. Krapivsky for comments on mathematical procedures and for providing us relevant references on multivariate complex analysis.

APPENDIX A: SINGULARITIES OF THE GENERATING FUNCTION $G(s_1,s_2)$

In this appendix we address the procedure of calculating the asymptotic behavior of the coefficients of a generating function [compare to Eq. (3)]

$$G(s_1,s_2)=\sum_{n,m=0}^{\infty} a_{m,n}s_1^m s_2^n \tag{A1}$$

in the limit $m+n=N\to\infty$ and $m/(m+n)\equiv\alpha=$const. The procedure suggested below is similar to that developed in [25], and we advise the reader to address this reference for accurate mathematical proofs. However, the algorithm under discussion is not in common use, so it seems instructive to give a short explanation of the physical level of rigor. Note that we pay attention only to the exponentially increasing multipliers in the asymptotic behavior of the free energy, thus neglecting the logarithmic corrections.

First of all, let $G(s_1,s_2)$ defined by Eq. (A1) be an analytic function of s_1,s_2 in the circle $\sqrt{|s_1|^2+|s_2|^2}<\epsilon$ for some $\epsilon>0$ near the origin $s_1=s_2=0$. Assume also, as is usual in statistical mechanics, the dominant singularities of G with respect to each of the variables (for fixed another one) to be real and positive.

Now, it is known that to calculate a function $G_1^*(s)$ defined as

$$G_1^*(s)=\sum_{i=0}^{\infty} a_{i,i}s^{2i}, \tag{A2}$$

whose dominant singularity defines the asymptotic of $a_{i,i}$ at $i\to\infty$, one can use the following trick (see, for example, [26]). Consider $G(sz,\frac{s}{z})$ as a function of a variable z. If s is small enough, this function is analytic with respect to z inside some ring where it can be represented by the Laurent series

$$G\left(sz,\frac{s}{z}\right)=\sum_{k=-\infty}^{\infty} b_k(s)z^k. \tag{A3}$$

Note now that the coefficient $b_0(s)$ in this expansion is exactly equal to the desired function $G_1^*(s)$, and therefore this function can be evaluated using the Cauchy theorem:

$$G_1^*(s)=\frac{1}{2\pi i}\oint_C G\left(sz,\frac{s}{z}\right)\frac{dz}{z}, \tag{A4}$$

where the curve C encloses the point $z=0$ exactly once and belongs to the region of analyticity of the function $G(sz,\frac{s}{z})$.

This consideration can be easily generalized in the following way. Suppose, we are interested in calculating the function

$$G^*\left(s,\frac{p}{q}\right)=\sum_{j=0}^{\infty} a_{jp,jq}s^{j(p+q)}, \tag{A5}$$

where p and q are two mutually simple natural numbers. To perform the summation in Eq. (A5), we consider a function $G(sz^q,sz^{-p})$ and take the coefficient of order 0 in its Laurent series:

$$G^*\left(s,\frac{p}{q}\right)=\frac{1}{2\pi i}\oint_C G\left(sz^q,\frac{s}{z^p}\right)\frac{dz}{z}, \tag{A6}$$

where the integration curve belongs again to the region of analyticity of $G(sz^q,sz^{-p})$.

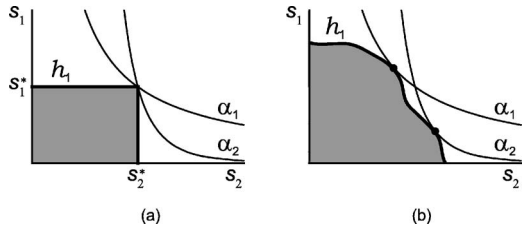

FIG. 10. Examples of possible domains of analyticity. The curves (A7) corresponding to different values of α ($\alpha_1<\alpha_2$) are also shown. The touching points may (b) or may not (a) depend on α.

This formalism allows one in principle to calculate the generating function $G^*(s,\frac{p}{q})$ for any given ratio of block lengths of the copolymer under discussion. However, on the one hand, the explicit calculation of G^* is still rather cumbersome and, on the other hand, it is actually not needed. What we are really interested in is the *asymptotic* behavior of the coefficients of G^* or, in other words, the position and the nature of its dominant singularity (or singularities). The task of finding these singularities can be accomplished without explicit knowledge of G^*.

To do that, let us first introduce some additional definitions. Let H be the 3D surface of singular points of G (it may consist of several branches, of course) in four-dimensional space (Re s_1, Re s_2, Im s_1, Im s_2) and h be the intersection of H with the "real positive plane" [i.e., the quarter-plane of (s_1,s_2) where $\mathrm{Arg}(s_1)=0$ and $\mathrm{Arg}(s_2)=0$]. Now, the surface H separates the space (s_1,s_2) into several domains. We pay particular attention to the shape of the domain containing the origin $(s_1=s_2=0)$. Let us call its edge H_1 ($H_1\subset H$). It follows immediately from the condition of single-variable dominant singularities to be real and positive that for any given absolute value $|s_1|,|s_2|$, the point of the edge of the domain H_1 closest to the origin is the point with $\mathrm{Arg}(s_1)=\mathrm{Arg}(s_2)=0$ and, thus, it belongs to the curve h. In what follows we can restrict ourselves therefore to consideration of real and positive singularities of G.

In a (s_1,s_2) quarter-plane there exists some domain of analyticity of G which includes the origin as shown in Fig. 10; its edge is designated by h_1.

Let us fix now some "direction" $\alpha=\frac{p}{p+q}$. We are interested in calculating the generating function $G^*(s,\alpha)$ corresponding to this particular direction as defined by Eq. (A5). We are thus to integrate the function $G(sz^q,sz^{-p})$ as prescribed in Eq. (A6). By fixing some particular value of s, we define therefore a connection between the values of $s_1=sz^q$ and $s_2=sz^{-p}$ in the integrand as follows:

$$s_1^p s_2^q=(s^p z^{pq})(s^q z^{-pq})=s^{p+q}(\text{or } s_1^\alpha s_2^{1-\alpha}=s). \tag{A7}$$

Obviously, the corresponding line on the (s_1,s_2) plane always intersects with our domain of analyticity if s is small enough. Therefore, for small s there is no problem in choosing the appropriate circle of integration $|z|=$const. One should just choose $|z|$ in such a way that the point $(s|z|^q,s|z|^{-p})$ belongs to the domain of analyticity. However, when one increases s, the intersection between this domain

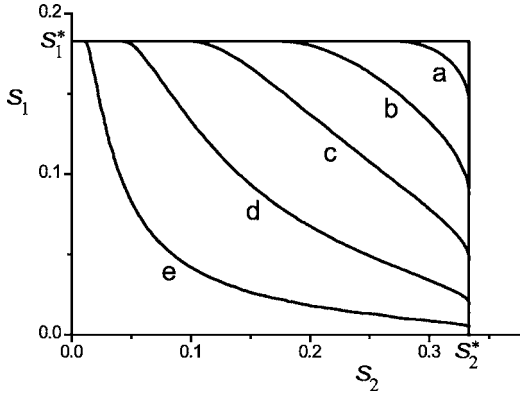

FIG. 11. Examples of the analyticity domains of Eq. (12) for $v_1=1, v_2=5$. Rectangular domain with borders $s_{1,2}^*=(1+2\sqrt{v_{1,2}})^{-1}$ corresponds to values of u below the necklace-cloverleaf transition (i.e., $u<u_{\mathrm{tr}}=\sqrt{5}$). The curves marked by a, b, c, d, and e are the pole lines corresponding to $u=5$, 10, 20, 50, and 200, respectively.

and the line (A7) may disappear at some point. In terms of the z plane, where the integration prescribed in Eq. (A6) is to be accomplished, it means that two singularities of the integrand, one being inside the integration curve and another outside, approach each other and finally merge. Note that they do merge indeed rather then just have equal absolute values since we have assumed these singularities to be real and positive. Thus, we see that the integral in Eq. (A6) has a singularity exactly at this point.

Thus, summarizing what was said above, we suggest the following procedure to find the singularities of the function $G^*(s,\alpha)$.

(i) Find the lines of singular points h of the initial function $G(s_1,s_2)$ in the real plane.

(ii) Define the domain of analyticity D which includes the origin $s_1=s_2=0$ and find its edge h_1.

(iii) For each value of $0<\alpha<1$ find the value of s for which the intersection between D and the curve (A7) ceases to exist for the first time. If h_1 is a smooth curve, it means just finding a common tangent line of h_1 and (A7); if there is a finite number of peculiar points on it (say, kinks or intersections of the branches), it includes also finding the conditions that (A7) intersects these particular points.

APPENDIX B: CALCULATION OF THE FREE ENERGY IN THE NO-LOOP-WEIGHT LIMIT

In this section we apply the strategy presented in Appendix A for calculating the free energy corresponding to the grand partition function (12)—i.e., that of the diblock copolymer with no loop weights.

As prescribed by the suggested procedure, we should first find the domain of analyticity which includes the point $s_1=s_2=0$. In Fig. 11 we have drawn these regions in the (s_1,s_2) plane for some particular sets of v_1,v_2 and different values of u. The edges of the region correspond to singularity lines defined by Eqs. (13) and (14). The square root singularities correspond to some straight lines $s_1=|s_1^{\mathrm{sqr}}|$ and $s_2=|s_2^{\mathrm{sqr}}|$ [where one should choose the solution of Eq. (13) with the lowest absolute value], so for any (s_1,s_2) with absolute values the rectangle $(0\le|s_1|<|s_1^{\mathrm{sqr}}|, 0\le|s_2|<|s_2^{\mathrm{sqr}}|)$ and the square roots in Eq. (12) are both analytical functions of their respective variables. Moreover, the pole singularity (14) defines some other curve in the (s_1,s_2) plane. This curve can either have no branch within the rectangle described above, and then this rectangle is the desired domain of analyticity itself, or it can intersect the rectangle (as shown by lines a—e in Fig. 11), and then one should take the part below and to the left of it as the desired domain of analyticity.

The further calculations can be much simplified by introducing the new variables

$$t_{1,2}=2\sqrt{v_{1,2}}\frac{s_{1,2}}{1-s_{1,2}}.$$

The square-root singularities now correspond just to $t_{1,2}=1$. The equation for the pole singularity (14) can be rewritten as follows:

$$U=\frac{u}{\sqrt{v_1v_2}}=\frac{t_1t_2}{(1-\sqrt{1-t_1^2})(1-\sqrt{1-t_2^2})}. \tag{B1}$$

Now, since the function

$$f(t)=\frac{t}{1-\sqrt{1-t^2}}$$

decreases monotonously from $+\infty$ to 1 with t increasing from 0 to 1, it is obvious that the solution of Eq. (B1) in the square $0\le t_1\le 1$, $0\le t_2\le 1$ exists for and only for $U\ge 1$—i.e., for $u\ge u_{\mathrm{tr}}=\sqrt{v_1v_2}$. Moreover, we easily see that for $u<u_{\mathrm{tr}}$ when the region of analyticity is just the rectangle shown in Fig. 8, the corridor of analyticity for the integrand in the integral expression for G^*, Eq. (A6), disappears exactly at the point when the line (A7) passes through the point $(s_1^{\mathrm{sqr}},s_2^{\mathrm{sqr}})$:

$$s_{\mathrm{sqr}}^*=(s_1^{\mathrm{sqr}})^{1-\alpha}(s_2^{\mathrm{sqr}})^{\alpha}. \tag{B2}$$

The free energy per link of the chain is given therefore by

$$f_{\mathrm{clov}}=\ln s_{\mathrm{sqr}}^*=-\alpha\ln(1+2\sqrt{v_1})-(1-\alpha)\ln(1+2\sqrt{v_2}), \tag{B3}$$

and the singularity of G^* is in this case of square-root nature which justifies our assertion that for $u<u_{\mathrm{tr}}$ the polymer is in the cloverleaf state.

In the case when $u>u_{\mathrm{tr}}$ the region of analyticity of Eq. (A6) disappears when the curves (A7) and (B1) have a common tangent line. To proceed further let us introduce the new variables $\phi_{1,2}$ as follows:

$$\phi_i=\arcsin t_i,\quad s_i=\frac{\sin\phi_i}{\sin\phi_1+2\sqrt{v_i}},\quad i=1,2. \tag{B4}$$

Equations (A7) and (B1) can be now rewritten as follows:

$$U^{-1}=\tan\frac{\phi_1}{2}\tan\frac{\phi_2}{2} \tag{B5}$$

and

$$s = (1 + 2\sqrt{v_1}/\sin\phi_1)^{-\alpha}(1 + 2\sqrt{v_2}/\sin\phi_2)^{\alpha-1}, \quad \text{(B6)}$$

where $\phi_{1,2} \in [0, \pi/2]$.

The condition to have the common tangent line gives rise to the additional equation

$$\frac{\sin\phi_1}{\sin\phi_2} = \frac{1-\alpha}{\alpha}\sqrt{\frac{v_2}{v_2}}\frac{\sin\phi_1\cos\phi_2(\sin\phi_1 + 2\sqrt{v_1})}{\sin\phi_2\cos\phi_1(\sin\phi_2 + 2\sqrt{v_2})}. \quad \text{(B7)}$$

One can easily exclude two of the three unknown variables (s, ϕ_1, ϕ_2) from Eqs. (B6) and (B7) and arrive finally at the following equation:

$$4\sqrt{v_1 v_2}\sinh(2\xi) + 4(1-2\alpha)\sqrt{v_1 v_2}\sinh(2\vartheta) - \alpha\sqrt{v_1}\sinh(\vartheta - \xi) + (1-\alpha)\sqrt{v_2}\sinh(\vartheta + \xi) = 0 \quad \text{(B8)}$$

for a new variable $\xi = -\frac{1}{2}\ln U - \ln\tan\frac{\phi_1}{2}$ with $\vartheta = -\frac{1}{2}\ln U$ and additional condition $-\vartheta \le \xi \le \vartheta$. It is easy to check that for any $\vartheta > 0$ (i.e., for any $u > u_{tr}$) the left-hand side of this equation is an increasing function of ξ and changes sign inside the interval $(-\vartheta, \vartheta)$. Equation (B8) has therefore exactly one solution with respect to ξ. The fact that this solution always stays within the interval $(-\vartheta, \vartheta)$ and never reaches its limits means that the dominate singularity in the necklace phase is always (independently of u and α) a regular point of the pole line, not the intersection of the pole and square-root line. This means that there is no coexistence of the necklace and cloverleaf phases in the system for any u and α. Moreover, after substituting this solution ξ^* into the original equation for the singularity of s, Eq. (B6), we obtain the final expression for the free energy in this (necklace) phase:

$$f_{neck} = \ln s^*_{pole} = -\alpha\ln[1 + 2\sqrt{v_1}\cosh(\vartheta - \xi^*)] - (1-\alpha)\ln[1 + 2\sqrt{v_2}\cosh(\vartheta + \xi^*)]. \quad \text{(B9)}$$

APPENDIX C: CALCULATION OF THE FREE ENERGY IN THE IDEAL-LOOP-WEIGHT LIMIT

In this section we apply the procedure suggested in the Appendix A to the calculation of the free energy of the block copolymer with "ideal" loop weights. To that end we investigate the singularities of the grand partition function Eq. (22) with X_1 and X_2 defined by Eq. (20).

As we have mentioned already, there are two possible types of singularities of this partition function. The first type corresponds to the singularities of the homopolymer partition function, is of square-root nature, and corresponds to the solution of Eqs. (20) and (21). The second one is a pole singularity corresponding to the zero of the denominator in the integrand on the RHS of Eq. (22).

$$(X_1 - X_2)\left(\frac{X_1 - s_1}{v_1 s_1^2} - \frac{X_2 - s_2}{v_2 s_2^2}\right)^{-1} - u s_1 s_2 = 0, \quad \text{(C1)}$$

where s_i and X_i are connected via Eq. (20). Contrary to the first one, this singularity is of collective nature; i.e., it depends simultaneously on the values of both s_1 and s_2, and it also depends on the value of u.

Moreover, in full analogy with the no-loop-weight case, the square-root singularities define for small values of u the region of analyticity of a rectangular shape in the (s_1, s_2) plane. In this case (i.e., below the necklace-cloverleaf transition) the acting singularity point is, independently of α, just the intersection point of the two square-root singularity lines and the free energy equals

$$f_{clov}(v_{1,2}, \alpha) = -\alpha\ln s_1^* - (1-\alpha)\ln s_2^* \quad \text{(C2)}$$

[compare to Eq. (B3)], where $s^*_{1,2}$ are the solutions of Eqs. (20) and (21) with $v = v_1, v_2$.

Now, with increasing u the line of poles enters the rectangle of analyticity. It first appears at the intersection point of two square-root singularity lines, thus giving rise to the following equation for the necklace-cloverleaf transition point:

$$(X_1^* - X_2^*)\left(\frac{X_1^* - s_1^*}{v_1 (s_1^*)^2} - \frac{X_2^* - s_2^*}{v_2 (s_2^*)^2}\right)^{-1} = u_{tr} s_1^* s_2^*, \quad \text{(C3)}$$

where we have inserted the solutions of the set (20) and (21) into (C1). This equation leads after some algebra to Eq. (23).

To investigate properly the phase above the transition point (i.e., for $u > u_{tr}$) we have (i) to find the common tangent lines of the pole line (C1) and the line (A7) and (ii) to check whether there is a possibility for the acting singularity to lie at the intersection of the pole line and the square-root singularity line—i.e., to check whether there is a kink at this intersection. To do that we should first of all find the derivative of the pole line. A direct calculation gives

$$ds_1\left[X_1'\left(1 - \frac{u s_2}{v_1 s_1}\right) + u s_2\left(\frac{X_2 - s_2}{v_2 s_2^2} + \frac{X_1}{v_1 s_1^2}\right)\right] ds_2\left[X_2'\left(1 - \frac{u s_1}{v_2 s_2}\right) + u s_1\left(\frac{X_1 - s_1}{v_1 s_1^2} + \frac{X_2}{v_2 s_2^2}\right)\right], \quad \text{(C4)}$$

where $X'_{1,2}$ are the derivatives of $X_{1,2}$ with respect to $s_{1,2}$:

$$X_i' = \frac{dX_i}{ds_i} = X_i\frac{(4\pi)^{3/2}X_i + 2v_i s_i \mathrm{Li}_{3/2}(X_i)}{(4\pi)^{3/2}X_i(2X_i - s_i) - v_i s_i^2 \mathrm{Li}_{1/2}(X_i)}, \quad i = 1,2, \quad \text{(C5)}$$

and the auxiliary conditions (20) are taken into account.

Now, let us notice that along the square-root singularity line (say, for $s_1=s_1^*$), the corresponding derivative (X_1' in this case) diverges, while, if we are not approaching the point (s_1^*, s_2^*), all the other coefficients in Eq. (C4) remain finite. We conclude immediately that to comply with this equation, the derivative $\frac{ds_2}{ds_1}$ along the pole line should also diverge in the vicinity of the point $s_1=s_1^*$. Thus, the pole line has a vertical asymptote at $s_1=s_1^*$ and there is no kink at the intersection of the pole and square root lines (the case $s_2=s_2^*$ is absolutely similar). We see therefore that the acting singularity never lies at this intersection. This proves, similarly to the no-loop-weight case, that there is no possibility of a intramolecular necklace-cloverleaf phase separation in the system.

To calculate the free energy of the necklace phase as a function of the composition variable α we should now equalize the pole line derivative (C4) and the derivative of the line (A7) which is given by

$$\frac{ds_1}{ds_2} = -\frac{1-\alpha}{\alpha}\frac{s_1}{s_2}. \tag{C6}$$

Collecting Eqs. (C4) and (C6) together, we get

$$\alpha s_2\left[X_2'\left(1-\frac{us_1}{v_2s_2}\right)+us_1\left(\frac{X_1-s_1}{v_1s_1^2}+\frac{X_2}{v_2s_2^2}\right)\right] = (\alpha-1)s_1\left[X_1'\left(1-\frac{us_2}{v_1s_1}\right)+us_2\left(\frac{X_2-s_2}{v_2s_2^2}+\frac{X_1}{v_1s_1^2}\right)\right]. \tag{C7}$$

This equation, together with Eqs. (20) and (C1), completes the set which defines four unknown variables s_1, s_2, X_1, and X_2 at the pole singularity point.

The free energy is again

$$f_{\text{neck}} = -\alpha \ln s_1 - (1-\alpha)\ln s_2. \tag{C8}$$

It is calculated numerically, and the corresponding plots are shown in Fig. 8.

[1] S. F. Edwards, Proc. Phys. Soc. London **85**, 613 (1965).
[2] I. M. Lifshits, Sov. Phys. JETP **28**, 1280 (1969).
[3] P. de Gennes, Biopolymers **6**, 715 (1968).
[4] V. Pande, A. Grosberg, and T. Tanaka, Rev. Mod. Phys. **72**, 259 (2000).
[5] A. Montanari and M. Mezard, Phys. Rev. Lett. **86**, 2178 (2001).
[6] R. Bundschuh and T. Hwa, Phys. Rev. E **65**, 031903 (2002).
[7] F. Krzakala, M. Mezard, and M. Mueller, Europhys. Lett. **57**, 752 (2002); M. Mueller, F. Krzkala, and M. Mezard, Eur. Phys. J. E **9**, 67 (2002).
[8] M. Lassig and K. J. Wiese, Phys. Rev. Lett. **96**, 228101 (2006).
[9] V. Guttal and R. Bundschuh, Phys. Rev. Lett. **96**, 018105 (2006).
[10] R. A. Neher and U. Gerland, Phys. Rev. E **73**, 030902(R) (2006).
[11] I. M. Lishits, A. Yu. Grosberg, and A. R. Khokhlov, Sov. Phys. JETP **71**, 1634 (1976).
[12] A. N. Semenov and M. Rubinstein, Macromolecules **31**, 1373 (1998); I. Ya. Erukhimovich, M. V. Thamm, and A. V. Ermoshkin, *ibid.* **34**, 5653 (2001), and references therein.
[13] A. Yu. Grosberg and A. R. Khokhlov, *Statistical Physics of Macromolecules* (AIP, New York, 1994).
[14] Y. Kafri, D. Mukamel, and L. Peliti, Phys. Rev. Lett. **85**, 4988 (2000).
[15] I. Ya. Erukhimovich, Vysokomol. Soedin., Ser. B **20**, 10 (1978).
[16] M. Müller, Phys. Rev. E **67**, 021914 (2003).
[17] H. Orland and A. Zee, Nucl. Phys. B **620**, 456 (2002).
[18] A. M. Gutin, A. Yu. Grosberg, and E. I. Shakhnovich, J. Phys. A **26**, 1037 (1993).
[19] I. Erukhimovich and A. V. Ermoshkin, J. Chem. Phys. **116**, 368 (2002).
[20] M. V. Tamm and I. Ya. Erukhimovich, Polym. Sci., Ser. A Ser. B **44**, 320 (2002).
[21] M. V. Thamm and I. Ya. Erukhimovich, J. Chem. Phys. **119**, 2720 (2003).
[22] L. D. Landau and I. M. Lifshits, *Statistical Physics*, 2nd ed., Course of Theoretical Physics, Vol. 5 (Pergamon, New York, 1968), pp. 1.
[23] M. Tamm, I. Erukhimovich, and S. Nechaev (unpublished).
[24] M. Tamm and S. Nechaev (unpublished).
[25] R. Pemantle and M. C. Wilson, e-print math.CO/0512548.
[26] S. Lando, *Lectures on Generating Function* (AMS, Providence, RI, 2003).